THE INTERNATIONAL SOCIETY OF PRECISION AGRICULTURE PRESENTS THE
13th INTERNATIONAL CONFERENCE ON
PRECISION AGRICULTURE
July 31-August 4, 2016 • St. Louis, Missouri USA

# Field Potential Soil Variability Index to Identify Precision Agriculture Opportunity

**Christopher Bobryk, Matt Yost, and Newell Kitchen**

USDA–Agriculture Research Service, Cropping Systems and Water Quality Research Unit

269 Agriculture Engineering Building, University of Missouri, Columbia, MO 65211

**A paper from the Proceedings of the**
**13th International Conference on Precision Agriculture**
**July 31 – August 4, 2016**
**St. Louis, Missouri, USA**

***Abstract.*** *Precision agriculture (PA) technologies used for identifying and managing within-field variability are not widely used despite decades of advancement. Technological innovations in agronomic tools, such as canopy reflectance or electrical conductivity sensors, have created opportunities to achieve a greater understanding of within-field variability. However, many are hesitant to adopt PA because uncertainty exists about field-specific performance or the potential return on investment. These concerns could be better addressed by understanding where variability in soil physical and chemical properties may have the greatest effect on crop responses to inputs, such as nitrogen fertilizer. Therefore, identifying fields that exhibit the most variation in soil characteristics (e.g. clay and organic matter content) and developing an indicator of variation that has the potential to affect crop responses to inputs could greatly advance PA adoption and use. The objectives of this research were to: 1) quantify the amount of potential soil variability over a large region, 2) generate an index that numerically identified fields that exhibit degrees of field variability, and 3) evaluate spatial clustering of variability over the region. This analysis focused on soil variability in agricultural fields across Missouri, USA. We calculated a variability index (VI) for clay and organic matter content at two depth increments (0-30 and 0-120 cm) using soil information from the National Resources Conservation Service's (NRCS) Soil Survey Geographic database (SSURGO). Ranges in VI for clay at the two depth increments were 1-82 and 1-91 with an average of 2.4 and 2.2, respectively. Organic matter VI averaged 2.0 and 2.3 for the two increments with narrower ranges from 1-42 and 1-29, accordingly. Significant high clay VI clusters at both increments were observed mostly along the Missouri River floodplain and across southeastern Missouri along the Mississippi River. High organic matter VI clusters exhibited similar distributions along the Missouri and Mississippi River floodplains; however, significant clusters of low organic matter VI values occurred within the Central Claypan and Southern Mississippi River Alluvium major land resource areas. Output from this research could be used as a decision support tool to aide suppliers and practitioners in determining the greatest opportunities to implement PA.*

***Keywords: Soil variability, Clay, Organic matter, Missouri, Clustering, Decision-support***

The authors are solely responsible for the content of this paper, which is not a refereed publication. Citation of this work should state that it is from the Proceedings of the 13th International Conference on Precision Agriculture. EXAMPLE: Lastname, A. B. & Coauthor, C. D. (2016). Title of paper. In Proceedings of the 13th International Conference on Precision Agriculture (unpaginated, online). Monticello, IL: International Society of Precision Agriculture.

## Introduction

Sophisticated management of cropland is becoming critically important to ensure that agricultural production potential can meet future food demands of the estimated 9.6 billion people by 2050 (Godfray et al., 2010; U.S. Census Bureau, 2015). The agriculture industry is under pressure to implement more efficient strategies to cope with an increasing world population and dynamic commodity prices, while abating environmental degradation (Castle et al., 2016; Tey et al., 2012). Precision agriculture (PA) is an optimal solution to meet these current and future challenges by implementing reduced or targeted placement of agronomic inputs (e.g. fertilizer). Economic or production considerations often are the primary drivers for farmers to invest in PA technologies (McLoud et al., 2007), and new tools specifically aimed at increasing production efficiency are continually being developed; however, a lag exists between technological advancements and PA adoption.

The potential impacts of PA are broad and include input cost reduction, improved management zone selection, and enhanced environmental protection. As such, the interest in PA has increased over the last two decades, especially with advancements in technology and reduced costs of PA instrumentation (e.g. machinery, yield monitors, proximal remote sensors, and GPS guidance systems). Even though there are more opportunities now to invest in PA tools that effectively manage cropland at various spatial and temporal scales, few producers risk adopting PA technologies because of profitability concerns (Erickson et al., 2015; Griffin et al., 2004). Consequently, uniform applications of agronomic inputs are routinely implemented for ease and to prevent crop nutritional deficiencies; therefore, avoiding risks for significant crop yield losses (Tey et al., 2012). Understanding spatial variations in key factors relevant to crop input use (e.g. soil type or weather conditions) is an important prerequisite for influencing PA adoption (Karpinski et al., 2015; McBratney et al., 2005). Quantifying a baseline measure representing manageable variability within a production field could help practitioners better understand basic requirements needed for introducing PA tools into general production systems (Mintert et al., 2015).

Spatial patterns in crop productivity are influenced by variation in certain soil conditions found within a field (Cambardella et al., 1994). Numerous studies have investigated different types of indicators to help delineate patterns in soil, crop, and landscape characteristics to generate specific management practices dedicated to improve crop production (Kitchen et al., 2005). For instance, airborne hyperspectral remote sensing data has been used to calculate a vegetation index that used proxies for crop chlorophyll content to predict in-field variations of nitrogen (Haboudane et al., 2002). Additionally, an economic index, based on biophysical characteristics of crops captured from proximal remote sensors, was used as a tool to evaluate production variability (Oliveira et al., 2012). The outcome of this research was an ‘opportunity index’ that proved only useful in single-season assessments. Karpinski et al. (2015) developed another indicator for categorizing soil heterogeneity solely based on winter wheat yield data that helped identify optimal field conditions for applying site-specific crop management; ultimately helping guide producers toward an increased economic return. These studies, and many others (Adamchuk et al., 2004; Haboudane et al., 2002; Kitchen et al., 2005; Rabbi et al., 2014; Stadler et al., 2015), focused primarily on field-scale applications of indicators for variability, which require sophisticated equipment to collect data and generate accurate and precise output. Few studies provide decision-support tools for practitioners or agricultural input suppliers, who are in the nascent stages of deciding whether to adopt PA technologies or simply interested in acquiring important information about broad-scale variations in soils surrounding production areas.

Investing in PA involves many risks, which often are underestimated when variation in costs, yields, or soil conditions are assessed at higher levels of aggregation (e.g. county scales) (Tey et al., 2012; Ziadat et al., 2015). Additionally, there is a lack in accessibility to decision-support frameworks that can easily demonstrate where opportunities exist for using PA to manage crop yield spatial variability. Research is needed that focuses on simple and standardized indicators to support basic decision processes during initial phases of PA adoption (Oliveira et al., 2012). Therefore, developing decision-support tools that help link the need for PA management with site-specific characteristics may help diminish the knowledge gap between understanding field-scale changes in production functions and employing the tools necessary to optimize inputs for different management zones. We suggest here an approach for informing practitioners, producers, suppliers, farming communities, industry, and agronomic scientists about broad-scale measures of spatial patterns of variability in soil properties.

The purpose of this research was to develop a variability index (VI) of clay and organic matter content within individually managed fields across Missouri. Clay and organic matter are two major indicators for soil N availability; therefore, output from this research was intended to be used as a decision-support platform to make better recommendations for identifying fields that could improve fertilizer N management. Building a broad-scale understanding of variations in specific soil properties is important and may lead to further refinement in agricultural management practices; subsequently promoting the efficacy of PA practices. The objectives of this study were to: 1) quantify the amount of clay and organic matter variability by field, 2) generate a landscape map that illustrated the VI by field, and 3) evaluate spatial distributions of high and low clustering patterns of the VI across groups of fields. This last objective would be particularly meaningful for agricultural suppliers that were interested in identifying opportunity areas to target their PA products or services.

## Methods and Materials

### Study Area and Datasets

This study identified soil variability within individual row-crop agricultural fields across Missouri, USA (Fig. 1). Boundaries of crop fields were obtained from the analysis of Yan and Roy (2016) where the authors used automated computational processes to extract boundaries for the conterminous U.S. using time series satellite imagery and edge detection algorithms (Yan et al., 2016). Fields greater than 4 ha were identified as manageable and used for this analysis. Clay and organic matter content data were obtained from the National Resources Conservations Service's (NRCS) Soil Survey Geographic (SSURGO) database. The SSURGO and field boundary datasets were in vector format and intersected with each other to obtain a master dataset containing all soil profile information and associated unique field boundary identification numbers using ArcGIS 10.3.

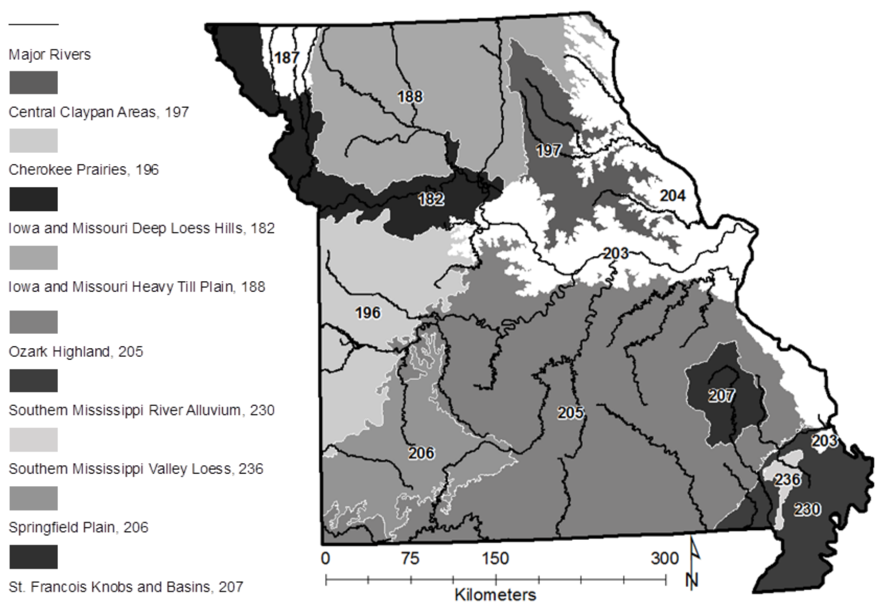


**Fig. 1: Study location (Missouri, USA) with major river systems illustrated. The state is dissected by major land resource area (MLRA) boundaries.**

### Soil Variability Index

The VI was determined at two depth increments (top: 0-30 and profile: 0-120 cm) for each variable (clay and organic matter) by calculating the ratio between the maximum and minimum soil content values from the pool of SSURGO map units for each field. Maximum and minimum values were calculated by first converting all variable percentages into content based on bulk density measures listed within the SSURGO dataset. The second step was to implement a depth-weighted function that sliced the whole profile into 1 cm increments, and the average value per variable was calculated for the top and profile. This function was performed for all SSURGO map units that intersected each field and a single VI value was attributed to individual fields. A constraint was applied to this ratio if either the maximum or minimum value did not represent at least 5% of the total area of the field. If the area of the maximum or minimum content did not match the 5% threshold, we subtracted the required proportion from the second highest or lowest content to match the 5% threshold. The remaining content value was added to the original maximum or minimum value to meet the 5% criterion. All calculations were conducted in R statistical software v3.2.1 (R Core Team, 2015) and the AQP package (Beaudette et al., 2013).

### Clustering Statistic

Anselin Local Moran's I cluster and outlier analysis (Anselin, 1995; Moran, 1950) was used to identify groups of multiple fields of similar VI values. Specifically, a local Moran's I was used to statistically identify, at the 95% confidence level, groups of high-high (HH; similar high VI values across multiple fields) or low-low (LL; similar low VI values across multiple fields) clustering, in addition to identifying anomalies where nearby fields exhibited dissimilarities either as low-high (LH) or high-low (HL). The Local Moran's I statistic was expressed as:

$$I_i = \frac{z_i - \bar{z}}{\sigma^2} \sum_{j=1,\, j \neq i}^{n} \left[ W_{ij} \left( z_j - \bar{z} \right) \right] \tag{1}$$

where $\bar{z}$ represented the mean value of VI ($z$) with the sample number of n determined by a distance threshold limit; $z_i$ was the value of VI at the location $i$; $z_j$ was the VI value at other locations (where $j \neq i$); $\sigma^2$ was the variance of $z$; and $W_{ij}$ can be a weighting factor between $z_i$ and $z_j$ that define predetermined spatio-temporal relationships among features. The spatial relationships of VI were identified in this investigation using an inverse distance squared conceptualization model with no weighted factors between values, and a distance threshold of 25 km. The distance parameter was used to minimize the number of fields identified without neighbors. All clustering analyses were performed using ArcGIS v10.3.

## Results and Discussion

### Soil Variability Index

A total of 176,000 crop production fields were used in the analysis. The VI for clay in the top increment (0-30 cm) had a wide range from 1 – 82 with an average of 2.4 and median of 1.7 (Table 1). The clay VI at the profile increment (0-120 cm) had a slightly broader range from 1 – 91, with a similar average and median VI of 2.2 and 1.5, respectively. The similarity between the median clay VI values at both depths was indicative of a wide-spread distribution of clay content with minimal variation throughout the general profile and across the landscape. The greatest range in clay content variability in both increments was delineated by several major land resource areas (MLRAs) (USDA-NRCS, 2006). The greatest VI values were found within the Deep Loess Hills of northeastern Missouri (along major rivers and tributaries) and along the western portion of the Central Mississippi Valley Wooded Slope areas (Fig. 2). Additionally, the southern Mississippi River Alluvium MLRA exhibited the most diversity in ranges of VI values. These particular MLRA sections possess extreme soil variability due to river system dynamics that drive fluctuations in soil erosion and deposition.

**Table 1: Variance Index descriptive statistics.**

| | Clay | | Organic Matter | |
|---|---|---|---|---|
| **Statistics** | **0-30 cm** | **0-120 cm** | **0-30 cm** | **0-120 cm** |
| Median | 1.67 | 1.50 | 1.59 | 1.80 |
| Mean | 2.36 | 2.18 | 2.00 | 2.29 |
| Variance | 10.34 | 9.45 | 2.33 | 2.34 |
| Standard Deviation | 3.22 | 3.07 | 1.53 | 1.53 |
| Coefficient of Variation | 1.36 | 1.41 | 0.76 | 0.67 |

Organic matter at the top increment ranged from 1 – 42 with an average of 2.0 and median of 1.6 (Table 1). The profile increment of organic matter had the narrowest range from 1 – 29 with an average and median VI of 2.3 and 1.8, respectively. The organic matter VI illustrated a more compressed range in variability compared to clay, but exhibited a much wider distribution of greater VI values (>3) across the state (Fig. 3). As was the case with clay VI, greater organic matter VI values were also concentrated within the Deep Loess Hills, Central Mississippi Valley Wooded Slope, and the Southern Mississippi River Alluvium MLRAs; however, high organic matter VI values were distributed widely throughout the landscape. These observed characteristics in the distribution and variability in organic matter suggested that this soil property was more heterogeneous than clay across the entire study area. Organic matter content may fluctuate as a function of changes in land-use management and surrounding weather conditions.

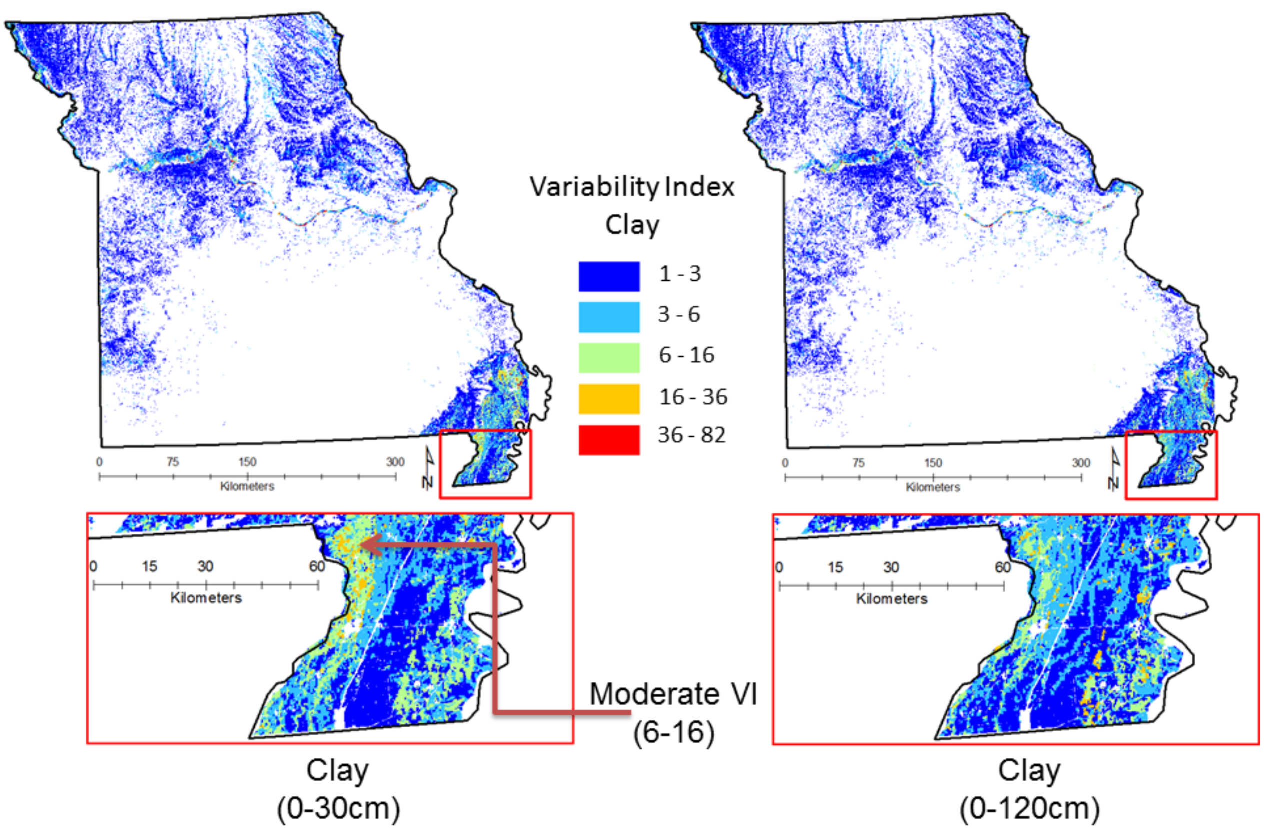


**Fig. 2: Variance Index (VI) values for clay at two depth increments (0-30 and 0-120 cm).**

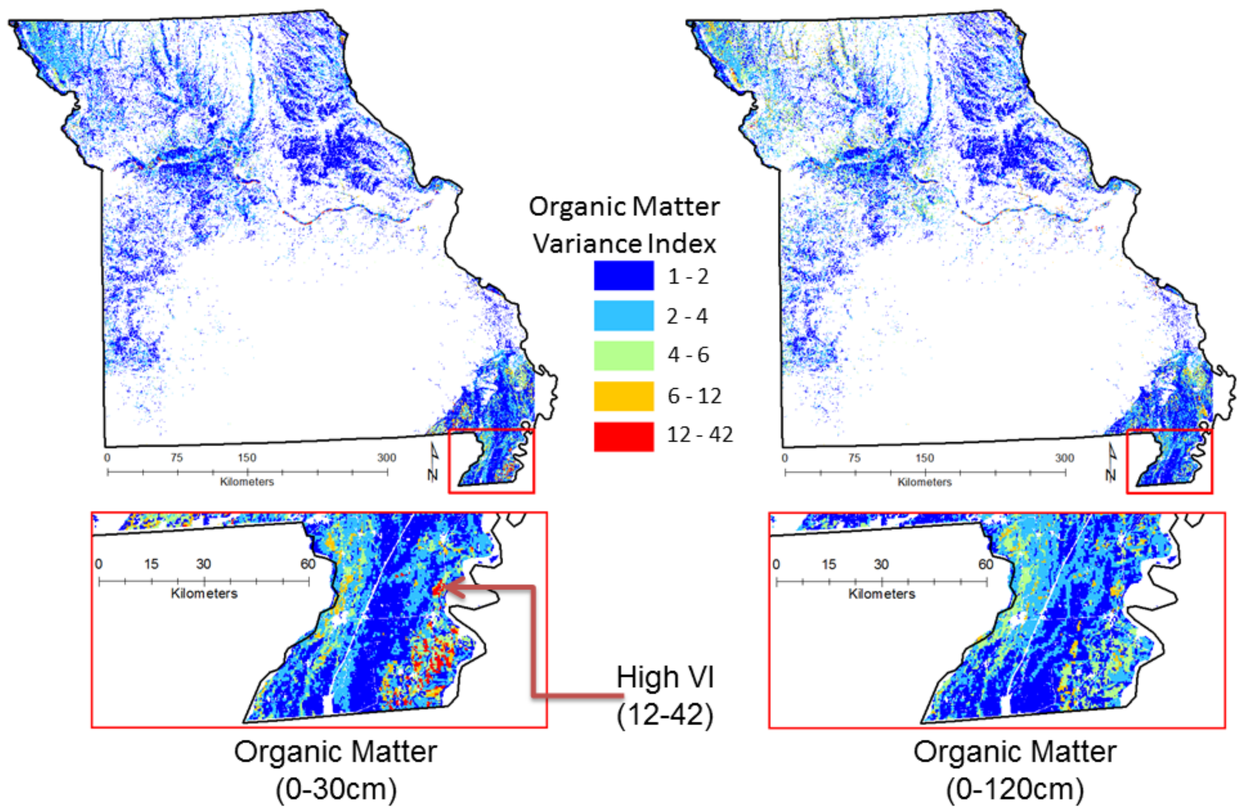


**Fig. 3: Variance Index (VI) for organic matter at two depth increments (0-30 and 0-120 cm).**

### Clustering of Soil Variability Index

The Anselin Local Moran's I function revealed several types of clustering in clay and organic matter VIs across the landscape. Of the fields significantly clustered, a total of 8% of clay VIs were categorized in each depth increment; whereas 13% and 25% of organic matter VIs were clustered in the top and profile increments, respectively (Table 2). The percentages of fields with LH and HL clusters was similar between clay and organic matter where each shared low numbers of fields in both increments.

**Table 2: The number and percentages of fields with significant clustering per depth increment within each Anselin Local Moran's I clustering category. Total represents the number of fields with significant clusters compared to the total number of fields (176000).**

| | Organic Matter | | Clay | |
|---|---|---|---|---|
| **Clustering** | **30 cm** | **120 cm** | **30 cm** | **120 cm** |
| High-High | 13215 (59%) | 20543 (47%) | 13081 (89%) | 13163 (91%) |
| High-Low | 320 (1%) | 859 (2%) | 29 (0.2%) | 36 (0.2%) |
| Low-High | 2436 (11%) | 3429 (8%) | 1604 (11%) | 1193 (8%) |
| Low-Low | 6483 (29%) | 18830 (43%) | 21 (0.1%) | 16 (0.1%) |
| Total | 22454 (13%) | 43661 (25%) | 14735 (8%) | 14408 (8%) |

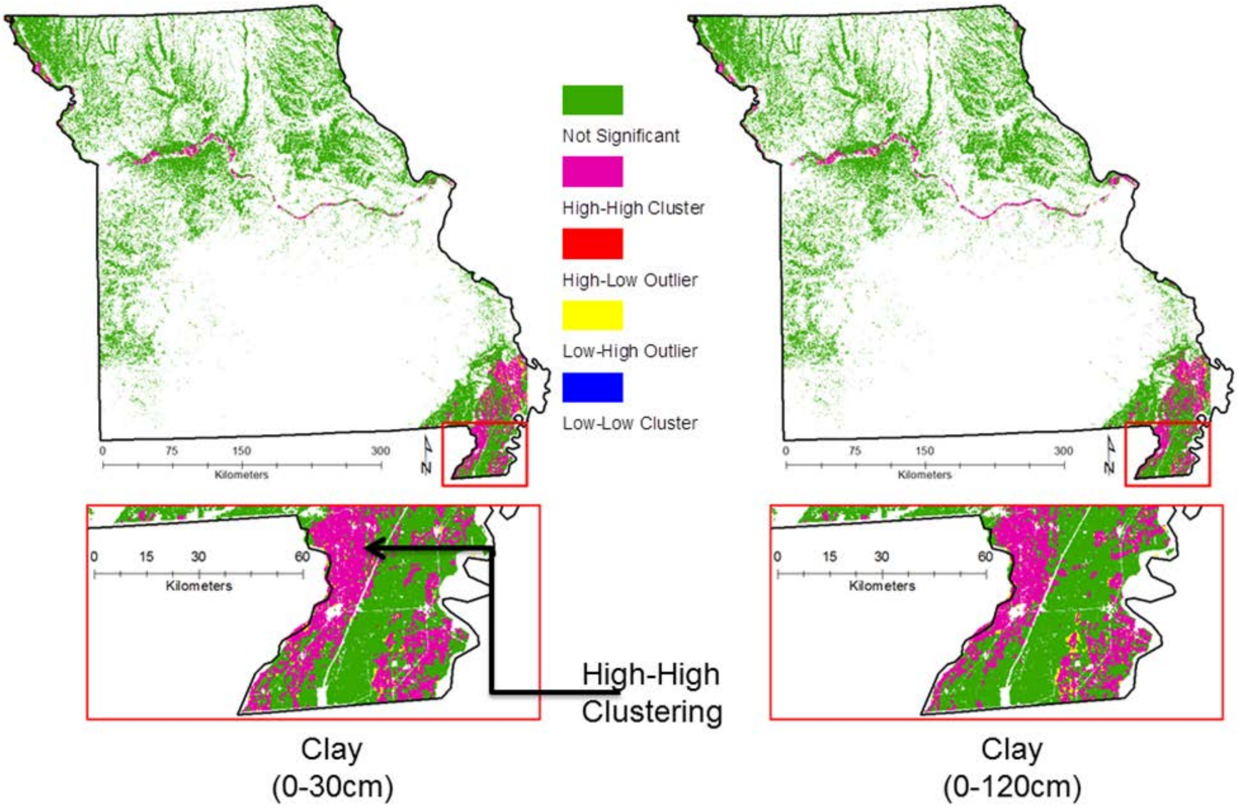


**Fig. 4: Clustering distributions for clay variability index values at top and profile depths.**

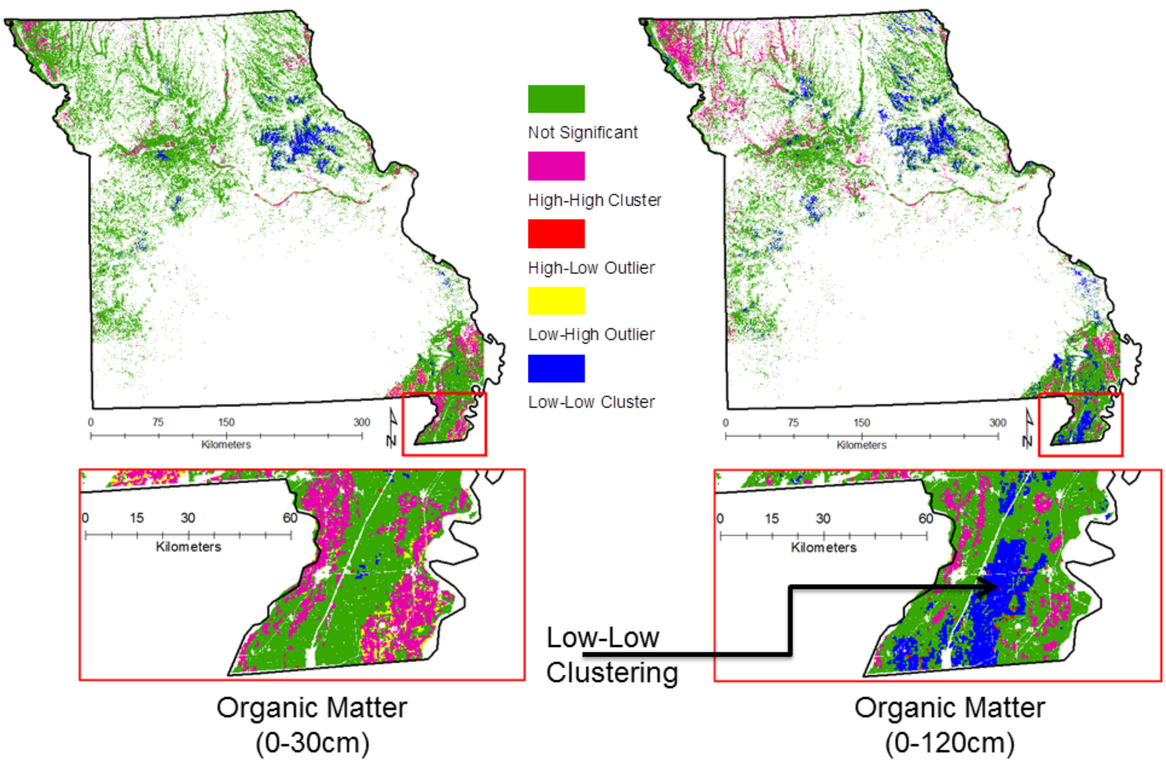


**Fig. 5: Clustering distributions for organic matter variability index values at top and profile depth increments.**

Clay VI clusters held tightly along the Missouri River and within the most southeastern portion of Missouri (Figure 4); whereas organic matter VI clustering was more widely distributed (Fig. 5). The most significant difference observed in clustering between organic matter and clay VI values was in the LL category. Only 0.01% of fields had LL clusters of clay VI for both the top and profile depth increments, while organic matter had about 4 and 11% of fields with LL clustering, respectively. The clustered areas with significantly low organic matter VI values were mainly demarcated by the Central Claypan MLRA. This region is generally defined by poorly drained soils with a restrictive claypan layer. In addition to the soil type, the weather characteristics (mean annual temperature around 12°C) and moisture regimes (mean annual precipitation around 1000 mm) may be contributing to an increase in cycling of organic matter that, in turn, lowers variability.

Soil properties can vary considerably given different types of management (e.g. crops, tillage regimes, intensity, fertilizer types, and rates in application) and soil forming factors (Jenny 1941). Consequently, the physical properties of the soil may respond by changing vertically with depth, laterally across fields, and even temporally due to weather conditions and human activities (Jung et al., 2006). Lastly, differences in the clay and organic matter VI, between the top (0-30 cm) and profile (0-120 cm), were determined to understand what depth increment captured the most variability. Fields with negative values indicated more variation throughout the 120-cm profile relative to the surface 30 cm; whereas positive values represented more variation in the top 30 cm than the 120-cm profile. The range of differences in clay VI between both depth increments was -26.5 to 35.3, with and average and standard deviation of 0.18 and 1.39, respectively. Organic matter VI had a smaller range in differences (-10 to 22.5), with an average and standard deviation of -0.29 and 1.13, correspondingly. Differences in clay VI values between depth increments showed that there was more variation in the top 30 cm where approximately 60% of fields had positive differences (Fig. 6). Conversely, only 39% of fields had

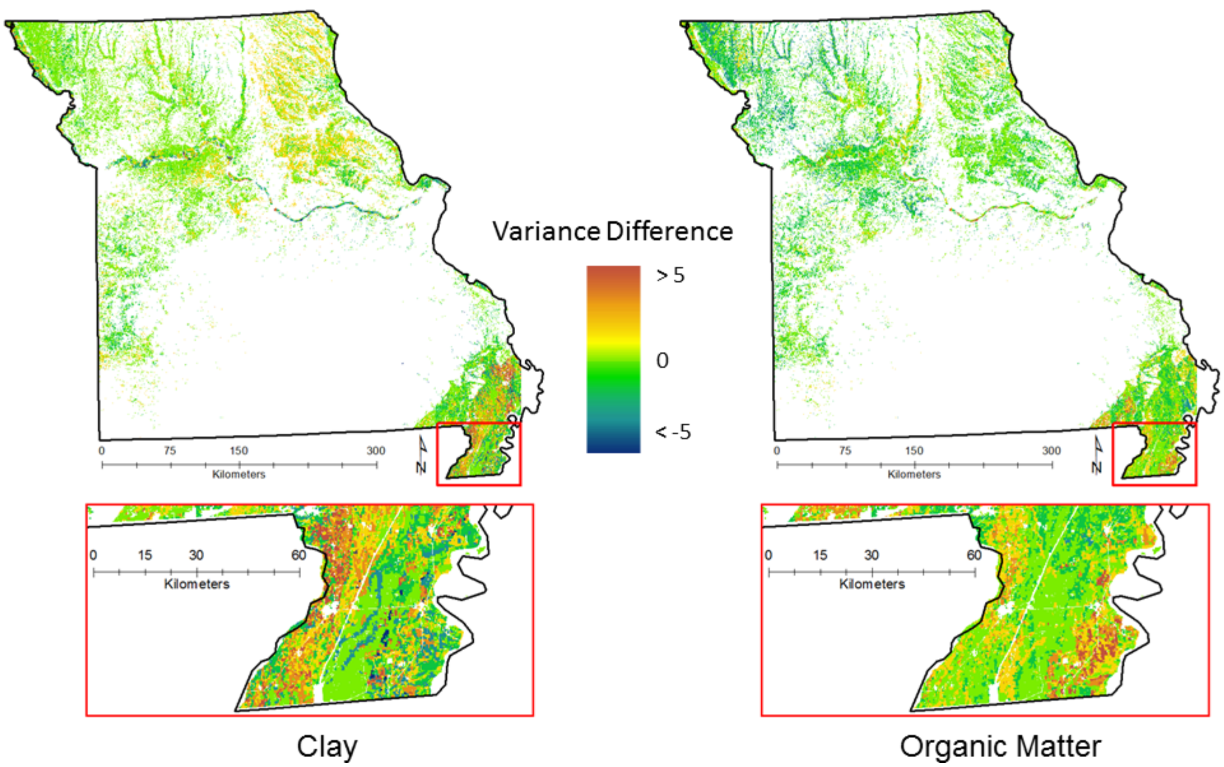


**Fig. 6: Distribution map of the differences in variability indices between top (0-30 cm) and profile (0-120 cm). Positive values indicate that the whole profile had lower VI values while negative values are indicative of greater variation in soil properties at the top depth increment.**

positive differences between the top and profile VI for organic matter. Thus, the topsoil generally captured more variability in clay content and the profile captured more variability in organic matter content; however, the differences varied from place to place as evidenced by the fact that the majority of fields with positive differences were found along the northeast side of Missouri extending into the most southeastern section along the Mississippi River. These results indicate that both the top and profile VIs have utility for identifying fields with the greatest and least potential for variation.

## Conclusion

This investigation calculated clay and organic matter VI values for individual fields for top (0-30 cm) and profile (0-120 cm) depth ranges across the state of Missouri, USA. We found that clay and organic matter VI ranged widely across the study region. The top depth for clay was found to have more variation than the profile, whereas organic matter varied more in the profile, rather than the top depth. Additionally, specific locations were identified across the state where groups of multiple fields had similar high or low VI values. The Missouri and Mississippi River alluvial floodplains shared similar groups of high clay and organic matter VI. The Central Claypan MLRA exhibited significant clustering of low VI values for organic matter. The results of this research highlight the ability of this methodology to identify clusters of variability, which may be used to target potential areas where the adoption of PA may have the greatest opportunity and return on investment.

Technological advancements are leading to new opportunities to access precision tools, techniques, and services, such as on-the-go sensors and high-resolution soil-landscape data. Individuals that manage large-scale agricultural fields exhibiting varying degrees of soil-landscape variation may be able to use results from this research to justify the cost of equipment needed to implement PA activities. Precision agriculture technology has decreased in cost over the last three decades and if this trend continues, coupled with more knowledge of potential field variation presented in this research, adoption may increase in the future. Whether PA solutions can be technically and economically used will continue to be important components to consider when evaluating adoption of PA technology. As such, this research provided a broad perspective on specific areas that exhibited different amounts of variation in soil properties across fields. Even though implementing PA technologies remain heavily situational and dependent on the amount of variation present within a field, this research highlighted fields and groups of fields with the greatest potential variation that could be used to facilitate future PA adoption.

## References

Adamchuk, V. I., Hummel, J. W., Morgan, M. T., & Upadhyaya, S. K. (2004). On-the-go soil sensors for precision agriculture. *Computers and Electronics in Agriculture, 44*, 71 - 91.

Anselin, L. (1995). Local indicators of spatial association - LISA. *Geographical Analysis, 27*, 93 - 115.

Beaudette, D. E., P., R., & O'Geen, A. T. (2013). Algorithms for Quantitative Pedology: A Toolkit for Soil Scientists. Retrieved from http://r-forge.r-project.org/projects/aqp/

Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Karlen, K. L., Turco, R. F., & Konopka, A. E. (1994). Field-Scale Variability of Soil Properties in Central Iowa Soils. *Soil Sci. Soc. Am. J., 58*, 1501 - 1511.

Castle, M. H., Lubben, B. D., & Luck, J. D. (2016). Factors Influencing the Adoption of Precision Agriculture Technologies by Nebraska Producers. *UNL Digital Commons*.

Erickson, B., & Widmar, D. A. (2015). Precision Agricultural Services Dealership Survey Results. *Sponsored by CropLife Magazine and the Center for Food and Agricultural Business*.

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., . . . Toulmin, C. (2010). Food Security: The Challenge of Feeding 9 Billion People. *Science, 327*, 812 - 818.

Griffin, T. W., Lowenberg-DeBoer, J., Lambert, D. M., Peone, J., Payne, T., & Daberkow, S. G. (2004). Adoption, Profitability, and Making Better Use of Precision Farming Data. *Staff Paper #04-06, Department of Agricultural*

*Economics*(Purdue University).

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment, 81*, 416 - 426.

Jung, W. K., Kitchen, N. R., Sudduth, K. A., & Anderson, S. H. (2006). Spatial Characteristics of Claypan Soil Properties in an Agricultural Field. *Soil Sci. Soc. Am. J., 70*, 1387 - 1397.

Karpinski, I., Schuler, J., & Müller, K. (2015). A new approach to support site-specific farming and economic decision making for precision agriculture in East Germany: The heterogeneity indicator. *Outlook on AGRICULTURE, 44*(4), 283 - 289.

Kitchen, N. R., Sudduth, K. A., Myers, D. B., Drummond, S. T., & Hong, S. Y. (2005). Delineating productivity zones on claypan soil fields using apparent soil electrical conductivity. *Computers and Electronics in Agriculture, 46*, 285 - 308.

McBratney, A., Whelan, B., & Ancev, T. (2005). Future Directions of Precision Agriculture. *Precision Agriculture, 6*, 7 - 23.

McLoud, P. R., Gronwald, R., & Kuykendall, H. (2007). *Precision Agriculture: NRCS Support for Emerging Technologies*. Greensboro, NC: East National Technology Support Center.

Mintert, J., Widmar, D., Langemeier, M., Boehlje, M., & Erickson, B. (2015). *The Challenges Of Precision Agriculture: Is Big Data The Answer?* Paper presented at the Southern Agricultural Economics Association, San Antonio, TX.

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika, 37*, 17 - 23.

Oliveira, R. P. d., Benites, V. d. M., Bernardi, A. C. d. C., Rabello, L. M., & Inamasu, R. Y. (2012). *Spatial Variability Index based on soil properties for notill and pasture site-specific management in Brazil.* Paper presented at the In Proceedings of the International Conference on Precision Agriculture, Indianapolis, IN.

R Core Team. (2015). R: A language and environment for statistical computing. Vienna, Austria. Retrieved from http://www.R-project.org/

Rabbi, S. M. F., Roy, B. R., Miah, M. M., Amin, M. S., & Khandakar, T. (2014). Spatial Variability of Physical Soil Quality Index of an Agricultural Field. *Applied and Environmental Soil Science, 2014*, 1 - 10.

Stadler, A., Rudolph, S., Kupisch, M., Langensiepen, M., van der Kruk, J., & Ewert, F. (2015). Quantifying the effects of soil variability on crop growth using apparent soil electrical conductivity measurements. *Europ. J. Agronomy, 64*, 8 - 20.

Tey, Y. S., & Brindal, M. (2012). Factors influencing the adoption of precision agricultural technologies: a review for policy implications. *Precision Agric, 13*, 713 - 730.

U.S. Census Bureau. (2015). International Data Base: World Population: 1950-2050. from http://www.census.gov/population/international/data/idb/worldpopgraph.php

USDA-NRCS. (2006). *Land Resource Regions and major Land Resource Areas of the United States, the Caribbean, and the Pacific Basin*.

Yan, L., & Roy, D. P. (2016). Conterminous United States crop field size quantification from multi-temporal Landsat data. *Remote Sensing of Environment, 172*, 67 - 86.

Ziadat, F. M., Dhanesh, Y., Shoemate, D., Srinivasan, R., Narasimhan, B., & Tech, J. (2015). Soil-Landscape Estimation and Evaluation Program (SLEEP) to predict spatial distribution of soil attributes for environmental modeling. *Int J Agric & Biol Eng, 8*(3), 158 - 172.