Received: 26 April 2019 | Accepted: 7 August 2019 | Published online: 10 January 2020

DOI: 10.1002/agj2.20048

Agronomy Journal

ARTICLE

Soil Fertility & Crop Nutrition

Guiding soil sampling strategies using classical and spatial statistics: A review

Patrick G. Lawrence[1] | **Wayne Roper**[2] | **Thomas F. Morris**[2] | **Karl Guillard**[2]

[1]Repliculture LLC, Bozeman, MT 59715

[2]Dep. of Plant Science and Landscape Architecture, Univ. of Connecticut, 1376 Storrs Rd, Unit 4067, Storrs, CT 06269

Correspondence
Patrick G. Lawrence, Repliculture LLC, Bozeman, MT 59715.
Email: Patrick@repliculture.com

Funding information
Northeast SARE, Grant/Award Number: LNE18-363

Abstract

Soil analysis is a key practice to increase the efficiency of nutrient management in agriculture. Since the early 20th century, increasingly sophisticated methods have been developed to describe and manipulate the inherent spatial variability in soil chemical properties within the realms of classical and spatial statistics. In this paper, we reviewed design-based (classical) and model-based (geostatistical) sampling to suggest field-scale sampling strategies consistent with common agronomic management goals in annual crop production systems. To assess the relevance of common sampling methods in relation to practice, current extension recommendations across the United States were compared with results from peer-reviewed literature. Despite decades of research, specific recommendations for sample sizes, sampling depths, numbers of soil cores, and layouts were highly variable for classical and geostatistical approaches. Mobile nutrients, such as NO_3, are frequently lacking in spatial structure and rarely are recommended for site-specific management. Nonmobile nutrients, such as P, are more spatially dependent and exhibit nested spatial structures that are inconsistent across fields. For these reasons, we recommend design-based sampling in most situations for simplicity, cost, and objectivity. The common design-based sampling protocol prescribes collection of individual cores in a zig-zag pattern that are combined to produce a composite sample. This protocol should be amended because it is not sufficiently randomized and is inadequate for log-normally distributed variables. To facilitate site-specific management, we recommend structured approaches for delineating management zones or strata and for researchers to systematically enumerate confounding variables while explicitly defining the scope of inference for future soil sampling studies.

1 | INTRODUCTION

Surface water eutrophication, nutrient pollution, and inflated input costs from improper fertilizer application have plagued row-crop agriculture since fertilizer use became widespread in the second half of the 20th century (Carpenter et al., 1998). Only 16% of applied P is taken up by cereal crops (Dhillon, Guilherme, Driver, Figueiredo, & Raun, 2017), and 47% of applied N is taken up by all crops (Ladha et al., 2016), with the rest remaining in croplands or lost via runoff, leaching, and gaseous emissions. The 4Rs concept (i.e., choosing the right fertilizer to apply at the right rate, right time, and in the

Abbreviations: EOSD, economic optimum sample density; MZ, management zone; SOC, soil organic carbon; SOM, soil organic matter; SRS, simple random sampling; TRV, Theory of Regionalized Variables.

right place) has frequently been used to minimize the over-application of fertilizer (Peterson & Bruulsema, 2019). Based on continuing high levels of surface water pollution and low percentages of fertilizer that end up in crop biomass, there remains room for improving application of the 4Rs across heterogeneous farm landscapes.

Three of the 4Rs depend on knowing the background level of soil nutrient availability at specific dates across farms and within fields, which in turn relies on proper soil sampling protocols and techniques. Given the inability to know exact nutrient concentrations at every point in space, uncertainty in field and subfield scale statistics is inevitable. Laboratory techniques for nutrient analysis typically contribute only 1.8 to 5% to the overall uncertainty of nutrient concentration estimates, depending on the nutrient measured and its overall concentration (Gustavsson, Luthbom, & Lagerkvist, 2006; Hurisso, Culman, & Zhao, 2018), with the remaining uncertainty arising from inherent spatiotemporal variability and agronomic management. Measuring or accounting for the spatial variability in nutrients and understanding its implications for fertilizer application is crucial for moving closer to the ideal of the 4Rs.

Beginning in the 1970s, recommendations for soil sampling have become more sophisticated, with numerous sources (e.g., extension publications, agronomic consultants, and private laboratory recommendations) providing suggestions for site-specific or stratified soil sampling methods. These recommendations are mirrored by an increase in the number of academic publications investigating the strength and nature of spatial variability in a wide variety of landscapes and cropping systems. In all sources (academic and industry), recommendations are wide ranging. For example, in a notably large-scale study of spatial variability within 23 farm fields, the density of soil sampling required to adequately characterize spatial variation in soil P ranged from <30 to 105 m (Lauzon, O'Halloran, Fallow, Bertoldi, & Aspinall, 2005). Sampling requirements thus differ for different locales and can vary for different nutrients measured in the same field (Cahn, Hummel, & Brouer, 1994).

The potential causes of variability are numerous and lead to differing sampling requirements. Soil type, biota, climate, and agronomic practices may all influence the scale and scope of variation. The goal of sampling, whether field-scale characterization or site-specific mapping, can drastically affect the sampling plan. Finally, differences in study methodologies, such as depth of sampling, sampling month, and statistical analysis, make it difficult to develop sampling plans that provide reliable estimates of the expected accuracy and precision.

An understanding of optimal soil sampling methods across the major crops, soils, and environments will improve the uptake efficiency of nutrients applied to grow our food, feed, and fiber and reduce the unsustainable amount of water and air pollution that occurs during crop production. The objectives of this review are to (i) provide a historical overview of past and current soil sampling strategies used in production agriculture, (ii) summarize how past and current soil sampling strategies are used to guide nutrient applications in production agriculture, and (iii) guide future research and practical recommendations for soil sampling. The scope of the review was limited to annual row-crop agriculture and within-field sampling.

Core Ideas

- Design-based (classical) sampling is recommended in most fields.
- Zig-zag sampling patterns and soil core compositing should be avoided where possible.
- Fewer soil cores per unit area are needed as field size increases.
- Extrapolation from grid-based sampling is limited by confounding variables.

We first reviewed historical peer-reviewed soil sampling literature in row-crop agriculture to contextualize current recommendations and practice. We also categorized soil sampling methods and described recent research that has influenced recommendations about the configuration and density of sampling. The dichotomy of design-based and model-based sampling (classical stratification and randomization versus grid-based geostatistical sampling, defined in more detail below) was used to guide the practitioner toward sampling configurations consistent with their desired end use. Where appropriate, we quantitatively compared findings across agricultural and bioclimatic environments. We used the United States as a case study to characterize disparities between common sampling recommendations and academic understanding. In this context, the US Land Grant University Cooperative Extension System served as the source of recommendations. Finally, we present suggestions for soil sampling protocols to guide fertilizer and manure applications and to meet the needs for regulatory compliance, specifically concerning P pollution.

2 | METHOD

We searched agricultural extension websites in each US state for specific pages or publications on soil sampling to capture the current state of soil sampling recommendations. Soil sampling protocols and recommendations were included in our analysis only if they referenced production agriculture (not home lawns or gardening). Three US-focused, peer-reviewed, and extension-oriented journals were also searched to assess

whether public-targeted extension publications were consistent with academic extension literature. For this reduced scope, the search terms were limited to "soil sampling," and the journals included were the *Journal of Production Agriculture*; *Journal of Extension*; and *Crop, Forage and Turfgrass Management* (including their previously separate constituent journals).

For the peer-reviewed academic literature review, we searched the Academic Search Premier and CAB Abstracts academic databases using the terms "soil AND (sampling OR [spatial OR variogram OR kriging]) AND (nutrient OR nitrogen OR phosphorus) AND agriculture." Our search was limited to papers and books written in English. The search was conducted in December 2018, and any duplicate references were removed.

To be included in the analysis, each study (i) examined nutrient management for annual row-crop agriculture; (ii) studied within-field variation (not between-field, although multiple fields could be included); (iii) reported at least 10 within-field observations to allow for descriptive statistics; (iv) determined the variability of measured nutrients; (v) measured at least one of the following nutrients/soil parameters: NO_3–N, NH_4–N, plant mineralizable N, plant-available N, P, total C, total N, soil organic matter (SOM), or soil organic carbon (SOC), Ca, Mg, Al, B, Cu, Fe, Mn, Zn, S, Pb, and pH; and (vi) used either empirical or modeling-based methods of analysis. If the study was modeling based, the criterion of 10 within-field observations was omitted.

The database search was made more comprehensive by gathering first- and second-order references cited by the initial set of publications. Titles and abstracts were read to assess whether each publication fit the inclusion criteria. If so, the text was evaluated for the same criteria.

Initially, 5346 publications were accessed in the database search. From those and the publications accessed by gathering first- and second-order references, 136 met our inclusion criteria, and 42 provided statistical data sufficient for quantitative comparison of results. Four of the 42 publications provided nonspatial results, with the remaining 38 providing results that included spatial parameters and variation. However, the 38 remaining publications did not consistently overlap in measured soil properties, sampling designs, crop types, and agro-environmental contexts; therefore, a rigorous meta-analysis was not possible. Instead, the 42 and 38 publications provided a foundation for understanding the possible structures of spatial variation and the numbers of samples required to characterize soil chemical properties. The entire set of 136 initially accessed publications was used to trace historical trends in soil sampling, to understand the effects of specific agronomic factors on spatial variation, and to identify optimal soil sampling methods as informed by modeling studies.

3 | TRENDS AND STATE OF PEER-REVIEWED SPATIAL SOIL SAMPLING LITERATURE

In this section, we provide a brief overview of historical trends in spatial soil sampling to contextualize current practices and to understand the origin of sampling recommendations commonly provided to farmers.

Spatial variability of soil nutrients has been an important topic in soil science literature since the early 20th century. Early studies explicitly recognized the lack of uniformity across fields and conceptualized nutrient variability as resulting from laboratory error and field error (Robinson & Lloyd, 1915). As the corpus of soil sampling research expanded, the field component of variability was further decomposed into vertical, horizontal, and temporal variability (Cline, 1944), which could all exist at multiple scales (Reed & Rigney, 1947).

Early soil sampling literature used classical statistical concepts of randomization and stratification to account for the spatial variability of soil properties, which are still widely used today. These studies aimed to characterize inter- and intra-class correlation for soil series and properties (Heuvelink & Webster, 2001; Webster & Beckett, 1968) and did not attempt to map or predict in units smaller than singular strata, despite the recognition that average test values used for fertilizer recommendations resulted in overapplication in some areas and underapplication in others (Peck & Melsted, 1967). Most research provided recommendations for the number of soil cores necessary to create a sufficiently representative composite sample, with the typical goal to estimate field-scale averages. Between 10 and 20 soil cores were frequently recommended for all nutrients for a field size of 4 to 8.1 ha (10–20 acres) (Hemingway, 1955; Tisdale and Nelson, 1956), although this varied by nutrient with nitrate requiring many more cores if higher precision was desired (45 cores for an 80% confidence interval within 10% of the mean on 75% of sampled areas) (Meisinger, 1984) and may have been partially based on presumed economic feasibility rather than statistical validity. When one set of soil cores was used to characterize field-scale means of multiple soil properties, the property requiring the greatest intensity of sampling determined the required sampling intensity for all properties (Reed & Rigney, 1947).

Before 1970, spatial designs for soil sampling followed the classical concepts of simple random sampling (SRS), stratified random sampling (Cochran, 1977), or systematic sampling (Madow & Madow, 1944). These sampling methodologies can be briefly described as random selection of sampling points throughout a field, division of a field into relatively homogeneous areas before random selection, and the use of a grid or pattern for choosing sampling locations. Using these methodologies, cores were often analyzed

individually to quantify variation for research purposes with the knowledge that, in practice, all cores would be composited before analysis.

The inability of compositing to provide any measure of variability was noted as a serious limitation (Peck & Melsted, 1967) even though the practice was widely used and recommended. Other limitations in the application of classical methodologies were also recognized. Cline (1944) noted that, "walking 'at random' over an area when selecting sampling units is far from complete randomization and is subject to strong personal bias" and warned about potential bias if sampling grids were aligned parallel to systematic variation in the soil such as crop rows.

Although motivation existed to overcome the limitations of classical sampling and exploit variability to improve crop productivity, the required statistical methodology had not yet been developed. Early researchers were aware of the tendency for neighboring observations of soil properties to be more similar than observations separated by large distances (spatial autocorrelation), but it remained a nuisance rather than an object of study.

The use of geostatistics to describe spatial variation was initiated in the 1960s with the publication of the Theory of Regionalized Variables (TRV) (Matheron, 1963). The TRV provided a statistical framework for quantitatively analyzing localized continuous variation of biophysical properties. Early applications were limited to mining geology; however, by the 1980s, the TRV had been extended to soil sampling (Burgess & Webster, 1980). The foundation for TRV is the concept of the variogram, which describes the variance between all point pairs separated by a lag interval h. The semivariance $\gamma(h)$ (Eq. (1)), representing the per-observation variance, is half the expected value of the value of the variance and is more commonly used to describe the continuity between observations $z(i)$ across a two-dimensional plane. To estimate $\gamma(h)$ for integer values of h (assuming a constant mean within the lag width around h), the semivariance is calculated as:

$$\gamma(h) = \frac{1}{2} E\left\{[z(i) - z(i+h)]^2\right\} \quad (1)$$

To describe changes in the semivariance as h increases, the semivariance for each lag interval (e.g., 0–10 m, 10–20 m, etc.) is calculated and plotted against h to produce the semivariogram. At times, the correlogram is used instead of the semivariogram and is related to the semivariogram:

$$\rho(h) = 1 - \frac{\gamma(h)}{\sigma^2} \quad (2)$$

where $\rho(h)$ is the autocorrelation at lag h, and σ^2 is the variance.

Plots of semivariance versus h are typically created using field-collected data to create empirical semivariograms, to which curves are fit and then used for field-scale spatial interpolation. Both the semivariogram and correlogram assume intrinsic stationarity: that the mean is constant and the covariance between observations at different locations is dependent only on their spatial separation, not their absolute location. In addition, isotropy, a lack of directional trends to the correlations, is often assumed to create these models of spatial variation, although given enough data directional (anisotropic) models may also be fit.

Refinement of the methods for estimating variograms (Cressie, 1985; Yfantis, Flatman, & Behar, 1987) and for predicting soil properties in unsampled locations via inverse distance weighted interpolation and kriging (Burgess & Webster, 1980; McBratney & Webster, 1981; Olea, 1984; Warrick & Myers, 1987) occurred in the 1980s. Sampling on a triangular grid was recognized to be the most efficient way to collect spatial soil samples (McBratney & Webster, 1981; Yfantis et al., 1987), but square grids were nearly as efficient, much easier to design, and almost ubiquitous. Following research by Lockman and Molloy (1984), the effect of temporal fluctuations in nutrients on fertilizer recommendations received empirical attention, culminating in the now-common recommendation to perform sampling in the fall after harvesting crops (excepting spring nitrate in semiarid locations and the Pre-Sidedress Nitrogen Test in humid temperate regions).

Empirical research during the 1970s and 1980s focused on describing soil properties in terms of single parameters (mean and variation) (Cameron, Nyborg, Toogood, & Laverty, 1971; Reuss, Soltanpour, & Ludwick, 1977) and mapping them as continuous surfaces (Dahiya, Anlauf, Kersebaum, & Richter, 1985; Tabor, Warrick, Myers, & Pennington, 1985; van Meirvenne & Hofman, 1989; Webster & McBratney, 1987). This expansion of empirical work continued into the 1990s, with the advent of mechanized Precision Agriculture/Site-Specific Crop Management gaining traction in the middle of the decade when Global Positioning System mapping equipment became commercially available. Previously, variation in soil properties could be spatially managed only on small scales, but, with the prospect of using larger automated fertilization equipment, spatial nutrient availability could be manipulated at a fine scale over large landscapes. Nevertheless, although research during this period explicitly recognized the possibilities for site-specific fertilization, the focus remained on refining grid sampling strategies and interpolation (Franzen & Peck, 1995; Mallarino, 1996; Penney, Nolan, McKenzie, Goddard, & Kryzanowski, 1996).

Once the means for describing and manipulating spatial variation in soil properties were available, subsequent research provided smaller iterations on the core geostatistical concepts. However, a disconnect was apparent between classical and geostatistical approaches: no clear methodology existed for determining the advantages and disadvantages of each or when each should be used. To fill this need, de Gruijter and ter Braak (1990) and Brus and de Gruijter (1997)

published two theoretical papers (see Wang et al. [2012] for a review) that categorized spatial soil sampling into "design-based" (classical) or "model-based" (geostatistical) approaches. The next section discusses the steps required to create a soil sampling plan and how these two approaches can help guide their development. Concurrently, other soil sampling research is reviewed that has implications for the timing, layout, and density of soil sampling.

4 | DEVELOPMENT OF A SOIL SAMPLING PLAN

4.1 | Goals

The core foundation of a soil sampling plan consists of clearly defined goals for nutrient management. The most common goals are to apply fertilizer and manure in a manner that increases profits by reducing costs and increasing yields while minimizing loss of nutrients to the environment. Depending on a farmer's ability to detect, interpret, and take advantage of spatial differences in nutrient availability in a cost-effective manner, the goal of increased profitability and reduced loss of nutrients to the environment could involve whole-field or site-specific management.

Another goal for nutrient management that is becoming more common is to meet requirements for regulatory compliance, most frequently with respect to P. As currently defined by state regulatory agencies, only whole-field estimates of nutrient concentrations are required for compliance, for example as an input to the Phosphorus Index (Lemunyon & Gilbert, 1993; Sharpley et al., 2003). As explored below, although a whole-field estimate is required in most cases, site-specific estimates may be advantageous in some situations.

4.2 | Defining the population

Regardless of the management goal, if the management scale is the whole field, then the population to be estimated or predicted is the entire field. If the management scale is site-specific, then the population to be predicted is the entire set of field management units (grid cells or zones). With this defined, it is possible to determine the sampling population, sampling units, sample size, methods, and locations of sampling. Each of these parameters is inter-related, and decisions about each parameter are affected by the choice of a design- or model-based approach to spatial sampling. This choice is described and formalized with a decision tree containing nine criteria in Brus and de Gruijter (1997). We refer the reader to the full outline of the decision tree in Brus and de Gruijter (1997) and instead enumerate below the appropriate approach for the common agronomic goals described above.

4.3 | Choosing a sampling approach

4.3.1 | Sampling for regulatory compliance: Design-Based

For situations in which strict objectivity and lack of bias is required for the sampling design (which excludes systematic sampling and subjective "representative" sampling), the only appropriate approach is design-based sampling. This approach uses the classical statistical concepts of randomization and stratification (Carter and Gregorich, 2007), where inference is entirely based on the sampling design and independence induced by the random selection of sampling units (and does not rely on spatial independence of the underlying phenomenon). Design-based sampling allows bias to be eliminated or accounted for when equal-probability sampling, unequal-probability sampling, or design-based weights are used (de Gruijter and ter Braak, 1990). Accounting for bias is likely to be most important when sampling for regulatory compliance.

4.3.2 | Sampling for whole-field management: Design- or model-based

Outside of a regulatory environment, the decision to use design- or model-based sampling depends on several factors. For the goal of whole-field estimation of spatial means and/or variance, design-based sampling is indicated where the sample size is adequate and spatial autocorrelation is weak. If spatial autocorrelation is strong, a model-based inference can increase the efficiency of sampling. Unfortunately, the distinction between strong and weak is not well defined (Brus & de Gruijter, 1997).

Unlike design-based sampling, the model-based approach does not require spatial independence (random selection) of sampling locations. Systematic or grid-based sampling is typically used for this approach, and independence is introduced by assuming that the observations come from one realization of the underlying random process (the superpopulation). The model-based approach additionally assumes intrinsic stationarity (Schabenberger & Gotway, 2005). In many situations, it also assumes that the covariance is direction independent, or isotropic.

4.3.3 | Sampling for site-specific management: Design- or model-based

For the goal of sub-field or site-specific estimation, one must first determine whether management zones or grid-cells/points are desired. If quantities in management zones are to be estimated, then a design-based approach is indicated,

which typically requires an adequate number of samples within each zone. If all these conditions hold but it is not possible to obtain enough samples, then a model-based approach is indicated. If prediction is desired for grid cells or points, then a model-based approach is indicated in nearly all situations.

4.4 | Sampling layout: sample location, size, and layout

A design-based approach guides the practitioner to use randomization, stratification, and probability-based sampling to choose the sampling layout, whereas a model-based approach suggests systematic sampling. Many references exist to guide a design-based approach; SRS is a mature practice. Stratified random sampling is simply an extension of SRS; yet, in practice the delineation of strata can be exceedingly difficult. For both sampling methods, points must be randomly selected for the whole field or within each stratum.

4.4.1 | Design-Based approach

Defining strata

Choosing the optimal number of strata is a trade-off between increased precision and increased cost and complexity (Lohr, 2010). This trade-off is difficult to navigate in agricultural settings because it is not clear which spatial data should be useful for stratification. Soil map units are the most intuitive data layer to use for stratification but may not provide improved estimates over SRS for the whole field (Mueller, Pierce, Schabenberger, & Warncke, 2001). Topography provides an appealing alternative to soil map units, yet sampling within topographic zones can provide inconsistent results across years and is only advised where manure has not been applied, fertility is relatively low, and mobile nutrients are of interest and when yield and remote sensing data are correlated to topography (Franzen, Cicacek, Hofman, & Swenson, 1998). Where multiple data layers are available, it may be advantageous to use Latin Hypercube Sampling (Minasny & McBratney, 2006) for stratifying the feature space. This method algorithmically generates samples from the multivariate distribution of available spatial data. In the context of soil sampling, the distribution could be comprised of soils data, remote sensing imagery, yield monitor measurements, or other ancillary data that are difficult to visualize in multidimensional space and challenging to efficiently stratify using traditional methods.

For regulatory purposes, stratification may be advantageous to account for areas of fields that may be at higher risk of nutrient loss or areas that can receive applications of a nutrient at a higher rate. In the case of P, critical source areas may only occupy a small portion of the landscape while generating a significant amount of the P runoff and leaching. Areas where manure has been applied, that are high in soil-test P above agronomic or environmental critical levels (McCormick, Jordan, & Bailey, 2009), or that have high levels of erosion (Sharpley et al., 2003), all of which may define a critical source areas, may be inadequately sampled using the SRS approach.

Management zones (MZs) are similar to strata and are frequently used to define distinct areas for site-specific management. Whereas the intention with defining strata is to approximate the underlying soil property of interest, the intention with MZs is to define locations where management intervention can have a clear effect on crop responses (Taylor, McBratney, & Whelan, 2007). As a result, MZs may be spatially similar to or divergent from classical strata and hence may not be consistent with the requirements for strict objectivity. Many layers have been used for stratification, including farmer observations (Taylor et al., 2007), electrical conductivity measurements (Peralta & Costa, 2013), previous years' yields (Flowers, Weisz, & White, 2005), and remotely sensed spectral measurements (Song et al., 2009). A full review of MZ literature is out of the scope of this paper, yet we note that the results for performance of MZs are mixed depending on the data layers used, agronomic history, and biophysical setting (Flowers et al., 2005; Mzuku et al., 2005; Sawchik & Mallarino, 2007; Schepers et al., 2004; Vasu et al., 2017).

Choosing the sample size

When a design-based approach is used to estimate whole-field or within-stratum spatial means and variances, determining the appropriate sample size is important for achieving the desired precision and accuracy. The required sample size for an acceptable margin of error when no spatial autocorrelation is present can be calculated as follows:

$$\text{Sample size} = \left(\frac{t_{1-\alpha/2} \times \sigma}{\text{ME}} \right)^2 \quad (3)$$

where $t_{1-\alpha/2}$ is the t-statistic for a desired precision level α and degrees of freedom equal to the input sample size minus one, σ is the standard deviation, and ME is the margin of error (Cochran, 1977). The inputs are derived from a pilot dataset or data from an analogous location. This assumes that no finite population correction is necessary because the population of possible sample locations is nearly infinite and that the distribution of nutrient concentrations is approximately normal. The margin of error is often calculated as a percentage of the mean.

If moderate to strong spatial autocorrelation exists in the field, which is often the case, the sample size can be reduced

if the aim is to estimate the mean. In such cases, the sample size is calculated as:

$$n = n_{\text{classic}} \left[1 - \frac{\text{cov}(y_i / y_j)}{\sigma^2}\right] \quad (4)$$

where n_{classic} is the previously specified sample size, $\text{cov}(y_i/y_j)$ is the covariance between all possible pairs of points, and σ^2 is the dispersion variance of the population, equal to $\text{E}[y - \text{E}(y)]^2$ (Wang, Stein, Gao, & Ge, 2012), where E[] represents the expected value of articles contained within the brackets.

For most fields, local data to estimate the sample size will not be available; therefore, data from other studies must be used. Herein we synthesize articles collected during the literature review to estimate appropriate sample sizes and densities for our soil chemical properties of interest. Individual observation data needed to account for spatial autocorrelation were not provided with the selected research. Consequently, we used the classical sample size estimator presented in Eq. (3). The basic characteristics of each study are presented in Supplementary Tables S1 and S2.

Only four of the catalogued studies used random sampling, which is required in a design-based approach, as a means of data collection. Our collective 60 years of experience with soil sampling for nutrient recommendations leads us to the conclusion that random sampling is used most frequently when soil samples are collected for nutrient recommendations. Some of the increase in the number of soil samples collected for nutrient recommendations reported in 831,000 soil samples in the 2015 report "Soil Test Levels in North America" (IPNI, 2015) was assumed to be the result of grid sampling or systematic sampling to describe management zones, which may have had the objective to describe soil spatial variability. However, the large majority of the samples were thought to be collected for whole-field nutrient analysis, which is typically done using SRS (IPNI, 2015).

The data collected for this study, therefore, are mostly geostatistical in nature, with the objective to use applied model-based inference, and therefore used versions of systematic grid-based sampling. Despite the biased nature of systematic sampling, we aggregated the data from individual fields in each study to obtain an estimate for the range of soil sampling densities that would be required to achieve a margin of error of 10% of each study's mean (Figure 1) at a 5% precision level. For NO_3–N, P, K, SOC/SOM, and pH, the number of qualifying studies was 16, 24, 19, 14, and 17, respectively, which represented 35, 88, 75, 30, and 64 fields. Multiple fields were allowed from each study, but repeated observations from each field were not. In such cases, the first observation that occurred in the fall was retained. In situations where data were reported as lognormal, we calculated the number of samples required to compute the geometric mean.

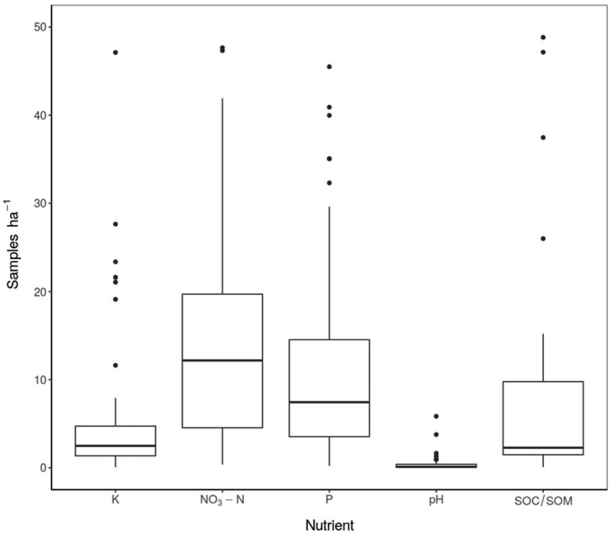

FIGURE 1 Boxplot of estimated required sampling densities for (a) phosphorus (n = 24 studies; 88 fields), (b) NO_3–N (n = 16; 35 fields), (c) potassium (n = 19; 75 fields), (d) soil organic carbon (SOC)/soil organic matter (SOM) (n = 14; 30 fields), and pH (n = 17; 64 fields); NO_3–N had four outliers between 50 and 6361, P had nine outliers between 50 and 8523, K had one outlier at 300, and SOC/SOM had one outlier at 730

Nitrate–N and P require the highest median number of samples per hectare (15 and 8.4, respectively, excluding outliers), and pH requires the least, with only 0.1. As the median sample density increased, the width of the distribution increased as well. Depending on the overall field size, these sampling densities could be time and cost prohibitive to collect to achieve a 5% precision level: for small fields only 4 ha (10 acres) in size, collecting 30 samples for P would be too expensive, and even more so for a 40-ha (100-acre) field. Furthermore, the wide distributions of required sample sizes suggest that, depending on the nutrient of interest and an individual field, many more samples would be required to achieve the desired level of statistical accuracy.

Fortunately, the required sampling density decreases with increasing field sizes (Figure 2), enabling an exponential decay function to be fit (SS sample size $= \alpha \times e^{(-k \times \text{field size})}$, where α and k are shape parameters). For example, instead of requiring 300 samples if a linear function was used for a 40-ha field for P, a sample density of only 2.4 samples ha^{-1} is required, equating to 96 samples total. To derive the exponential decay functions, outlier sample size densities over 100 ha^{-1} (13 of the 292 data points) were omitted.

The ability to estimate sample size requirements based on field size could be valuable because in many cases the site-specific nutrient variability, which is a better indicator of sampling requirements (Clay, Carlson, Brix-Davis, Oolman, & Berg, 1997), is unknown. However, the relationship between required sampling densities and field size is heavily affected by a small number of large-field observations and,

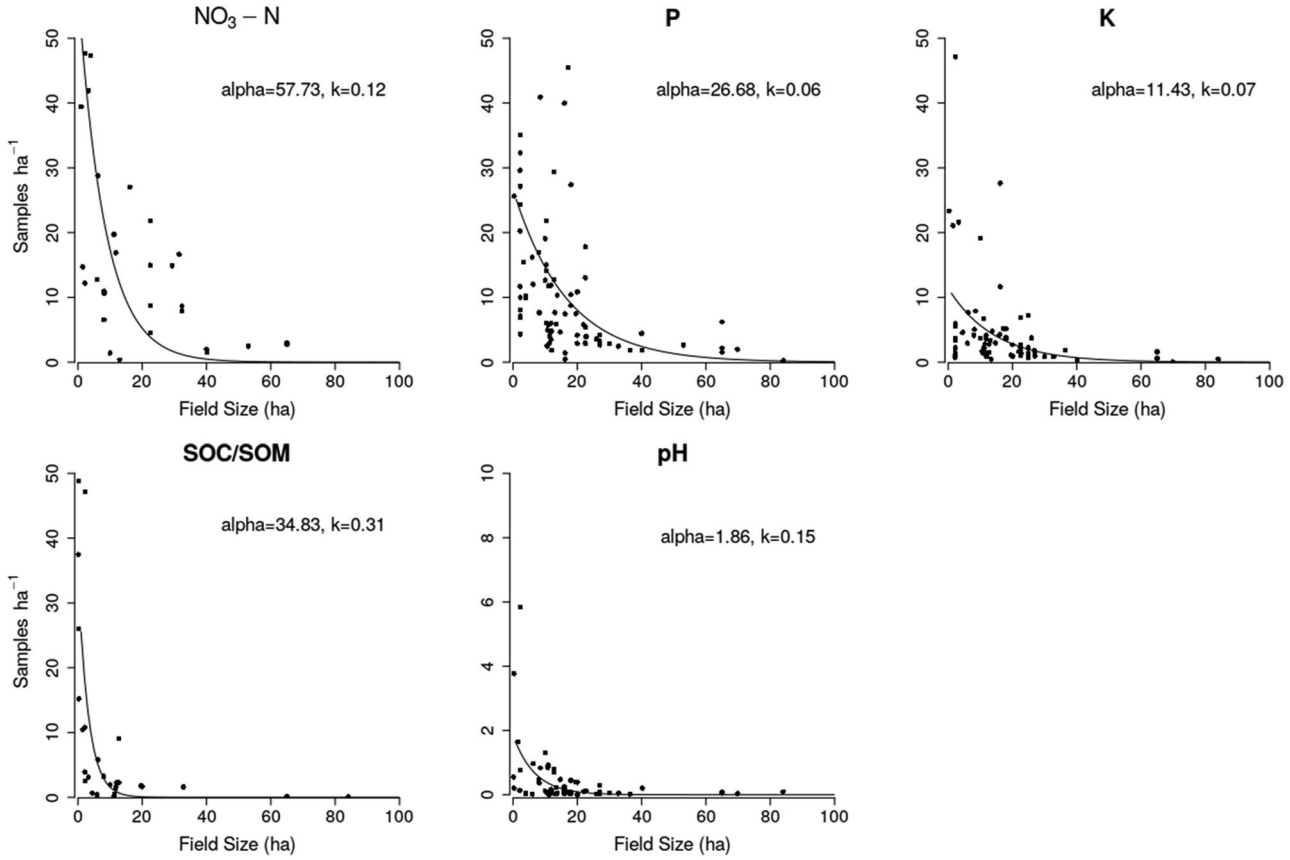

FIGURE 2 Effect of field size on required sampling density for (a) phosphorus (n = 24 studies; 88 fields), (b) NO_3–N (n = 16; 35 fields), (c) potassium (n = 19; 75 fields), (d) soil organic carbon (SOC)/soil organic matter (SOM) (n = 14; 30 fields), and pH (n = 17; 64 fields). Points represent the required sampling density calculated from single fields in each study. Functions fit with field sampling densities >100 samples ha^{-1} were removed for clarity

in the case of NO_3, a few observations of small fields that displayed high variability. Therefore, although these limited data suggest there may be a decline in sampling requirements as field sizes increase, more data would assist in confirming this hypothesis.

Compositing and log-normality

Compositing soil samples before analysis is a common method for reducing soil analysis costs and is sensible when sampling distributions are normal. However, for the 13 studies that explicitly tested for skewness and kurtosis, 10 found at least one or more chemical properties that deviated from normality. Given this high percentage, compositing, which effectively calculates the mean of a number of soil cores, is unlikely to represent the true population parameter of interest and will typically present a higher estimate of nutrient concentrations than is accurate.

In locations where "hot spots" of nutrients are encountered, the distribution of a nutrient would be positively skewed, which would lead the farmer to infer that the test value of the composited soil property is high. For fertilizer management, this may cause the farmer to apply an inadequate amount of fertilizer for optimum crop growth. In a regulatory context, the farmer may be penalized for excess nutrient concentrations.

To avoid this outcome, it is important to have prior evidence that a nutrient will be normally distributed in a field before assuming that compositing will adequately represent the average nutrient concentration. Where log-normality is present, farmers may implement some type of systematic grid-based sampling with variable-rate fertilization to avoid over-application in hot spots and under-application in areas with lower nutrient concentrations. If variable-rate fertilization is not possible, then a simple outlier-resistant measure of central tendency, such as the median (or, equivalently, the geometric mean), may be used to choose the uniform fertilizer rate.

4.4.2 | Model-Based approach

Defining grids

The configuration and density required for systematic grid sampling is dependent on the strength and structure of spatial variation, which in turn are described by the semivariogram or correlogram. We refer the reader to key geostatistical

references for detailed descriptions of the semivariogram and its use in interpolation (Burgess & Webster, 1980; Cressie, 1993; Hengl, Rossiter, & Stein, 2003; Kerry, Oliver, & Frogbrook, 2010; Matheron, 1963; Oliver & Webster, 2014; Schabenberger & Gotway, 2005) and instead describe the methods used for acquiring grid data, the resulting spatial structures, and the implications for future sampling.

The grid sizes used to assess spatial variability of soil chemical properties varied from 1 to 100 m in the 39 assessed studies that collected grid-based data. Most samples were collected at the centroids or corners of the grids, although several studies collected samples spread throughout grid cells, referred to as the "grid-cell" method. Grid-cell sampling has mixed results because it is sometimes more efficient for estimating nutrient concentrations (Chang, Clay, Carlson, Clay, & Malo, 2003; Mallarino & Wittry, 2004), but in other instances it is no more effective than a whole-field approach (Mueller et al., 2001).

To effectively use a model-based approach that relies on grid sampling, the strength of spatial autocorrelation must be moderate to strong (Kravchenko, 2003). Cambardella et al. (1994) proposed assessing the degree of spatial dependence using the ratio of the semivariogram nugget to the sill, a metric that was used in a later study by Cambardella and Karlen (1999) and subsequently adopted by other researchers (Kravchenko, 2003; Mueller, Pusuluri, Mathias, Cornelius, & Barnhisel, 2004). This metric may be calculated using pilot data, extrapolated from other studies, or estimated using ancillary data (Kerry & Oliver, 2008). Ratios <0.25 suggest strong spatial dependence, 0.25 to 0.75 suggest moderate spatial dependence, and >0.75 suggest weak spatial dependence. Figure 3 displays the spatial dependence of NO_3–N, P, K, SOC/SOM, and pH for the assessed studies.

The strength of spatial dependence is extremely inconsistent between studies; however, P, pH, and SOC/SOM seem to be more strongly spatially dependent than NO_3–N and K. Given its lower degree of spatial dependence and high mobility in the soil, grid-based sampling of NO_3–N is unlikely to be advantageous for whole-field or site-specific management. Plant mineralizable N, NH_4–N, Zn, total N, Mg, Al, S, and Ca had average spatial dependencies of 0.85, 0.55, 0.5, 0.14, 0.13, 0.13, 0.12, and 0.08, respectively, but had wide ranges and low accuracy because only one to nine individual values were recorded for each property.

In contrast to NO_3–N, P has much lower mobility in fertility-limited soils and a high degree of spatial dependence; hence, P is likely to be a better candidate for grid-based sampling. Several factors may amplify the spatial variability of P. Hot spots from manure application or old homesteads can either increase (Cambardella & Karlen, 1999) or decrease (Grandt, Ketterings, Lembo, & Vermeylen, 2010) spatial variability. Surficial or deep-banding of P fertilizer can increase lateral (Kitchen, Westfall, & Havlin, 1990) and vertical stratification (Rehm, Scobbie, Randall, & Vetsch, 1995) of P and K, although the patterns of stratification depend on the tillage system used. In no-till systems, banded applications may result in cyclical patterns perpendicular to rows (Mallarino, 1996). To guard against varying degrees of lateral and vertical P concentration when grid sampling in banded fields, within-band samples should be collected in proportion to the area of the fields the bands occupy (Kitchen et al., 1990) to a depth of 15 to 20 cm (Mallarino & Borges, 2006), or unaligned systematic grids should be used (Wollenhaupt, Wolkowski, & Clayton, 1994). If the location of the band is not known, samples can be collected in pairs separated by a distance of half the band width (Kitchen et al., 1990).

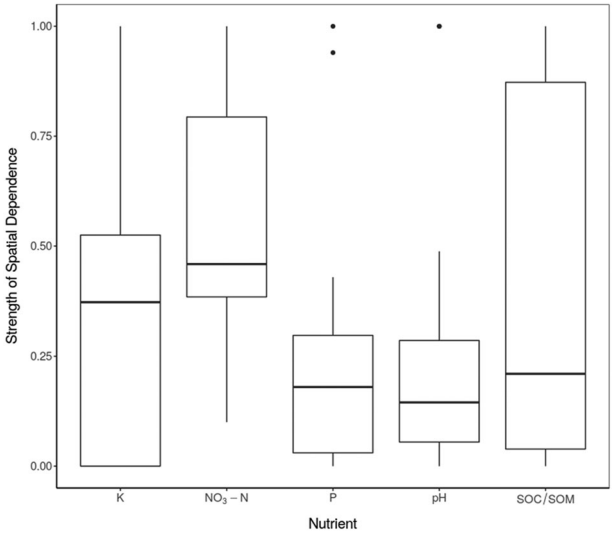

FIGURE 3 Strength of spatial dependence of nutrients for all studies for (a) phosphorus (n = 13 studies; 22 fields), (b) NO_3–N (n = 9; 15 fields), (c) potassium (n = 9; 15 fields), (d) soil organic carbon (SOC)/soil organic matter (SOM) (n = 9; 12 fields), and pH (n = 10; 16 fields); <0.25 = strong spatial dependence, 0.25 to 0.75 = moderate spatial dependence, >0.75 = weak spatial dependence

For other chemical properties, there is little information on the effect of various agronomic practices on the strength of spatial dependence. Tillage may increase the distance over which soil chemical properties are correlated (Robertson, Crum, & Ellis, 1993) as compared with uncultivated grassland sites, but it may be difficult to make further generalizations about spatial structure in some fields because periodic trends and nested scales of variation may obscure true spatial patterns (Cahn et al., 1994; Mallarino & Borges, 2006).

Across all studies, NO_3–N generally had the smallest median range of spatial correlation, whereas P had the largest median range of spatial correlation (Figure 4). When collecting grid samples for multiple chemical properties, the property with the smallest range will likely strongly influence the required grid size, although co-kriging can be

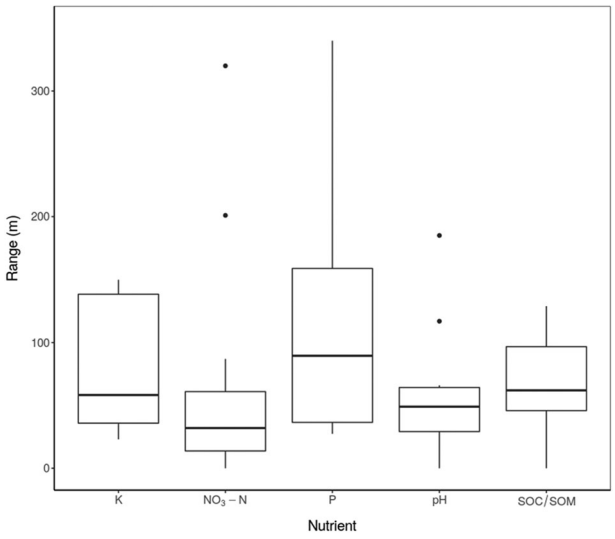

FIGURE 4 Range of spatial correlation for (a) phosphorus (n = 11 studies; 17 fields), (b) NO_3–N (n = 8; 13 fields), (c) potassium (n = 7; 11 fields), (d) soil organic carbon (SOC)/soil organic matter (SOM) (n = 3; 6 fields), and pH (n = 6; 11 fields). Studies that did not exhibit a sill were excluded from this graph. One study-field each for K and SOC/SOM displayed only a nugget effect (no spatial autocorrelation) and were also excluded

used to augment data for soil properties only available at lower resolutions (McBratney & Webster, 1983).

To decide on the final grid size to use for sampling, the range of spatial correlation can help guide the sampling density, but the sampling design ultimately depends on the predefined objectives. Three objectives are typically specified as (i) minimizing the estimation error variance from kriging, (ii) achieving equal spatial coverage, or (iii) achieving equal coverage in feature space (Wang et al., 2012), where the feature space is defined as the joint multivariate distribution of spatial variables from a field. The first objective requires an assumed variogram to optimize spatial allocation of sampling points.

McBratney, Pringle, and McBratney (1999) developed a decision tree for collecting pilot samples to optimize spatial allocation of sampling points and a proposal for averaging variograms from other locations for use in the target location. Once a variogram has been obtained or assumed, grid points may be placed and then algorithmically shifted to minimize the objective function using simulated spatial annealing (van Groenigen, 2000; van Groenigen, Siderius, & Stein, 1999; Vašát, Heuvelink, & Borůvka, 2010). Alternatively, because it is most important to have high number of point pairs at small lags and to achieve other distance-angle dispersion criteria, algorithmic options are available for choosing sampling locations (Warrick & Myers, 1987).

If equal spatial coverage is desired and a variogram is not available, it is possible to optimize sampling locations with the goal of minimizing the distance from unsampled points to sampled points (van Groenigen et al., 1999; Wang et al., 2012). For achieving equal coverage in feature space, latin hypercube sampling is another option (described in Design-Based Approach: Defining Strata) (Minasny & McBratney, 2006).

Another consideration in determining the grid size when nutrient recommendations are desired is the economic optimum sample density (EOSD). In agricultural settings, the EOSD is dependent on the net economic outcome of increasing or decreasing sample numbers. The agronomic factors that influence the EOSD are the accuracy and precision of the economically optimum nutrient rate (Cerrato & Blackmer, 1990; Morris et al., 2018), which includes (i) the applicability of the economically optimum nutrient rate to the field location of interest, (ii) the rate at which uncertainty is reduced as more samples are collected, (iii) prior information about nutrient concentrations available from the location of interest that can inform sampling decisions, and (iv) other genetic, environment, and management factors (Hatfield & Walthall, 2015) that affect yield responses to N. Economic factors that influence the EOSD include the relative change in net revenue from additional N availability from manure or fertilizer, the fixed and variable costs of sampling, crop and nutrient prices, and the cost of variable rate fertilizer application if it is used to manage spatial variability. When all of these factors are considered, which would require considerable time and expense, the EOSD can be calculated as the value where the marginal costs of increased sampling equal the marginal benefits from the increased sampling. Increasing the sampling density beyond the EOSD will not reduce uncertainty enough to justify the increased spending on labor, equipment, and laboratory analysis fees. No literature was found that derived the EOSD in agricultural settings, but the EOSD is an important metric that merits attention in future research.

Once the grid size and configuration has been determined, it is important to check additional assumptions before proceeding with sampling or analysis. If anisotropy is expected, grid sampling should be modified, for example, by adjusting the sampling length along or across rows by the correlation length multiplied by the ratio of directional variability (i.e., Cv_x/Cv_y) (Gupta, Mostaghimi, McClellan, Alley, & Brann, 1997). Once data have been collected, it is critical to check the stationarity (i.e., constant spatial mean) and isotropy assumptions of model-based inference. In the studies assessed here, only 16% checked for stationarity and 13% checked for anisotropy. Correcting for nonstationarity is possible via median-polish kriging (Cressie, 1993; Mohanty & Kanwar, 1994) or by estimating the trend and variogram of random residuals simultaneously (Lark, 2012; Lark, Cullis, & Welham, 2006; Webster & Oliver, 2007), and directional semivariograms may be used to account for anisotropy (Mueller et al., 2001).

4.5 | Limitations of cross agro-environmental comparisons for design- and model-based inference

A high degree of variability was observed in the ranges of spatial correlation, strengths of spatial dependence, validation of assumptions, and required design-based sample sizes across all studies. Considering the vast array of potential environmental, agronomic, sampling, and statistical confounding variables for each study, this uncertainty is not surprising and should give practitioners pause before extrapolating findings to new locations.

Research on spatial variability, especially at the field scale, cannot isolate confounding factors, as is possible in small-plot experiments. In fact, isolation would undermine the purpose of applied spatial studies because it would largely eliminate possibilities for spatial extrapolation. The alternative, however, is that many spatial soil sampling studies discover spatial patterns that are at best only applicable in a hyper-localized area, as discovered by Lauzon et al. (2005) across 23 farm fields.

Several developments would greatly aid the ability to extrapolate from single studies or to merge results from multiple studies, as suggested by the average variogram approach (McBratney et al., 1999). First, comprehensive background data collection on soil properties, cropping history, fertilization history, and climate and an exhaustive list of soil sampling protocols would help practitioners select research that was sufficiently similar to their own context. Second, a systematic research effort needs to be aimed at identifying the effect of various confounding variables on the spatial structure of soil chemical properties, isolating them one or several at a time. Only then can practitioners know whether those variables can be safely ignored when applying variograms from disparate environments. Third, for specific combinations of geographic confounding variables (e.g., soils and climate), systematic attempts to enumerate and sample spatial variability of nutrients would help identify the uncertainty of variogram parameters. Finally, if researchers were to propose a discrete scope of inference for each study, it would greatly aid practitioners in knowing whether the required sample sizes and/or variogram properties could be extrapolated to their own agro-environmental context.

5 | CURRENT STATE OF EXTENSION SOIL SAMPLING RECOMMENDATIONS

Extension recommendations necessarily must provide a more concrete answer to practitioners on how to collect soil samples in the presence of spatial variability compared with peer-reviewed academic literature.

The literature search on peer-reviewed, extension-oriented literature uncovered only five articles about spatial configuration of soil sampling. Due to this small volume of literature, it was not possible to infer whether public-oriented extension literature was consistent with the academic understanding of extension practitioners.

Production agriculture–focused soil sampling recommendations were found from only 39 of the land grant university extension systems in the 50 US states. The vast majority (87%) of those provided suggestions for forming MZs, and none mentioned stratification for design-based sampling (Supplemental Table S3). Of those suggesting the use of MZs, the recommended sampling densities ranged from 0.6 to 5 samples per hectare (0.25–2 samples per acre) (Supplemental Table S3), which is substantially below the median required number of samples for NO_3–N and P and above the median number for other soil properties at a precision level of 5% (Figure 1). For lower levels of precision, the recommended sampling densities may be more consistent with the peer-reviewed literature. The associated maximum recommended MZ size ranged from 2 to 16 ha (5–40 acres), and none recommended reducing sampling intensities for larger areas. To delineate MZs, a wide variety of data layers were suggested, including soil type, cropping history, erosion, slope, aspect, soil texture, soil depth, remote sensing, crop productivity, and past management. In the simplest case, one or two data layers may be aggregated and delineated by hand, but a quantitative approach and the incorporation of more data layers requires a Geographic Information System to manage, integrate, display, and derive benefit from the data.

To collect the samples within MZs or whole fields, 54% of extension sources suggested “zig-zag,” “Z,” or “W” sampling, with 44% mentioning the need to take “representative” samples (Supplemental Table S3). This suggestion is at odds with the requirements of a design-based approach in which random selection of locations is paramount. As a result, it is highly likely that these methods often result in biased results, especially if the samples are composited before analysis and the soil property is lognormally distributed.

For the purposes of site-specific agricultural management, 16 (40%) extension sources recommended the option of grid sampling. Of those, nine recommended point sampling, three recommended grid-cell sampling, and three recommended point or grid-cell sampling. Only five sources mentioned detailed suggestions to account for fertilizer banding; several more mentioned avoiding bands altogether when sampling. For no-till, many recommended taking samples closer to the surface (0–5 cm or 0–2 in.) in addition to a full-depth sample (typically 0–15 cm or 0–6 in.) to account for a lack of deep incorporation (Supplemental Table S3).

6 | RECOMMENDATIONS

Collecting a large number of soil cores is expensive and time consuming. Therefore, it is important to select a sampling strategy that is efficient and consistent with the goals of the practitioner or researcher. To aid in these decisions, we distilled current and historical research into a single decision tree (Figure 5) with nine possible decision points, described below.

1. Sampling for regulatory compliance?

If the intent of sampling is to comply with regulations, strict objectivity is required (see section Sampling for Regulatory Compliance: Design-based). This requirement can be fulfilled with a SRS design or stratified random sampling design.

2. Whole or sub-field management?

If one suspects that there is substantial sub-field variability that can be manipulated and that would be profitable to manage site-specifically, then sub-field management should be considered.

3i. Strong spatial autocorrelation?

The strength of spatial autocorrelation is defined as the ratio of the nugget to the sill of the semivariogram (see section Model-Based Approach: Defining Grids), which is often defined as >0.75 when strong (Cambardella et al., 1994), although "strong" spatial autocorrelation is not well defined and varies from paper to paper. This information should be derived from previous studies at the same location or high-quality spatial data collected from an analogous agronomic system. If the spatial autocorrelation is not known, then the assumption should be that there is no strong spatial autocorrelation.

3ii. Management scale: points/cells or management zones/strata?

If the goal is to manage on a sub-field scale and at a high spatial resolution, then this is affirmative. If the intention is to manage at an aggregated resolution (i.e., management zones), then the answer is no.

3iii. Suspected groupings of field areas with respect to soil property of interest?

Based on visual observations, soil maps, yield monitor data, or other ancillary spatial data sources, one may suspect that the soil property of interest is defined by homogeneous areas with discontinuous boundaries. If so, stratification should be considered.

4i. Are prior spatial data available?

Previously collected spatial data that are of high quality and from an analogous agronomic system (cropping history, soils, climate, fertilization history, etc.) can provide valuable information on the density and configuration required for model-based sampling.

4ii. Are spatial data layers available that are correlated with groupings?

Related to Decision Point 3iii, observations or spatial data layers must be available that are correlated with the soil property of interest to create strata. For example, if P is a yield-limiting factor, then spatial yield data may be used to derive strata for P. Similarly, if significant runoff has transported applied fertilizer to toe-slopes, then topographic data may be associated with available P and can therefore be used for stratification.

5. Is the goal to create baseline spatial data?

Without the availability of previously collected spatial data, it is important to explicitly recognize whether the sampling activity is intended to form a baseline dataset for future work. When creating a baseline, it is important to anticipate that the soil property may vary on a scale much smaller than the chosen sampling resolution, reducing the practical utility of the survey.

6. Costs of grid sampling are affordable?

Based on previously collected spatial data and/or a default resolution for baseline data collection of at least 100 data points for any field size (Webster & Oliver, 1992), it is straightforward to estimate the required sampling density and costs. If these are not manageable, then design-based sampling should be used.

6.1 | Suggestions for research applications

If these conditions are fulfilled and the sampling is performed for research purposes, we suggest the investigators catalog and publish the following minimum data:

1. Soil physical properties: soil order, top horizon depth, organic matter, texture, pH, anomalous soil properties (e.g., highly calcareous).
2. Agronomic history: 10-yr cropping and tillage history, 5-yr fertilization/manure history, significant presence of weeds with an estimated percentage infestation if possible, use of irrigation, presence of tiling, any other anomalous historical artifacts (e.g., fertilizer spills, stacking of manure).
3. Climate: average precipitation and temperature, growing-season precipitation.
4. Soil sampling: depth of sampling, probe diameter, number of fields, field size, size of sampling area, configuration of sampling, compositing parameters (number of cores, area sampled, configuration), month of sampling and relation to

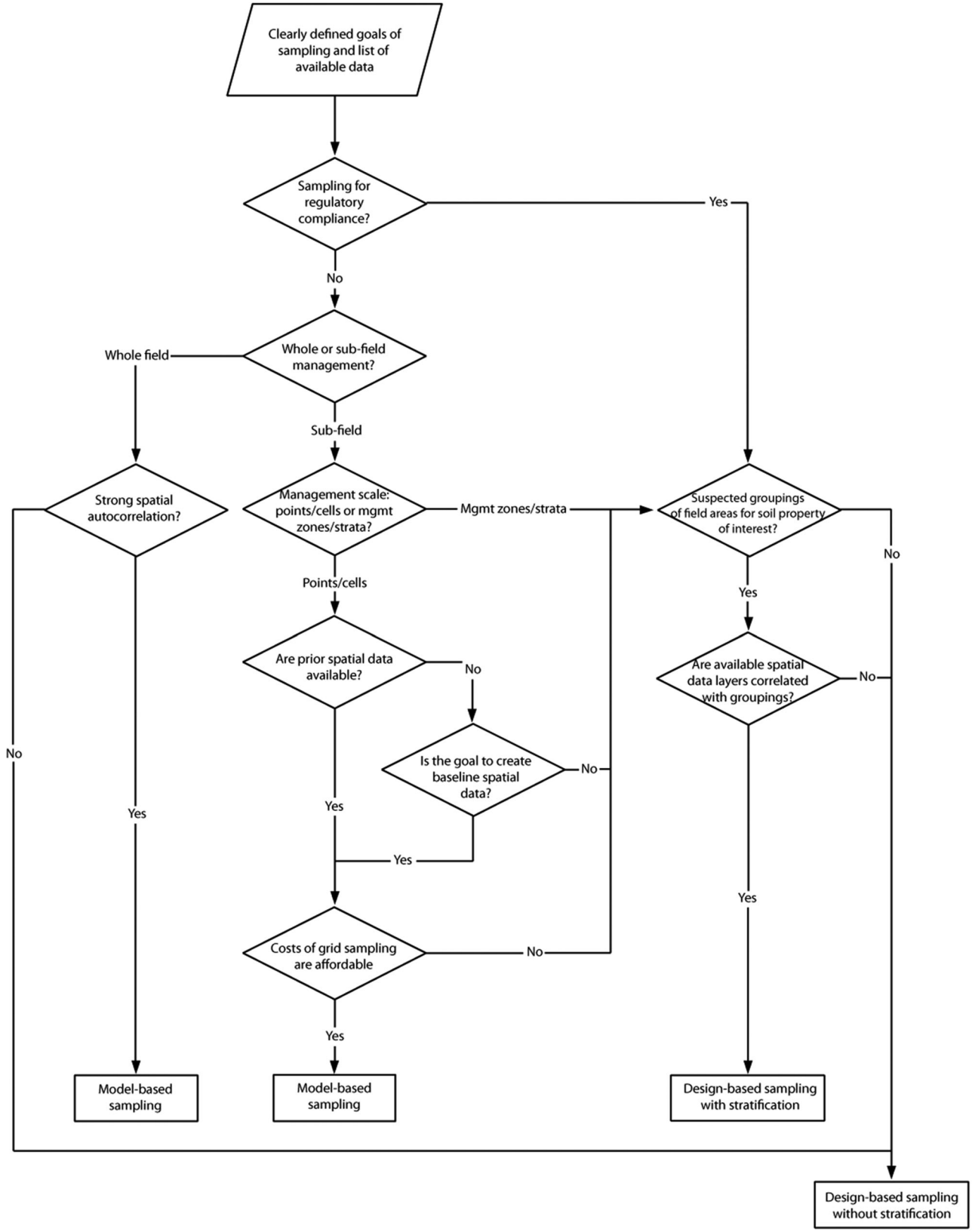

FIGURE 5 Decision tree to guide soil sampling for researchers and practitioners in selecting a sampling strategy

field operations (e.g., pre-fertilization, post-harvest), grid spacing, total number of sampled points.
5. Statistical properties: ideally the raw data would be available online because the costs of data storage are negligible. If this is not possible, researchers should provide the means, standard deviations or variance, sample size, and coefficient of variation. For spatial studies, this would also include semivariogram parameters (model form, nugget, sill, and range) and confirmation that assumptions about log-normality, stationarity, and isotropy were validated.
6. Scope of inference: For the use of other researchers or practitioners, to what set of conditions could this research be reasonably extended?

6.2 | General suggestions for design-based sampling

1. If no prior information on the soil property is available, log-normality should be assumed, and soil samples should not be composited. If the collected cores display log-normality, the geometric mean or median should be used instead of the arithmetic mean. Only in instances in which normality has previously been established should soil samples be composited.
2. Reduce the number of cores collected per unit area as the field size increases. The number of cores will always need to be balanced with cost constraints.

7 | CONCLUSION

The inherent spatial variability of soil chemical properties at multiple scales is challenging to characterize, even when significant resources are available for sampling and analysis. Statistical methods to estimate and predict those properties have become increasingly sophisticated, yet unclear sampling goals, conflicting information, and lack of structured data presentation have hampered the development of a comprehensive knowledge base on soil variability in agricultural landscapes. To make progress in soil sampling, it is important to describe site properties in detail, explore the effect of confounding variables, quantify spatial uncertainties, and define the scope of interest for each individual study.

Given the slow pace of advancements in knowledge about spatial variability of soil nutrients from soil sampling research, it is not surprising that extension recommendations often suggest practices such as compositing and Z-sampling, which do not have a strong foundation in peer-reviewed literature. Fortunately, there are theoretically simple ways to correct these recommendations, such as suggesting soil cores to be individually analyzed. However, the results from this review indicate that from a practical standpoint most fields likely will be too expensive to sample if collection of individual cores is required to obtain a value reasonably close to the mean at a 5% precision level; a 10 or 20% precision level may be more affordable. Using software for randomly allocating points within fields, strata, or MZs also would greatly increase the accuracy of design-based sampling. Creating more confident recommendations on the data layers to use for MZs is more problematic, and the current ad hoc approach would benefit from more structure such as might be provided with decision trees.

Until the research community develops a more structured approach toward applying model-based sampling in previously unsampled locations, it is difficult to know whether specific sampling densities or configurations would result in positive net economic returns for a farmer. Design-based sampling methods, with smaller required sample sizes and the ability to minimize bias, currently outweigh the promise of site-specific management in most fields unless the costs for sampling are greatly decreased and benefits to precision applications of nutrients are greatly increased. The results reported in this study indicate that design-based sampling is more likely than model-based sampling to provide benefits for farmers and society through more profitable applications of fertilizer and manure, more efficient use of nonrenewable natural resources, and minimization of pollution.

ACKNOWLEDGMENTS

This material is based on work that is supported by the National Institute of Food and Agriculture, USDA, under award number 2017-38640-26915 through the Northeast Region SARE program under project number LNE18-363. USDA is an equal opportunity employer and service provider. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the USDA.

ORCID

Patrick G. Lawrence
https://orcid.org/0000-0002-2324-5866

REFERENCES

Bonmati, M., Ceccanti, B., & Nanniperi, P. (1991). Spatial variability of phosphatase, urease, protease, organic carbon and total nitrogen in soil. *Soil Biology & Biochemistry*, *23*(4), 391–396.

Borges, R., & Mallarino, A. P. (1998). Variation of early growth and nutrient content of no-till corn and soybean in relation to soil phosphorus and potassium supplies. *Communications in Soil Science and Plant Analysis*, *29*(17–18), 2589–2605. https://doi.org/10.1080/00103629809370136

Brus, D. J., & de Gruijter, J. J. (1997). Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). *Geoderma*, *80*(1–2), 1–44. https://doi.org/10.1016/S0016-7061(97)00072-4

Bogunovic, I., Mesic, M., Zgorelec, Z., Jurisic, A., & Bilandzija, D. (2014). Spatial variation of soil nutrients on sandy-loam soil. *Soil and Tillage Research*, *144*, 174–183. https://doi.org/10.1016/j.still.2014.07.020

Burgess, T. M., & Webster, R. (1980). Optimal interpolation and isarithmic mapping of soil properties: I the semi-variogram and punctual kriging. *Journal of Soil Science*, *31*(2), 315–331. https://doi.org/10.1111/j.1365-2389.1980.tb02084.x

Cahn, M. D., Hummel, J. W., & Brouer, B. H. (1994). Spatial analysis of soil fertility for site-specific crop management. *Soil Science Society of America Journal*, *58*(4), 1240. https://doi.org/10.2136/sssaj1994.03615995005800040035x

Cambardella, C. A., & Karlen, D. L. (1999). Spatial analysis of soil fertility parameters. *Precision Agriculture*, *1*(1), 5–14. https://doi.org/10.1023/A:1009925919134

Cambardella, C. A., Moorman, T. B., Parkin, T. B., Karlen, D. L., Novak, J. M., Turco, R. F., & Konopka, A. E. (1994). Field-scale variability of soil properties in Central Iowa soils. *Soil Science Society of America Journal*, *58*(5), 1501. https://doi.org/10.2136/sssaj1994.03615995005800050033x

Cameron, D. R., Nyborg, M., Toogood, J. A., & Laverty, D. H. (1971). Accuracy of field sampling for soil tests. *Canadian Journal of Soil Science*, *51*(2), 165–175. https://doi.org/10.4141/cjss71-023

Carpenter, S. R., Caraco, N. E., Correll, D. L., Howarth, R. W., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications*, *8*(3), 559–568. https://doi.org/10.1890/1051-0761(1998)008[0559:NPOSWW]2.0.CO;2

Carter, M. R., & Gregorich, E. G. 2007. *Soil sampling and methods of analysis* (2nd ed.). Boca Raton, FL: CRC Press.

Cerrato, M. E., & Blackmer, A. M. (1990). Comparison of models for describing corn yield response to nitrogen fertilizer. *Agronomy Journal*, *82*, 138–143. https://doi.org/10.2134/agronj1990.00021962008200010030x

Chang, J., Clay, D. E., Carlson, C. G., Malo, D., & Clay, S. A. (1999). Precision farming protocols: Part 1. Grid distance and soil nutrient impact on the reproducibility of spatial variability measurements. *Precision Agriculture*, *1*, 277–289. https://doi.org/10.1023/A:1009921024451

Chang, J., Clay, D. E., Carlson, C. G., Clay, S. A., & Malo, D. D. (2003). Different techniques to identify management zones impact nitrogen and phosphorus sampling variability. *Agronomy Journal*, *95*(6), 1550–1559. https://doi.org/10.2134/agronj2003.1550

Clay, D. E., Carlson, C. G., Brix-Davis, K., Oolman, J., & Berg, B. (1997). Soil sampling strategies for estimating residual nitrogen. *Journal of Production Agriculture*, *10*(3), 446–452. https://doi.org/10.2134/jpa1997.0446

Cline, M. G. (1944). Principles of soil sampling. *Soil Science*, *58*(4), 275–288. https://doi.org/10.1097/00010694-194410000-00003

Cochran, W. G. 1977. *Sampling techniques*. New York: John Wiley & Sons.

Cressie, N. (1985). Fitting variogram models by weighted least squares. *Journal of the International Association for Mathematical Geology*, *17*(5), 563–586. https://doi.org/10.1007/BF01032109

Cressie, N. A. C. 1993. Statistics for spatial data. Revised ed. New York: John Wiley & Sons.

Dahiya, I. S., Anlauf, R., Kersebaum, K. C., & Richter, J. (1985). Spatial variability of some nutrient constituents of an Alfisol from loess II. *Geostatistical analysis. Z. Pflanzenernaehr. Bodenkd*, *148*(3), 268–277. https://doi.org/10.1002/jpln.19851480307

de Gruijter, J. J., & Braak, ter C. J. F. (1990). Model-free estimation from spatial samples: A reappraisal of classical sampling theory. *Mathematical Geology*, *22*(4), 407–415. https://doi.org/10.1007/BF00890327

Dhillon, J., Guilherme, T., Driver, E., Figueiredo, B., & Raun, W. R. (2017). World phosphorus use efficiency in cereal crops. *Agronomy Journal*, *109*, 1670–1677. https://doi.org/10.2134/agronj2016.08.0483

Flowers, M., Weisz, R., & White, J. G. (2005). Yield-based management zones and grid sampling strategies. *Agronomy Journal*, *97*(3), 968. https://doi.org/10.2134/agronj2004.0224

Franzen, D. W., Cicacek, L. J., Hofman, V. L., & Swenson, L. J. (1998). Topography-based sampling compared with grid sampling in the Northern Great Plains. *Journal of Production Agriculture*, *11*(3), 364–370. https://doi.org/10.2134/jpa1998.0364

Franzen, D. W., & Peck, T. R. (1995). Field soil sampling density for variable rate fertilization. *Journal of Production Agriculture*, *8*(4), 568–574. https://doi.org/10.2134/jpa1995.0568

Goovaerts, P., & Chiang, C. N. (1993). Temporal persistence of spatial patterns for mineralizable nitrogen and selected soil properties. *Soil Science Society of America Journal*, *57*, 372–381. https://doi.org/10.2136/sssaj1993.03615995005700020015x

Grandt, S., Ketterings, Q. M., Lembo, A. J., & Vermeylen, F. (2010). In-field variability of soil test phosphorus and implications for agronomic and environmental phosphorus management. *Soil Science Society of America Journal*, *74*(5), 1800. https://doi.org/10.2136/sssaj2008.0190

Gupta, R. K., Mostaghimi, S., McClellan, P. W., Alley, M. M., & Brann, D. E. (1997). Spatial variability and sampling strategies for NO_3–N, P, and K determinations for site-specific farming. *Transactions of the ASAE*, *40*(2), 337–343. https://doi.org/10.13031/2013.21278

Gustavsson, B., Luthbom, K., & Lagerkvist, A. (2006). Comparison of analytical error and sampling error for contaminated error. *Journal of Hazardous Materials*, *B138*, 252–260. https://doi.org/10.1016/j.jhazmat.2006.01.082

Hatfield, J. L., & Walthall, C. L. (2015). Meeting global food needs: Realizing the potential via genetics × environment × management interactions. *Agronomy Journal*, *107*(4), 1215–1226.

Hausherr Luder, R., Qin, R., Richner, W., Stamp, P., & Noulas, C. (2018). Spatial variability of selected soil properties and its impact on the grain yield of oats (*Avena sativa* L.) in small fields. *Journal of Plant Nutrition*, *41*(19), 2446–2469. https://doi.org/10.1080/01904167.2018.1527935

Hemingway, R. G. (1955). Soil-sampling errors and advisory analyses. *Journal of Agricultural Science*, *46*(1), 1–8. https://doi.org/10.1017/S0021859600039563

Hengl, T., Rossiter, D. G., & Stein, A. (2003). Soil sampling strategies for spatial prediction by correlation with auxiliary maps. *Australian Journal of Soil Research*, *41*(8), 1403–1422. https://doi.org/10.1071/SR03005

Hurisso, T. T., Culman, S. W., & Zhao, K. (2018). Repeatability and spatiotemporal variability of emerging soil health indicators relative to routine soil nutrient tests. *Soil Science Society of America Journal*, *82*, 939–948. https://doi.org/10.2136/sssaj2018.03.0098

Heuvelink, G. B., & Webster, R. (2001). Modelling soil variation: Past, present, and future. *Geoderma*, *100*(3–4), 269–301. https://doi.org/10.1016/S0016-7061(01)00025-8

International Plant Nutrition Institute. (2015). Soil test levels in North America: Summary update. Retrieved from http://nap.ipni.net/article/NAP-3018 (accessed 14 Apr. 2019).

Kerry, R., & Oliver, M. A. (2003). Variograms of ancillary data to aid sampling for soil surveys. *Precision Agriculture*, *4*, 261–278. https://doi.org/10.1023/A:1024952406744

Kerry, R., & Oliver, M. A. (2008). Determining nugget:sIll ratios of standardized variograms from aerial photographs to krige sparse soil data. *Precision Agriculture*, *9*(1–2), 33–56. https://doi.org/10.1007/s11119-008-9058-0

Kerry, R., Oliver, M. A., & Frogbrook, Z. L. (2010). Sampling in precision agriculture. In: M. A. Oliver (Ed.), *Geostatistical applications for precision agriculture* (pp. 35–63). Dordrecht, The Netherlands: Springer. https://doi.org/10.1007/978-90-481-9133-8_2

Kitchen, N. R., Westfall, D. G., & Havlin, J. L. (1990). Soil sampling under no-till banded phosphorus. *Soil Science Society of America Journal*, *54*(6), 1661. https://doi.org/10.2136/sssaj1990.03615995005400060026x

Kravchenko, A. N. (2003). Influence of spatial structure on accuracy of interpolation methods. *Soil Science Society of America Journal*, *67*(5), 1564. https://doi.org/10.2136/sssaj2003.1564

Ladha, J. K., Tirol-Padre, A., Reddy, C. K., Cassman, K. G., Verma, S., Powlson, D. S., … Pathak, H. (2016). Global nitrogen budgets in cereals: A 50-year assessment for maize, rice, and wheat production systems. *Scientific Reports*, *6*, 1–9. https://doi.org/10.1038/srep19355

Lamb, J. A., & Rehm, G. W. (2002). Short-term stability of soil test phosphorus in agricultural fields. *Canadian Journal of Soil Science*, *82*(2), 239–247. https://doi.org/10.4141/S01-031

Lark, R. M. (2012). Towards soil geostatistics. *Spatial Statistics*, *1*, 92–99. https://doi.org/10.1016/j.spasta.2012.02.001

Lark, R. M., Cullis, B. R., & Welham, S. J. (2006). On optimal prediction of soil properties in the presence of trend: The empirical best linear unbiased predictor (E-BLUP) with REML. *European Journal of Soil Science*, *57*, 787–799. https://doi.org/10.1111/j.1365-2389.2005.00768.x

Lauzon, J. D., O'Halloran, I. P., Fallow, D. J., Bertoldi, von A. P., & Aspinall, D. (2005). Spatial variability of soil test phosphorus, potassium, and pH of Ontario soils. *Agronomy Journal*, *97*(2), 524. https://doi.org/10.2134/agronj2005.0524

Lemunyon, J. L., & Gilbert, R. G. (1993). The concept and need for a phosphorus assessment tool. *Journal of Production Agriculture*, *6*(4), 483–486. https://doi.org/10.2134/jpa1993.0483

Leo, M. W. (1963). Heterogeneity of soil of agricultural land in relation to soil sampling. *Journal of Agricultural and Food Chemistry*, *11*(5), 432–434. https://doi.org/10.1021/jf60129a023

Lockman, R. B., & Molloy, M. G. (1984). Seasonal variations in soil test results. *Communications in Soil Science and Plant Analysis*, *15*(7), 741–757. https://doi.org/10.1080/00103628409367514

Lohr, S. L. (2010). *Sampling: Design and analysis*. Boston, MA: Brooks/Cole.

Lopez-Granados, F., Jurado-Exposito, M., Atenciano, S., Garcia-Ferrer, A., de la Orden, M. S., & Garcia-Torres, L. (2002). Spatial variability of agricultural soil parameters in southern Spain. *Plant and Soil*, *246*, 97–105. https://doi.org/10.1023/A:1021568415380

Madow, W. G., & Madow, L. H. (1944). On the theory of systematic sampling, I. *Annals of Mathematical Statistics*, *15*(1), 1–24. https://doi.org/10.1214/aoms/1177731312

Mallarino, A. P. (1996). Spatial variability patterns of phosphorus and potassium in no-tilled soils for two sampling scales. *Soil Science Society of America Journal*, *60*(5), 1473. https://doi.org/10.2136/sssaj1996.03615995006000050027x

Mallarino, A. P., & Borges, R. (2006). Phosphorus and potassium distribution in the soil following long-term deep-band fertilization in different tillage systems. *Soil Science Society of America Journal*, *70*(2), 702–707. https://doi.org/10.2136/sssaj2005.0129

Mallarino, A. P., & Wittry, D. J. (2004). Efficacy of grid and zone soil sampling approaches for site-specific assessment of phosphorus, potassium, ph, and organic matter. *Precision Agriculture*, *5*(2), 131–144. https://doi.org/10.1023/B:PRAG.0000022358.24102.1b

Maresma, A., & Ketterings, Q. M. (2017). In-field variability of the Illinois Soil Nitrogen Test and loss-on-ignition results for nitrogen management. *Soil Science Society of America Journal*, *81*(5), 1211–1221. https://doi.org/10.2136/sssaj2016.10.0334

Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, *58*(8), 1246–1266. https://doi.org/10.2113/gsecongeo.58.8.1246

McBratney, A. B., Pringle, M. J., & McBratney, A. (1999). Estimating average and proportional variograms of soil properties and their potential use in precision agriculture. *Precision Agriculture*, *1*(2), 125–152.

McBratney, A. B., & Webster, R. (1981). The design of optimal sampling schemes for local estimation and mapping of regionalized variables. I: Theory and method. *Computers & Geosciences*, *7*(4), 331–334. https://doi.org/10.1016/0098-3004(81)90077-7

McBratney, A. B., & Webster, R. (1983). Optimal interpolation and isarithmic mapping of soil properties: V. Co-regionalization and multiple sampling strategy. *Journal of Soil Science*, *34*(1), 137–162. https://doi.org/10.1111/j.1365-2389.1983.tb00820.x

McCormick, S., Jordan, C., & Bailey, J. S. (2009). Within and between-field spatial variation in soil phosphorus in permanent grassland. *Precision Agriculture*, *10*, 262–276. https://doi.org/10.1007/s11119-008-9099-4

Meisinger, J. J. (1984). Evaluating plant-available nitrogen in soil–crop systems. In R. D. Hauck (Ed.), *Nitrogen in crop production* (pp. 391–416). Madison, WI: ASA.

Miao, Y., Mulla, D. J., & Robert, P. C. (2006). Spatial variability of soil properties, corn quality and yield in two Illinois, USA fields: Implications for precision corn management. *Precision Agriculture*, *7*, 5–20. https://doi.org/10.1007/s11119-005-6786-2

Minasny, B., & McBratney, A. B. (2006). A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers & Geosciences*, *32*(9), 1378–1388. https://doi.org/10.1016/j.cageo.2005.12.009

Mohanty, B. P., & Kanwar, R. S. (1994). Spatial variability of residual nitrate-nitrogen under two tillage systems in central Iowa: A composite three-dimensional resistant and exploratory approach. *Water Resources Research*, *30*(2), 237–251. https://doi.org/10.1029/93WR02922

Morris, T. F., Murrell, T. S., Beegle, D. B., Camberato, J. J., Ferguson, R. B., Grove, J., … Yang, H. (2018). Strengths and limitations of nitrogen rate recommendations for corn and opportunities for improvement. *Agronomy Journal*, *110*, 1–37. https://doi.org/10.2134/agronj2017.02.0112

Mueller, T. G., Pierce, F. J., Schabenberger, O., & Warncke, D. D. (2001). Map quality for site-specific fertility management. *Soil*

Science Society of America Journal, *65*(5), 1547–1558. https://doi.org/10.2136/sssaj2001.6551547x

Mueller, T. G., Pusuluri, N. B., Mathias, K. K., Cornelius, P., & Barnhisel, R. I. (2004). Site-specific soil fertility management: A model for map quality. *Soil Science Society of America Journal*, *68*(6), 2031–2041. https://doi.org/10.2136/sssaj2004.2031

Mzuku, M., Khosla, R., Reich, R., Inman, D., Smith, F., & MacDonald, L. (2005). Spatial variability of measured soil properties across site-specific management zones. *Soil Science Society of America Journal*, *69*(5), 1572. https://doi.org/10.2136/sssaj2005.0062

Olea, R. A. (1984). Sampling design optimization for spatial functions. *Journal of the International Association for Mathematical Geology*, *16*(4), 369–392. https://doi.org/10.1007/BF01029887

Oliver, M. A., & Webster, R. (2014). A tutorial guide to geostatistics: Computing and modeling variograms and kriging. *Catena*, *113*, 56–69. https://doi.org/10.1016/j.catena.2013.09.006

Peck, T. R., & Melsted, S. W. (1967). Field sampling for soil testing. In G. W. Hardy (Ed.), *Soil testing and plant analysis. Part 1. Soil testing. Special Publication 2* (pp. 25–35). Madison, WI: SSSA.

Penney, D. C., Nolan, S. C., McKenzie, R. C., Goddard, T. W., & Kryzanowski, L. (1996). Yield and nutrient mapping for site specific fertilizer management. *Communications in Soil Science and Plant Analysis*, *27*(5–8), 1265–1279. https://doi.org/10.1080/00103629609369631

Peralta, N. R., & Costa, J. L. (2013). Delineation of management zones with soil apparent electrical conductivity to improve nutrient management. *Computers and Electronics in Agriculture*, *99*, 218–226. https://doi.org/10.1016/j.compag.2013.09.014

Peterson, H., & Bruulsema, T. (2019). Sustainable phosphorus management: Defining 4R practices. *Better Crops Plant Food*, *103*, 54–56. https://doi.org/10.24047/BC103154

Reed, J. F., & Rigney, J. A. (1947). Soil sampling from fields of uniform and nonuniform appearance and soil types. *Journal American Society of Agronomy*, *39*, 26–40. https://doi.org/10.2134/agronj1947.00021962003900010003x

Rehm, G. W., Scobbie, A. J., Randall, G. W., & Vetsch, J. A. (1995). Impact of fertilizer placement and tillage system on phosphorus distribution in soil. *Soil Science Society of America Journal*, *59*(6), 1661. https://doi.org/10.2136/sssaj1995.03615995005900060022x

Ruffo, M. L., Bollero, G. A., Hoeft, R. G., & Bullock, D. G. (2005). Spatial variability of the Illinois Soil Nitrogen Test. *Agronomy Journal*, *97*(6), 1485–1492. https://doi.org/10.2134/agronj2004.0323

Reuss, J. O., Soltanpour, P. N., & Ludwick, A. E. (1977). Sampling distribution of nitrates in irrigated fields. *Agronomy Journal*, *69*(4), 588. https://doi.org/10.2134/agronj1977.00021962006900040016x

Robertson, G., Crum, J. R., & Ellis, B. G. (1993). The spatial variabilty of soil resources following long-term disturbance. *Oecologia*, *96*, 451–456. https://doi.org/10.1007/BF00320501

Robinson, G. W., & Lloyd, W. E. (1915). On the probable error of sampling in soil surveys. *Journal of Agricultural Science*, *7*(2), 144. https://doi.org/10.1017/S0021859600002598

Sawchik, J., & Mallarino, A. P. (2007). Evaluation of zone soil sampling approaches for phosphorus and potassium based on corn and soybean response to fertilization. *Agronomy Journal*, *99*(6), 1564. https://doi.org/10.2134/agronj2006.0151

Schabenberger, O., & Gotway, C. A. (2005). *Statistical methods for spatial data analysis*. Boca Raton, FL: Chapman & Hall/CRC Press.

Schepers, A. R., Shanahan, J. F., Liebig, M. A., Schepers, J. S., Johnson, S. H., & Luchiari, A. (2004). Appropriateness of management zones for characterizing spatial variability of soil properties and irrigated corn yields across years. *Agronomy Journal*, *96*(1), 195. https://doi.org/10.2134/agronj2004.0195

Sharpley, A. N., Weld, J. L., Beegle, D. B., Kleinman, P. J. A., Gburek, W. J., Moore, P., & Mullins, G. (2003). Development of phosphorus indices for nutrient management planning strategies in the United States. *Journal of Soil and Water Conservation*, *58*(3), 137–152.

Song, X., Wang, J., Huang, W., Liu, L., Yan, G., & Pu, R. (2009). The delineation of agricultural management zones with high resolution remotely sensed data. *Precision Agriculture*, *10*(6), 471–487. https://doi.org/10.1007/s11119-009-9108-2

Stenger, R., Priesack, E., & Beese, F. (2002). Spatial variation of nitrate-N and related soil properties at the plot scale. *Geoderma*, *105*(3–4), 259–275. https://doi.org/10.1016/S0016-7061(01)00107-0

Tabor, J. A., Warrick, A. W., Myers, D. E., & Pennington, D. A. (1985). Spatial variability of nitrate in irrigated cotton: II. Soil nitrate and correlated variables. *Soil Science Society of America Journal*, *49*(2), 390. https://doi.org/10.2136/sssaj1985.03615995004900020024x

Taylor, J. A., McBratney, A. B., & Whelan, B. M. (2007). Establishing management classes for broadacre agricultural production. *Agronomy Journal*, *99*(5), 1366. https://doi.org/10.2134/agronj2007.0070

Tisdale, S. L., & Nelson, W. L. 1956. *Soil fertility and fertilizers*. New York: Macmillan.

van Groenigen, J. W. (2000). The influence of variogram parameters on optimal sampling schemes for mapping by kriging. *Geoderma*, *97*(3–4), 223–236. https://doi.org/10.1016/S0016-7061(00)00040-9

van Groenigen, J. W., Siderius, W., & Stein, A. (1999). Constrained optimisation of soil sampling for minimisation of the kriging variance. *Geoderma*, *87*(3–4), 239–259. https://doi.org/10.1016/S0016-7061(98)00056-1

van Meirvenne, M., & Hofman, G. (1989). Spatial variability of soil nitrate nitrogen after potatoes and its change during winter. *Plant and Soil*, *120*(1), 103–110. https://doi.org/10.1007/BF02370296

Vašát, R., Heuvelink, G. B. M., & Borůvka, L. (2010). Sampling design optimization for multivariate soil mapping. *Geoderma*, *155*(3–4), 147–153. https://doi.org/10.1016/j.geoderma.2009.07.005

Vasu, D., Singh, S. K., Sahu, N., Tiwary, P., Chandran, P., Duraisami, V. P., … Kalaiselvi, B. (2017). Assessment of spatial variability of soil properties using geospatial techniques for farm level nutrient management. *Soil and Tillage Research*, *169*, 25–34. https://doi.org/10.1016/j.still.2017.01.006

Wang, J.-F., Stein, A., Gao, B.-B., & Ge, Y. (2012). A review of spatial sampling. *Spatial Statistics*, *2*, 1–14. https://doi.org/10.1016/j.spasta.2012.08.001

Warrick, A. W., & Myers, D. E. (1987). Optimization of sampling locations for variogram calculations. *Water Resources Research*, *23*(3), 496–500. https://doi.org/10.1029/WR023i003p00496

Webster, R., & McBratney, A. B. (1987). Mapping soil fertility at Broom's Barn by simple kriging. *Journal of the Science of Food and Agriculture*, *38*(2), 97–115. https://doi.org/10.1002/jsfa.2740380203

Webster, R., & Oliver, M. A. (1992). Sample adequately to estimate variograms of soil properties. *European Journal of Soil Science*, *43*(1), 177–192. https://doi.org/10.1111/j.1365-2389.1992.tb00128.x

Webster, R., & Oliver, M. A. (2007). *Geostatistics for environmental scientists* (2nd ed.). Chichester: John Wiley & Sons. https://doi.org/10.1002/9780470517277

Webster, R., & Beckett, P. H. T. (1968). Quality and usefulness of soil maps. *Nature*, *219*, 680–682. https://doi.org/10.1038/219680a0

Wollenhaupt, N. C., Wolkowski, R. P., & Clayton, M. K. (1994). Mapping soil test phosphorus and potassium for variable-rate fertilizer application. *Journal of Production Agriculture*, *7*(4), 441–448. https://doi.org/10.2134/jpa1994.0441

Yfantis, E., Flatman, G. T., & Behar, J. V. (1987). Efficiency of kriging estimation for square, triangular, and hexagonal grids. *Mathematical Geology*, *19*(3), 183–205. https://doi.org/10.1007/BF00897746

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Lawrence PG, Roper W, Morris TF, Guillard K. Guiding soil sampling strategies using classical and spatial statistics: A review. *Agronomy Journal*. 2020;112:493–510. https://doi.org/10.1002/agj2.20048