CROP MANAGEMENT

Crop, Forage & Turfgrass Management

Implications of choosing different interpolation methods: A case study for soil test phosphorus

Emma G. Matcham[1] | **Sakthi K. Subburayalu[2]** | **Steven W. Culman[3]** |
Laura E. Lindsey[4] (iD)

[1] Dep. of Agronomy, Univ. of Wisconsin–Madison, Madison, WI 53706, USA

[2] Agricultural Research Development Program, Central State Univ., Wilberforce, OH 45384, USA

[3] School of Environment and Natural Resources, The Ohio State Univ., Wooster, OH 44691, USA

[4] Dep. of Horticulture and Crop Science, The Ohio State Univ., Columbus, OH 43210, USA

Correspondence
Laura E. Lindsey, Department of Horticulture and Crop Science, The Ohio State University, Columbus, OH 43210, USA.
Email: lindsey.233@osu.edu

Assigned to Associate Editor D. Arnall

Abstract

Many farmers have their fields grid soil sampled to plan for variable rate P fertilizer application. Grid soil samples are often interpolated to create fertilizer application maps. However, most farmers and other practitioners do not compare interpolation methods. The objective of this study was to evaluate the performance of different grid soil sampling interpolation methods on P fertilizer prescription maps. Grid soil samples were collected from six fields and interpolated via geographically weighted regression (GWR), random forest (RF), and inverse distance weighting (IDW). At four out of six site–years, the root mean square error of soil test P (STP) was 7 to 16% lower for the GWR method compared with the RF method, and GWR identified the low-STP areas better than RF. Geographically weighted regression may outperform RF and IDW because of its lower error and reduced sensitivity to individual high- and low-STP values. Evaluating multiple interpolation techniques and comparing maps both visually and using global error rates can improve the decisions made by farmers and other practitioners.

1 | BACKGROUND ON SOIL SAMPLE INTERPOLATION

Soil sampling and corresponding test values estimate plant-available nutrients and are the basis for P fertilizer application recommendations (Culman et al., 2020). In precision nutrient management, farmers or their consultants may use a grid soil sampling approach to locate and quantify in-field variations in soil nutrient levels. Interpolated maps from grid sampled soils are used to develop a variable rate fertilizer prescription. Each field subunit in the prescription typically receives different rates of fertilizer with the goal of increased nutrient stewardship, use efficiency, and profitability (Wittry & Mallarino, 2004).

Most of the commonly used agricultural software programs, such as SMS (Ag Leader) and Trimble Ag Software (Trimble Inc.), provide options to the end user when it comes to the choice of the interpolation method. However, guidelines for choosing one method over the other are often general and lack any measure of uncertainty associated with the estimates. For fertilizer application at variable rates to improve nutrient stewardship, the interpolation method used to develop prescription maps must be carefully selected to accurately predict the nutrient needs of the crop while minimizing the potential negative environmental outcomes from excess application of P fertilizer. At high soil test P (STP) levels, P loss as runoff increases and contributes to non-point-source pollution (Carpenter et al., 1992; Dayton et al., 2014). If the interpolation underestimates the STP level, fertilizer may be overapplied, which increases the fertilizer cost without increasing yield (Culman et al., 2020). In cases where interpolation overestimates the STP level, fertilizer may be underapplied, leading to financial risk arising from reduced yield (Brooker et al., 2017; Fulford & Culman, 2018).

Abbreviations: GWR, geographically weighted regression; IDW, inverse distance weighting; RF, random forest; STP, soil test phosphorus.

Past regional-scale interpolation studies of soil organic matter and soil organic carbon that used elevation, slope, and other terrain attributes as covariates along with the soil test value have found that geographically weighted regression (GWR)-based interpolations had lower error than multiple linear regression, kriging, and regression kriging (Mishra et al., 2010; Wang et al., 2012). However, it has been reported that GWR is sensitive to multicollinearities between the covariates used as predictor variables at both the global and local scales (Wheeler & Tiefelsdorf, 2005). Random forest (RF) is another method that has proven to be useful in regional-scale spatial modeling applications and is widely used in digital soil mapping (Subburayalu & Slater, 2013). Random forest, a machine learning algorithm, creates an ensemble of decision trees using a random subset of observations and covariates to prevent the overfitting issues common in other decision tree algorithms and is particularly well-suited for handling multicollinearities (Grimm et al., 2008; Guio Blanco et al., 2018). However, both GWR and RF are seldom used for variable-rate application scales. Hence, a study that specifically compares these interpolation methods at in-field scales is important for evaluating the utility of these techniques for precision agriculture applications.

Global error measures, or metrics that summarize the error rate across the entire field, can be used to quantify interpolation accuracy and are available when interpolating via RF, GWR, and other spatial regression algorithms. This case study compared GWR and RF, two methods with comparable error metrics and covariates, with inverse distance weighting (IDW). Inverse distance weighting is a common interpolation method among farmers and their consultants and is used in agricultural software programs such as SMS (Ag Leader) and Trimble Ag Software (Trimble Inc.). Kriging was not considered in this study because the data did not meet the assumptions of stationarity (Hengl, 2009).

Quantifying error estimates of different interpolation techniques will help farmers and their consultants select the appropriate technique for variable-rate fertilizer applications. The objective was to compare three different interpolation techniques (GWR, RF, and IDW) on six fields in Ohio, to demonstrate as a case study, the importance of evaluating multiple interpolation techniques. Maps were compared visually and on the basis of the global error rate. Non-interpolated grid soil sampling maps were also included in the comparisons.

2 | METHODS FOR SAMPLING STP

2.1 | Sites

Within Ohio, two fields were sampled in 2017 and four fields were sampled in 2018, for a total of six site–years, prior to planting soybean [*Glycine max* (L.) Merr.]. Fields ranged

---

Core Ideas

- The field area with low soil test P (STP) varied among interpolation methods.
- Geographically weighted regression identified more low-STP areas than random forest.
- Inverse distance weighting was sensitive to outlier STP values.

---

from 44.5 to 116.1 acres. Site–years were denoted by location within the state (C = central, W = western, S = southern, and NW = northwestern) and year (17 = 2017 and 18 = 2018). Site descriptions are provided in Table 1. Cooperating farmers managed the fields according to standard practices. Soil sampling points were laid out in an even grid, where each grid cell was 0.50 acre in size and had one sampling location at its center. Small grid size is a practice being promoted within the industry (Integrated Ag Services, 2020). Within 10 ft of each grid point, three soil cores (8 inches deep and 0.75 inches wide) were collected (Culman et al., 2020). Cores from the same grid point were mixed and air-dried. Soil samples were analyzed for STP with a Mehlich-3 extractant (North-Central Regional Committee for Soil Testing and Plant Analysis-13, 2015).

The soil test results were mapped without interpolation, with an individual soil test result being used to represent the fertility level within the boundaries of each grid cell (see the example shown in Figure 1). Non-interpolated grid maps were compared with the maps derived via the three interpolations methods (GWR, RF, and IDW).

2.2 | Interpolating STP

Digital elevation models (2.5 ft resolution) were downloaded from the Ohio Geographically Referenced Information Program database (Ohio Office of Information Technology, 2018) to generate the following terrain attributes for each field in SAGA GIS (Conrad et al., 2015): slope, aspect, relative slope position (scaled from 0 to 1), and topographic wetness

TABLE A Useful conversions

To convert Column 1 to Column 2, multiply by	Column 1 suggested unit	Column 2 SI unit
0.405	acre	ha
2.54	inches	cm
1	ppm	mg kg^{-1}

TABLE 1 Site descriptions including soil taxonomic class and soil test values for organic matter (OM), pH, and soil test P (STP), with ranges shown in parentheses. Site–year names are generated from the year (2017 or 2018) and region (C, central Ohio; N, northern Ohio; W, west and central Ohio; S, southern Ohio; NW, northwest Ohio)

Site–year	Field size	Soil taxonomic class	pH median	OM median	STP median
	acre			%	ppm
C17	49	Fine, illitic, mesic Aeric Epiaqualfs	6.4(5.4–7.7)	3.4(2.3–6.2)	25(10–622)
C18	116	Fine, mixed, active, mesic Typic Argiaquolls	6.5(5.4–7.9)	3.8(2.1–5.7)	25(6–237)
W17	59	Fine, mixed, active, mesic Aeric Epiaqualfs	6.0(4.8–7.6)	3.2(2.1–4.7)	21(5–107)
W18	44	Fine, mixed, active, mesic Aquic Hapludalfs	6.8(5.1–7.9)	2.4(1.8–4.6)	28(5–236)
NW18	77	Fine, illitic, nonacid, mesic Typic Endoaquepts	6.8(5.9–7.9)	3.8(2.7–5.3)	101(28–196)
S18	124	Fine, mixed, superactive, mesic Typic Argiaquolls	7.2(5.2–8.2)	3.0(1.5–5.6)	23(4–270)

Legend
- Sampling Locations

Phosphorous (ppm)
- 21.0 – 45.8
- 45.8 – 70.6
- 70.6 – 95.4
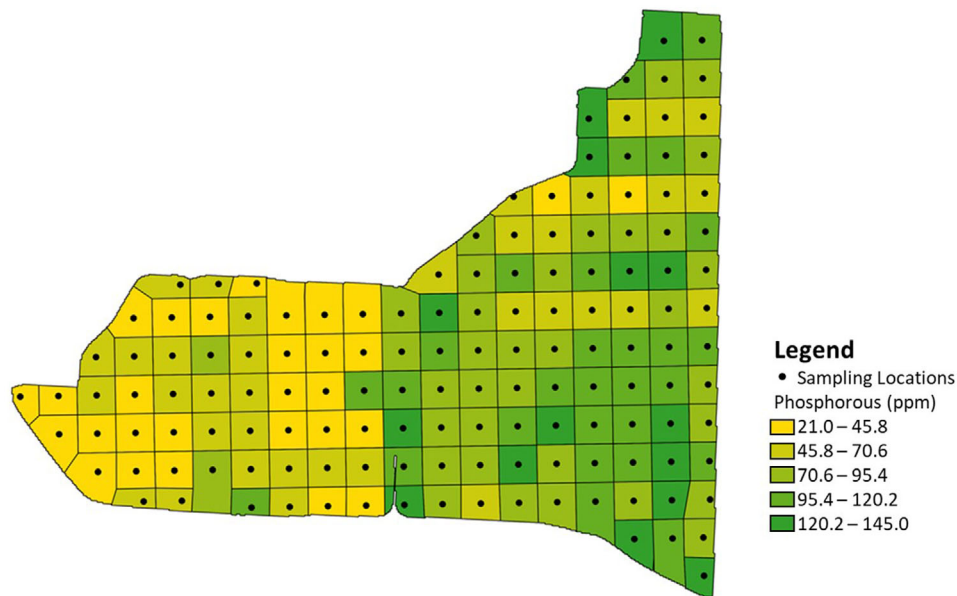- 95.4 – 120.2
- 120.2 – 145.0

FIGURE 1 Soil test P level without interpolation for site NW18, shown as an example. The field was split into approximately 0.5-acre grid cells, and a soil sample taken near the center (black dots in the figure) was used to represent the fertility of the entire cell. ppm, parts per million

index. Terrain attributes were used as covariates in the GWR and RF models.

Geographically weighted regressions were run in SAGA GIS according to the procedures detailed in Matcham et al. (2020) with a Guassian kernel and bandwidth = 1. Random forest models were built by the package *randomForest* within R (3.4.2) with the parameters ntree = 500 and mtry = 2, which are the values recommended by the package for a dataset of our size (Liaw & Wiener, 2002). For RF model evaluation, the use of out-of-bag error estimates has the same accuracy as the use of a testing dataset that is the same size as the training dataset (Breiman, 2001). Out-of-bag error estimates were calculated using *randomForest* and were used for RMSE calculations. The RMSE summarized the error for each field and was calculated from the soil test results of soil sampling and the estimated soil test results of interpolation at each soil sampling location. The RMSE was calculated for the GWR and RF interpolations by the package *Metrics* for RF and GWR interpolations (Hamner et al., 2018).

Inverse distance weighting interpolation was performed by SAGA GIS with the minimum number of points being one and a maximum of 20 for all sites (Conrad et al., 2015). The weighting function for inverse distance was to a power of two. The RMSE was not calculated for IDW interpolations because
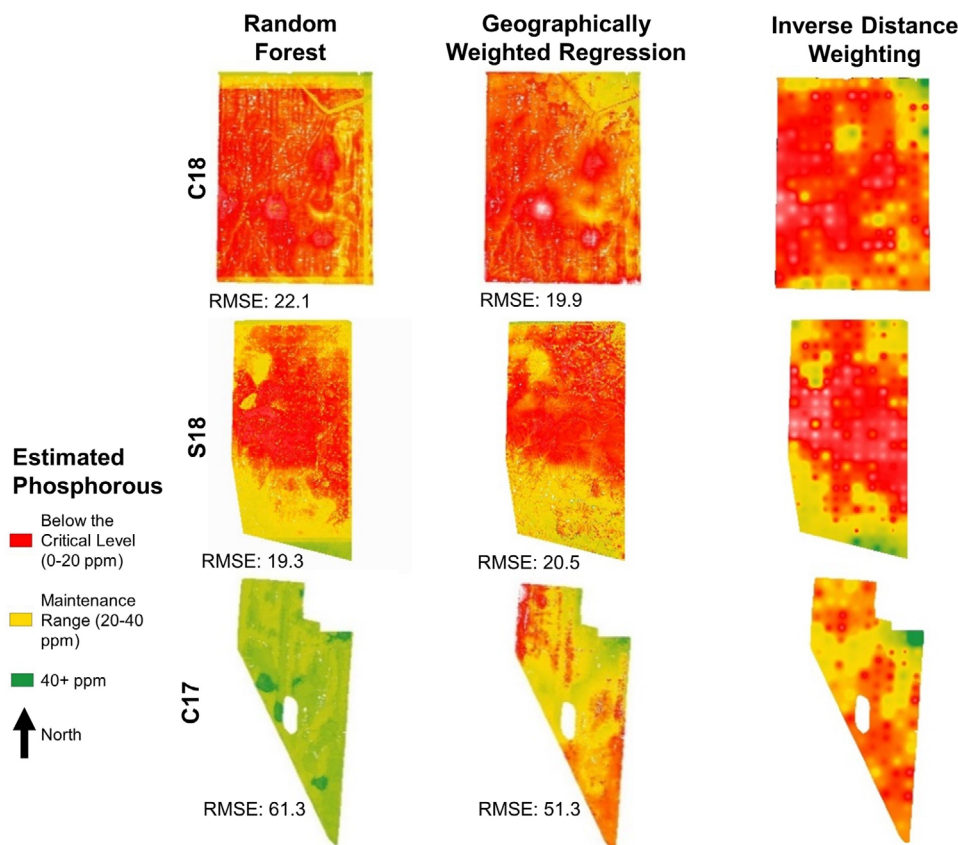
FIGURE 2 Interpolated soil test P (STP) maps using random forest, geographically weighted regression, and inverse distance weighting methods at three site–years in central and south-central Ohio. The critical level [20 parts per million (ppm)] and maintenance range (20–40 ppm) are defined in the Tri-State fertilizer recommendations (Culman et al., 2020). Field portions below the critical level are likely to have crop yields limited by low STP levels

the predicted STP value would have matched the observed STP value, resulting in a RMSE value of zero.

3 | COMPARISON OF INTERPOLATED STP MAPS

In four out of six site–years, the RMSE of STP was lowest for the GWR method (Figure 2, Figure 3). The RMSE for RF interpolations may have been higher, since RF minimizes the overall error rates without considering the error at the subfield scale, and RF is optimized to accurately predict the values that are most represented in the dataset. The full range of soil fertility levels or covariate values within the field may not have been evenly represented by grid soil sampling, and RF can have high error rates for unbalanced datasets (Subburayalu & Slater, 2013). The impact of unbalanced datasets on RF estimates seemed to override any potential improved handling of multicollinearity in the RF estimates compared with the GWR estimates.

Across site–years, RF and GWR identified similar areas that were within the maintenance range for STP [20 to 40 parts per million (ppm) Mehlich-3 P] (Culman et al., 2020)

(Figure 2, Figure 3). Inverse distance weighting predicted a wider range of soil fertility levels across fields than the other interpolation methods and predicted less of each field to be within the STP maintenance range. However, GWR identified a lower proportion of the field as falling above the maintenance range than RF or IDW at site–years C18, W17, W18, and S18.

Inverse distance weighting identified more of the field as being above the maintenance range than RF or GWR. A single high-STP sample tended to increase the estimated STP over a larger area of the field in IDW maps thn the GWR or RF maps. Geographically weighted regression tended to smooth over narrow portion of high STP.

The site–year C17 exhibited more differences between the RF and GWR maps, probably because of the high RMSE of both models at that site (Figure 2). The high RMSE and the larger discrepancy between the interpolation methods at C17 could be a result of having relatively few sampling points south of the tree patch. The IDW map at the C17 site was more similar in average fertility level to the GWR map at this site than the RF map (Figure 2). The NW18 site–year had minimal differences among the maps, probably because all soil samples at NW18 were above the critical STP level (Figure 3).
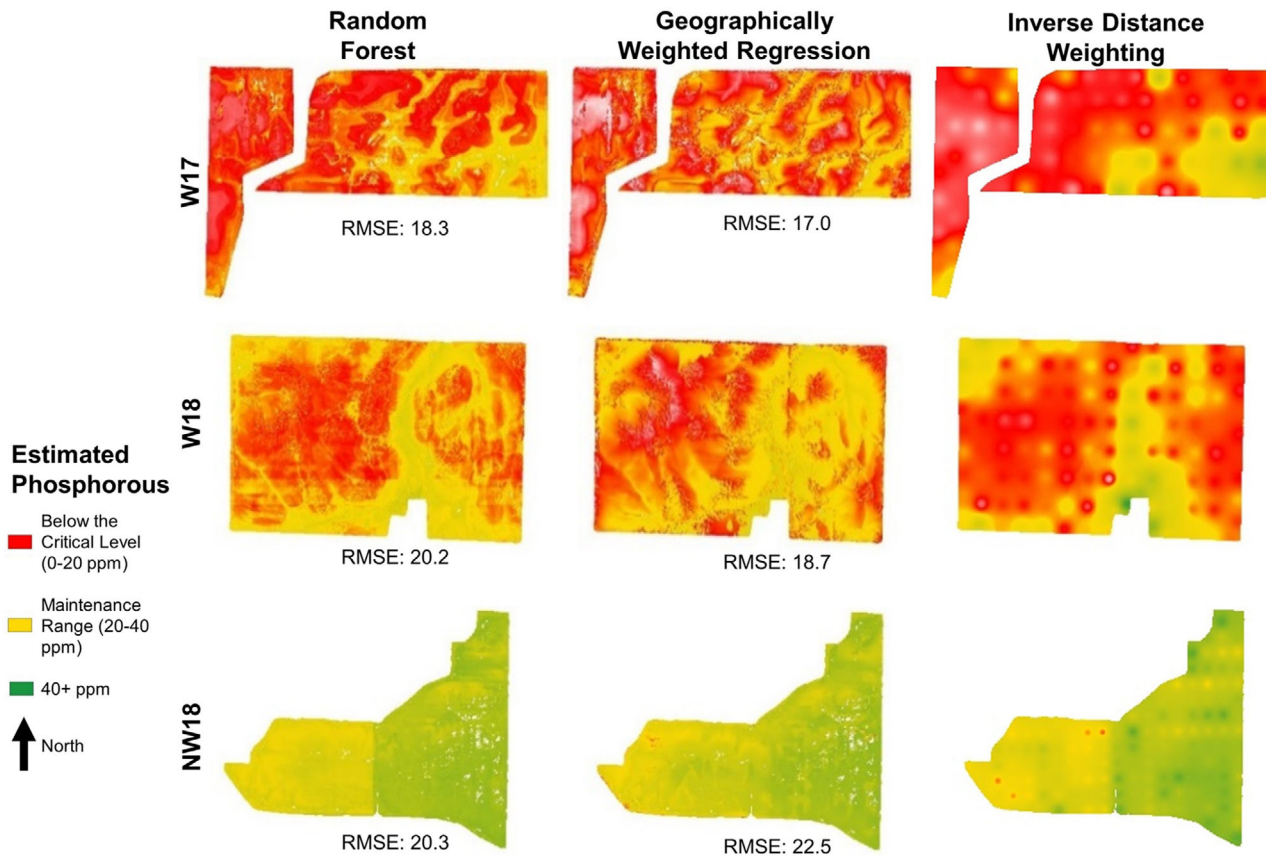
FIGURE 3 Interpolated soil test P (STP) maps via random forest, geographically weighted regression, and inverse distance weighting methods at three site–years in western and northwestern Ohio. The critical level [20 parts per million (ppm)] and maintenance range (20–40 ppm) are defined in the Tri-State fertilizer recommendations (Culman et al., 2020). Field portions below the critical level are likely to have crop yields limited by low STP levels

TABLE 2 Area below the critical soil test P (STP) level of <20 parts per million based on random forest, geographically weighted regression, inverse distance weighting, and noninterpolated grids. Site–years are denoted by location within the state (C, central; W, western; S, southern; NW, northwestern) and year (17, 2017; 18, 2018)

		Field area below the critical level			
Site–year	Field size	Random forest	Geographically weighted regression	Inverse distance weighting	Noninterpolated grid
		————————————————acres————————————————			
C18	116	9.4	12.4	24.2	35.6
W17	59	11.9	14.8	22.7	29.1
W18	44	0	3.2	1.5	7.7
S18	79	9.4	6.7	11.1	31.1
NW18	77	0	0	0	0

4 | IMPLICATIONS FOR SOIL PHOSPHORUS MANAGEMENT

When the STP is below the critical level of 20 ppm, P fertilizer is recommended to build up the STP levels (Culman et al., 2020). In our study, the field area below the critical level varied among the soil interpolation methods and the non-interpolation method (Table 2). Among the interpolation methods, RF tended to result in the smallest amount of field area below the critical level, whereas IDW tended to have the largest amount of field area below the critical level. The non-interpolated maps identified the greatest amount of field area below the critical level, so applying P based on the non-interpolated grid maps would maximize the per-field fertilizer

application rate and minimize the risk of yield loss caused by low STP limitations. However, over time, this method may increase the environmental risk of phosphorous run-off caused by high STP soils. Maps created with IDW were very similar to the non-interpolated grid maps and skewed towards higher P application recommendations. Additionally, maps generated by IDW had larger sections of the field influenced by individual high- or low-STP samples than GWR or RF.

5 | RECOMMENDATIONS

Geographically weighted regression estimated more low-STP areas of the field than RF and had a lower error, as measured by RMSE, than RF for four of six site–years. Taken in combination, the low error, the limited sensitivity to individual high- and low-STP values, and the ability to identify low-STP areas suggest that GWR may often outperform RF and IDW and be a reliable method of choice. However, GWR was not universally the error-minimizing method. We recommend comparing maps both visually and on the basis of RMSE before applying fertilizer to improve decision-making by farmers and other practitioners. Visual comparisons near environmentally sensitive areas are particularly important when using IDW or noninterpolated grid maps that may be biased by an individual high- or low-STP sample and can help evaluate other interpolation methods outside the scope of this paper, such as kriging. Although grid soil sampling approaches can improve fertilizer efficiency and farm profitability, our results show that additional gains can be realized by validating the interpolation methods when generating prescription fertilizer application maps. Although this case study specifically focused on P application maps, the implications of our findings can probably be extended to other nutrients managed in a variable rate approach.

CONFLICT OF INTEREST
The authors declare that there is no conflict of interest.

AUTHOR CONTRIBUTIONS
Emma G. Matcham: conceptualization, formal analysis, investigation, methodology, writing—original draft. Sakthi Kumaran Subburayalu: methodology, writing—review and editing. Steven Culman: methodology, writing—review and editing. Laura Lindsey: funding acquisition, supervision, writing—review and editing.

ORCID
Laura E. Lindsey https://orcid.org/0000-0001-7026-0949

REFERENCES
Breiman, L. (2001). *Random forests*. UC Berkley, Statistics Department. https://www.stat.berkeley.edu/%7Ebreiman/randomforest2001.pdf

Brooker, A. P., Lindsey, L. E., Culman, S. W., Subburayalu, S. K., & Thomison, P. R. (2017). Low soil test phosphorus and potassium limit soybean grain yield in Ohio. *Crop, Forage, and Turfgrass Management*, *3*, 1–5. https://doi.org/10.2134/cftm2016.12.0081

De Donato, P., Cases, J. M., Humbert, B., Lutgen, P., & Feyder, G. (1992). Eutrophication of aquatic ecosystems: Bistability and soil phosphorus. *Proceedings of the National Academy of Sciences*, *102*, 10002–10005. https://doi.org/10.1002/polb.1992.090301201

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., & Böhner, J. (2015). System of Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geoscientific Model Development*, *8*, 1991–2007. https://doi.org/10.5194/gmd-8-1991-2015

Culman, S., Fulford, A., Camberato, J., Steinke, K., Lindsey, L., LaBarge, G., Watters, H., Lentz, E., Haden, R., Herman, B., Hoekstra, N., Thomison, P., Minyo, R., Dorrance, A., Rutan, J., & Warncke, D. (2020). *Tri-State fertilizer recommendations for corn, soybeans, wheat, and alfalfa*. (Bulletin 974). The Ohio State University.

Dayton, E. A., Whitacre, S. D., & Holloman, C. H. (2014). Demonstrating the relationship between soil phosphorus measures and phosphorus solubility: Implications for Ohio phosphorus risk assessment tools. *Journal of Great Lakes Research*, *40*, 473–478. https://doi.org/10.1016/j.jglr.2014.04.001

Fulford, A. M., & Culman, S. W. (2018). Over-fertilization does not build soil test phosphorus and potassium in Ohio. *Agronomy Journal*, *110*, 56–65. https://doi.org/10.2134/agronj2016.12.0701

Grimm, R., Behrens, T., Märker, M., & Elsenbeer, H. (2008). Soil organic carbon concentrations and stocks on Barro Colorado Island—Digital soil mapping using Random Forest analysis. *Geoderma*, *146*, 102–113. https://doi.org/10.1016/j.geoderma.2008.05.008

Guio Blanco, C. M., Brito Gomez, V. M., Crespo, P., & Ließ, M. (2018). Spatial prediction of soil water retention in a Páramo landscape: Methodological insight into machine learning using random forest. *Geoderma*, *316*, 100–114. https://doi.org/10.1016/j.geoderma.2017.12.002

Hamner, B., Frasco, M., & LeDell, E. (2018). *Metrics*. Github. https://github.com/mfrasco/Metrics

Hengl, T. (2009). *A practical guide to geostatistical mapping* (2nd ed.). EUR 22904 EN Scientific and Technical Research series report. Office of Publications of the European Communities.

Integrated Ag Services. (2020). *HD soil sampling*. Integrated Ag Services. https://integratedag.net/services/high-density-automated-precision-soil-sampling

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*, 18–22.

Matcham, E. G., Hamman, W. P., Hawkins, E. M., Fulton, J. P., Subburayalu, S., & Lindsey, L. E. (2020). Soil and terrain properties that

predict differences in local ideal seeding rate for soybean. *Agronomy Journal*, *112*, 1981–1991. https://doi.org/10.1002/agj2.20179

Mishra, U., Lal, R., Liu, D., & Van Meirvenne, M. (2010). Predicting the spatial variation of the soil organic carbon pool at a regional scale. *Soil Science Society of America Journal*, *74*, 906–914. https://doi.org/10.2136/sssaj2009.0158

North-Central Regional Committee for Soil Testing and Plant Analysis-13. (2015). *Chemical soil test procedures for the North Central region*. North Central Regional Research Publication No. 221. North-Central Regional Committee for Soil Testing and Plant Analysis. https://p2infohouse.org/ref/17/16690.pdf#page=33

Ohio Office of Information Technology. (2018). *Ohio geographically referenced information program*. Ohio Office of Information Technology https://ogrip.oit.ohio.gov

Subburayalu, S. K., & Slater, B. K. (2013). Soil series mapping by knowledge discovery from an Ohio county soil map. *Soil Science Society of America Journal*, *77*, 1254–1268. https://doi.org/10.2136/sssaj2012.0321

Wang, K., Zhang, C., & Li, W. (2012). Comparison of geographically weighted regression and regression kriging for estimating the spatial distribution of soil organic matter. *GIScience & Remote Sensing*, *49*, 915–932.

Wheeler, D., & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems*, *7*, 161–187. https://doi.org/10.1007/s10109-005-0155-6

Wittry, D. J., & Mallarino, A. P. (2004). Comparison of uniform- and variable-rate phosphorus fertilization for corn-soybean rotations. *Agronomy Journal*, *96*, 26–33. https://doi.org/10.2134/agronj2004.0026