Do soil health tests match farmer experience? Assessing biological, physical, and chemical indicators in the upper Midwestern United States

Article *in* Soil Science Society of America Journal · February 2021

DOI: 10.1002/saj2.20233

Christine Dazil Sprunger
The Ohio State University

Received: 14 June 2020 | Accepted: 27 January 2021

DOI: 10.1002/saj2.20233

Soil Science Society of America Journal

NUTRIENT MANAGEMENT & SOIL & PLANT ANALYSIS

Do soil health tests match farmer experience? Assessing biological, physical, and chemical indicators in the Upper Midwest United States

Brendan O'Neill[1] | **Christine D. Sprunger**[2] | **G. Philip Robertson**[3,4]

[1] School for Environment and Sustainability, Univ. of Michigan, Ann Arbor, MI 48109, USA

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

[3] W.K. Kellogg Biological Station, Michigan State Univ., Hickory Corners, MI 49060, USA

[4] Dep. of Plant, Soil, and Microbial Sciences and Great Lakes Bioenergy Research Center, Michigan State Univ., East Lansing, MI 48824, USA

Correspondence
Brendan O'Neill, School for Environment and Sustainability, Univ. of Michigan, Ann Arbor, MI USA
Email: obrendan@umich.edu.

Assigned to Associate Editor David Herring Hardy.

Abstract

Soil health testing provides an integrated assessment of biological, physical, and chemical attributes to inform the sustainable management of farm fields. However, it is unclear how tests reflect farmers' own assessments of soil quality and agronomic performance, which may disproportionately influence farm management practices. We asked farmers in three regions of Michigan to identify three fields to compare their own assessments against soil health tests: a "best," a "worst," and a "non-row crop" reference field. Each field was tested for soil aggregate stability, available water capacity, soil organic matter (SOM), mineralizable carbon (MinC), permanganate oxidizable carbon (POXC), pH, P, and K. We evaluated soil health scores using paired *t* tests to compare results from contrasting fields with farmers' assessments of each field. Across all farms, the overall soil health test score for cropped fields was significantly higher on fields farmers rated as "Best." This result was driven solely by physical and biological (including C) parameters; inorganic chemical tests did not distinguish among field types. On reference fields in all regions, biological parameters were consistently higher, but inorganic chemical and physical measures were not. The performance of soil C measures was inconsistent: SOM and MinC consistently detected significant differences between "Best" and "Worst" cropped fields, but POXC did not. Our results suggest that common soil health assays for physical and biological attributes generally align well with farmers' assessments of their fields. That soil health tests match farmer experience reinforces the value of these tests as a meaningful guide for soil management decisions.

1 | INTRODUCTION

The environmental and social costs of intensive agricultural production in the United States have led to calls for more ecologically based approaches to management (Drinkwater & Snapp, 2007; Robertson et al., 2014; Schipanski et al., 2016). Ecological management practices are designed to

Abbreviations: AS, aggregate stability; AWC, available water capacity; CEC, cation exchange capacity; CND, cumulative normal distribution; MinC, mineralizable carbon; MSU, Michigan State University; NRC, non-row crop; PMN, potentially mineralizable N; POXC, permanganate oxidizable carbon; SOM, soil organic matter; SR, surface resistance; SSR, subsurface resistance maintain crop productivity while also delivering a range of ecosystem services both on-farm and to society at large (Power, 2010; Robertson et al., 2014). Soil health is broadly defined as the continued capacity of soil to function as a vital living system that sustains plants, animals, and humans (Doran & Parkin, 1994); thus, soil health has emerged as a framework for linking soil management practices to agronomic performance and ecosystem function (Culman et al., 2013; Karlen et al., 2008; Lal, 2016; Wade et al., 2020). In practice, the soil health paradigm has shaped a new soil testing regime that is more closely linked to principles of ecological management and that could potentially lead to outcomes such as improved crop growth, soil carbon (C) sequestration, and reduced nutrient leaching (Cherry et al., 2008; Karlen et al., 2006; Minasny et al., 2017).

Farmers often have detailed knowledge of their long-term managed fields (Gruver & Weil, 2007), and this knowledge logically informs management decisions related to nutrients, tillage, and residue. Soil testing is rightly seen as a tool used in conjunction with farmers' own knowledge to guide field management (Andrews et al., 2003). As with other soil tests, farmers require actionable management decisions to be based on acceptance of soil health test results. If specific test parameters accord with farmers' experience for a given field, this may advance farmers' acceptance of testing results and ultimately translate into ecological management practices. In turn, understanding how soil health test results align with farmers' assessments of a field's agronomic performance can inform recommendations that follow from testing results.

Traditional soil testing for row crops is primarily focused on soil inorganic chemistry and in particular pools of plant nutrients (i.e., N, P, K, and micronutrients) and soil pH. Soil health tests include these parameters but also include them with key measures of biological and physical properties that together drive ecosystem functions such as soil C accumulation and aggregation. Integrated measures may better correspond to characteristics that farmers use to describe their own fields. For example, farmers often describe physical features, such as how fields respond to precipitation, cultivation, or seed set. The advancement of integrated soil health indicators, such as available water capacity (AWC) and aggregate stability (AS) and surface hardness, is an acknowledgment of this disconnect in soil testing approaches (Fine et al., 2017; Moebius-Clune et al., 2016).

Even total soil organic matter (SOM), which is often measured in standard soil tests, may not align with farmers' field assessments. Soil organic matter, comprised mostly of older, stable fractions of C, may impart soil qualities such as greater cation exchange capacity (CEC) and water holding capacity. Although farmers may recognize and value the importance of maintaining SOM, standard SOM measures do not typically inform yearly management decisions and may not explain variation in field performance (O'Neill, 2017; Sprunger, 2015). Indeed, measures of dynamic C fractions often correspond better to fertility status than do measures of total SOM (Culman et al., 2013; Wander, 2004) and can be more sensitive to management and thus potentially provide farmers with a more integrated assessment of soil functioning. For example, physical metrics, like AS, are intricately linked to biological indicators because stable aggregates emerge from microbial activity and root production (Chantigny et al., 1997; Tiemann & Grandy, 2015). Although research has shown that integrated metrics of soil function distinguish between management practices (Idowu et al., 2009; Morrow et al., 2016), we lack understanding as to how these measures accord with farmers' knowledge of their fields.

Core Ideas

- Soil health test results generally aligned with farmers' field assessments of soil quality.
- Biological and physical indicators best detected differences across fields.
- Across regions and soil types, mineralizable C best aligned with farmers' assessments.

To date, much of the validation of soil health indicators has occurred on controlled experimental field trials (e.g., Culman et al., 2013; Hurisso et al., 2016; Morrow et al., 2016; Roper et al., 2017; Sprunger et al., 2019) and less so on farmers' fields (e.g., Williams et al., 2020), which limits our understanding of how soil health indicators can guide farmers as they make critical management decisions (Karlen et al., 2017). Many row-crop growers in the United States farm over 180 ha, often on multiple fields that are miles apart. Generally, this means that each field presents unique challenges based on field-by-field variation in soil quality and management history.

In contrast to experimental field trials, farmers' field management is often dynamic, with multiple management practices implemented season to season over several decades based on a range of practical considerations that are tailored to specific fields. Yet, to our knowledge, dynamic measures of soil C, such as mineralizable C (minC) and permanganate oxidizable C (POXC), have not been assessed in relation to how farmers rate field performance. Farmers are usually knowledgeable of their fields and commonly label them as "good" or "poor," indicative of factors such as agronomic performance and soil quality (Gruver & Weil, 2007). The alignment of farmers' knowledge of soil characteristics and function to the interpretation of soil health tests results is important for the implementation and adoption of soil health management practices.

Here we address this knowledge gap by asking how well—if at all—soil health indicators reflect farmers' knowledge and assessments of their fields. Additionally, testing soil health on working farms provides an opportunity to assess the sensitivity of soil health tests indicators to detect variability across fields and on a range of soil types, which can ultimately guide new research questions and inform outreach and recommendations for farmers.

We combined biophysical data with qualitative data from interviews with farmers to assess how chemical, physical, and biological metrics of soil health align with farmers' knowledge of their fields. Our specific research objectives were (a) to quantify variability in on-farm soil health scores across three agricultural regions in Michigan and (b) to evaluate the degree to which soil health parameters align with farmers' assessments of field performance. We hypothesize that physical and biological soil health indicators will better align with farmers' field assessments than will chemical assessments due to a better ability to differentiate among fields that lack measurable nutrient deficiencies.

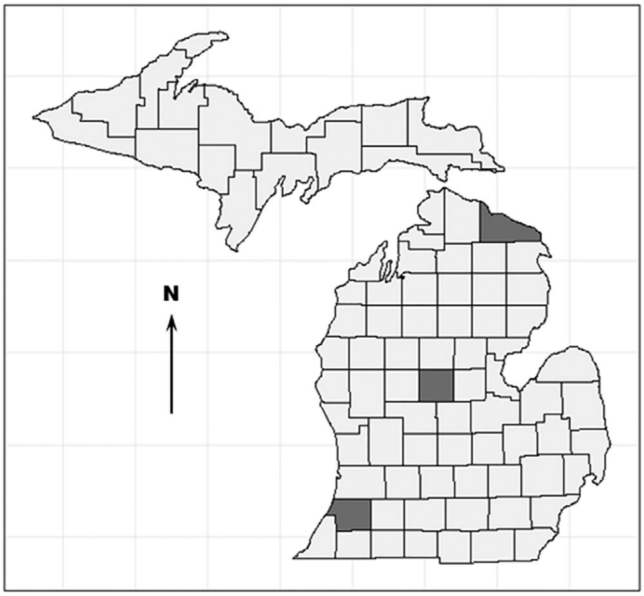

FIGURE 1 Map of Michigan with sampled regions (North, Central, and South) shaded in dark gray.

2 | MATERIALS AND METHODS

2.1 | Participant selection

Our study is grounded in a participatory research framework that included Michigan farmers, Michigan State University (MSU) Extension staff, and MSU researchers. We asked staff from MSU extension and conservation districts in each region to recommend farmers who might be willing to be interviewed in exchange for free soil testing. Eligibility was limited to farmers with conventionally managed row-crops in order to best represent regional agricultural land use. Median farm size was 172 ha, and none of the participants had previously undertaken soil health testing on their fields. The participating farms were operations that primarily grow grains and some cover crops.

We asked each participating farmer to identify three fields, including a "Best" field, a "Worst" field, and a reference field that was currently not in row crops (non-row crop [NRC]), such as a pasture, land under conservation, or a buffer strip field margin. The NRC field served as a within-farm, low-management-intensity reference field as compared to cropped fields. Fields under perennial vegetation typically score higher on soil health metrics (De et al., 2020) and thus can be useful for comparing metrics across cropped fields and for comparing differences across regions.

In total we evaluated 40 fields from three field types: Best, Worst, and NRC. These represent 13 farms from the north, central, and southwestern regions of Michigan, which have distinct climates and soil types (Figure 1; Table 1). This yielded a broad range of fields on which to test and compare soil health parameters while allowing for a participatory approach to engage directly with farmers. After asking farmers to identify these fields and sampling each, we met with individual farmers to discuss what properties defined their characterization of each field and to gather management histories. Research activities were compliant with the MSU Human Research Protection Program, and this study was classified as exempt (IRB #i046108).

2.2 | Field sampling

The cropped fields sampled ranged in area from 2 to 28 ha (median, 12 ha). We measured management-sensitive parameters of soil health to examine how well soil health tests and specific metrics characterized the soils on farm fields as compared to standard soil fertility tests offered by the MSU Plant, Soil and Nutrient Laboratory. This objective informed our sampling approach. Soil samples were removed according to the Cornell Soil Health Assessment guidelines (Moebius-Clune et al., 2016), which includes assessment of field variability or anomalies, soil conditions, and crop management. In each field, five representative locations were selected; at each location, bulk soil samples (~4 cm by 9 cm to 15-cm depth) were excavated from the sides of each of two shovel-dug pits. At each location, two penetrometer readings (Imants) at 15- and 45-cm depths were taken to assess surface and subsurface compaction, respectively. For a given field, each of the 10 bulk soil samples was composited, thoroughly mixed, subsampled (~2 kg), placed in a plastic bag, and stored on ice until

TABLE 1 Sand content, textural class designation, field area, soil series, and soil classification for each field type (Best, Worst, non-row crop [NRC]) in all regions with associated latitude and longitude

Region	Farm no.	Field type	Sand%	Textural class	Areaha	Soil series[a]	Soil classification
North (45˚42′ N, 83˚81′ W)	1	Best	88	loamy sand	11	Emmet	coarse-loamy, mixed, active, frigid Inceptic Hapludalfs
		Worst	82	loamy sand	6	Emmet	coarse-loamy, mixed, active, frigid Inceptic Hapludalfs
		NRC	85	loamy sand	6	Cheboygan	coarse-loamy, mixed, active, frigid Alfic Haplorthods
	2	Best	81	loamy sand	12	Omena	coarse-loamy, mixed, active, frigid Haplic Glossudalfs
		Worst	81	loamy sand	10	Omena	coarse-loamy, mixed, active, frigid Haplic Glossudalfs
		NRC	83	loamy sand	6	Omena	coarse-loamy, mixed, active, frigid Haplic Glossudalfs
	3	Best	74	sandy loam	6	Omena	coarse-loamy, mixed, active, frigid Haplic Glossudalfs
		Worst	64	sandy loam	16	Omena	coarse-loamy, mixed, active, frigid Haplic Glossudalfs
		NRC	82	loamy sand	5	Omena	coarse-loamy, mixed, active, frigid Haplic Glossudalfs
	4	Best	70	sandy loam	3	Ossineke	fine-loamy, mixed, semiactive, frigid Oxyaquic Glossudalfs
		Worst	71	sandy loam	2	Ossineke	fine-loamy, mixed, semiactive, frigid Oxyaquic Glossudalfs
		NRC	76	sandy loam	5	Slade	fine-loamy, mixed, active, frigid Aquic Glossudalfs
Central (43˚60′ N, 84˚76′ W)	5	Best	42	loam	28	Ithaca	fine, mixed, semiactive, mesic Aquic Glossudalfs
		Worst	38	loam	26	Ithaca	fine, mixed, semiactive, mesic Aquic Glossudalfs
		NRC	60	sandy loam	1	Ithaca	fine, mixed, semiactive, mesic Aquic Glossudalfs
	6	Best	63	sandy loam	21	Conover	fine-loamy, mixed, active, mesic Aquic Hapludalfs
		Worst	71	sandy loam	17	Parkhill	fine-loamy, mixed, semiactive, nonacid, mesic Mollic Epiaquepts
		NRC	71	sandy loam	2	Parkhill	fine-loamy, mixed, semiactive, nonacid, mesic Mollic Epiaquepts

(Continues)

TABLE 1 (Continued)

Region	Farm no.	Field type	Sand%	Textural class	Areaha	Soil series[a]	Soil classification
	7	Best	66	sandy loam	5	Conover	fine-loamy, mixed, active, mesic Aquic Hapludalfs
		Worst	80	loamy sand	12	Conover	fine-loamy, mixed, active, mesic Aquic Hapludalfs
		NRC	78	loamy sand	7	Parkhill	fine-loamy, mixed, semiactive, nonacid, mesic Mollic Epiaquepts
	8	Best	66	sandy loam	14	Ziegenfuss	fine, mixed, semiactive, nonacid, mesic Mollic Epiaquepts
		Worst	64	sandy loam	16	Ithaca	fine, mixed, semiactive, mesic Aquic Glossudalfs
		NRC	91	sand	0.5	Ithaca	fine, mixed, semiactive, mesic Aquic Glossudalfs
	9	Best	86	loamy sand	12	Onekama	fine, mixed, active, mesic Haplic Glossudalfs
		Worst	71	sandy loam	16	Ithaca	fine, mixed, semiactive, mesic Aquic Glossudalfs
		NRC	45	loam	2	Ithaca	fine, mixed, semiactive, mesic Aquic Glossudalfs
South (42°21′ N, 85°89′ W)	10	Best	82	loamy sand	2	Coloma	mixed, mesic Lamellic Udipsamments
		Worst	83	loamy sand	6	Oshtemo	coarse-loamy, mixed, active, mesic Typic Hapludalf
		NRC	83	loamy sand	4	Coloma	mixed, mesic Lamellic Udipsamments
	11	Best	65	sandy loam	12	Oshtemo	coarse-loamy, mixed, active, mesic Typic Hapludalf
		Worst	63	sandy loam	9	Oshtemo	coarse-loamy, mixed, active, mesic Typic Hapludalf
		NRC	65	sandy loam	8	Oshtemo	coarse-loamy, mixed, active, mesic Typic Hapludalf
	12	Best	85	loamy sand	13	Oshtemo	coarse-loamy, mixed, active, mesic Typic Hapludalf
		Worst	92	sand	3	Oshtemo	coarse-loamy, mixed, active, mesic Typic Hapludalf
		NRC	86	loamy sand	2	Riddles	fine-loamy, mixed, active, mesic Typic Hapludalfs
	13	Best	85	loamy sand	15	Oshtemo	coarse-loamy, mixed, active, mesic Typic Hapludalf
		Worst	81	loamy sand	2	Oshtemo	coarse-loamy, mixed, active, mesic Typic Hapludalf
		NRC	86	loamy sand	2	Oshtemo	coarse-loamy, mixed, active, mesic Typic Hapludalf

[a]Represents the dominant soil series, by area, in each field.

further processing. We noted each field location, and sampling locations were identified by use of NRCS Web Soil Survey (https://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm). All samples were collected on the same day for each region (between 26 May and 12 June 2014) and maintained at 4 ˚C until processed.

2.3 | Soil testing

Samples were sieved to <8 mm to remove stones, and a subsample of sieved soil was submitted to the MSU Plant, Soil and Nutrient Laboratory for analysis. Soil pH was determined in a 1:1 soil and water solution, total SOM by loss on ignition, and P with Bray 1 extractant. Soil K^+ was extracted with 1 N ammonium acetate, and all cation concentrations were determined via flame emission spectroscopy. Cation exchange capacity was calculated through summation of cations plus the contribution of pH (buffer index + meq of K^+). We grouped these parameters into chemical measures. We analyzed the remaining soil for soil texture, AS, and AWC (physical parameters) and for POXC, mineralizable carbon (MinC), and potentially mineralizable N (PMN) (biological parameters).

Soil texture was determined using a rapid method (Kettler et al., 2001) on soil dried overnight at 60 ˚C. A 14-g portion of soil was weighed into a 50-ml Falcon tube containing 42 ml of 3% hexametaphosphate solution; tubes were placed on their side on a shaker at 120 rpm for 2 h. Contents of each tube were poured through a 0.053-mm sieve and thoroughly washed with 600 ml deionized water into a catch basin. The sand fraction on the sieve was washed into a previously tared drying tin. Particles in the catch basin were thoroughly resuspended and allowed to settle for 4–6 h, after which the clay particles in suspension were decanted. The settled silt particles were washed into another tared drying tin. All tins were dried overnight at 105 ˚C, and contents were weighed. Texture was calculated as: sand % = (oven dry sand mass/original sample mass) × 100%; silt % = (oven dry silt mass/original sample mass) × 100%; and clay % = 100 − (sand % + silt %).

Wet AS was determined from soil dried to constant weight at 40 ˚C for 1–2 d in the oven followed by isolation of aggregate size fractions from 0.25 to 2 mm (Moebius et al., 2007). Ten grams of soil aggregates were spread evenly on a 0.25-mm mesh, 125-mm-diameter sieve. The sieve was placed on a funnel containing previously weighed filter paper, all atop a ring stand. Sieves were exposed to a rain simulator (rate previously calibrated) for 5 min, after which material retained on the sieve was thoroughly washed through the sieve. Remaining particles (e.g., small stones) were washed off the sieve surface into a drying tin. The tin and filter paper with slaked soil were oven-dried for 1 d at 105 ˚C, and AS was calculated as the percentage of soil retained on the sieve (difference from what was not slaked onto the filter) and adjusting for the mass of unsieved particles.

To determine AWC, another portion of soil dried to 60 ˚C was sieved to <2 mm. Two 15-g portions were spread evenly inside brass rings situated on ceramic plates with known porosity under water saturation. Plates were placed into high-pressure chambers: 10 kPa (field capacity) and 1,500 kPa (permanent wilting point). After equilibration, soils were weighed, dried at 105 ˚C, and reweighed. Available water capacity was calculated as soil water loss between samples at 10 and 1,500 kPa and reported as g water per g soil.

To determine labile C as POXC, duplicate 2.5-g samples of air-dried soil were mixed with buffered 0.02 M $KMnO_4$ solution in 50-ml conical tubes, shaken at 120 rpm for 2 min, and allowed to settle for 8 min (Weil et al., 2003). From this reaction, 0.5 ml of supernatant was diluted with 49.5 ml of deionized water. The degree of oxidation was measured colorimetrically at 550 nm on a Fisher Scientific Thermo Multiskan microplate reader and standardized to a series of known $KMnO_4$ standards.

To determine MinC, 10 g of air-dried soil from each field sample was placed in loosely capped Mason jars and brought to 50% water-filled pore space (Franzluebbers et al., 2000; Robertson et al., 1999). The jars were then incubated for 24 h at 25 ˚C, after which they were capped tightly. A CO_2 reading was taken immediately by injecting 0.5 ml of headspace gas into an infrared gas absorption analyzer (LI-CO7R LI-820, LI-COR Biosciences). Three subsequent readings were taken over the following 90 min, and a flux was calculated by regressing the change in CO_2 against the incubation period. Two analytical replicates were used for each field sample. Final fluxes were calculated by averaging analytical replicates.

Potentially mineralizable N was determined from field-moist soil sieved to <2 mm (Drinkwater et al., 1996). For each of the five field samples, $NH_4{}^+$ was extracted from duplicate 8-g soil aliquots using 1 M KCl while shaking (rotary shaker) at 120 rpm for 1 h. Two additional 8-g replicates of soil were placed in conical tubes, 10 ml of deionized water was added, and dinitrogen gas was used to replace tube headspace air and bubbled into the slurry for 1 min prior to sealing with butyl rubber stoppers. Sealed tubes were incubated at 25 ˚C. After 7 d, the stoppers were removed; buffer was added to bring the slurry to 1 M KCl; and samples were shaken, filtered, and stored on ice. Concentrations of $NH_4{}^+$ were determined colorimetrically at 630 nm (Sinsabaugh et al., 2000). Potentially mineralizable C was determined from the concentration of $NH_4{}^+$ from incubated soil minus $NH_4{}^+$ from an initial soil extraction of the same soil.

2.4 | Farmer interviews

After compiling all soil testing data, we conducted interviews with participant growers. In the first phase of the interview, we discussed the management history of each field type, including crop rotation, tillage, farmer-specific management decisions, and criteria used to categorize fields (i.e., Best and Worst). In the second phase, we discussed specific test results for all fields and held open discussions aimed at integrating soil test results with farmers' knowledge and practical experience for each field type. All of the interviews were recorded, and notes were transcribed within 24 h of each interview. Recordings were transcribed, analyzed for common themes, and coded based on specific soil test parameters, on different approaches to soil management for each field type, and on the influence of soil testing on management practices (Saldaña, 2015).

2.5 | Scoring and statistical analysis

We used the mean and standard deviation for each parameter measured to calculate normal distributions in R (R Core Team, 2020). These were calculated from the Best, Worst, and NRC fields and from an additional field for which farmers desired soil health results, for a total of 52 fields across 13 farms. For parameters that indicate greater health with a higher value (AS, AWC, SOM, POXC, MinC, PMN, CEC), we used a cumulative normal distribution (CND):

$$p = f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx \qquad (1)$$

where p is the probability (between 0 and 1) that the indicator value x falls within the distribution with mean μ and standard deviation σ. The probability was multiplied by 100 to scale indicator scores from 0 to 100 for each soil health metric (Fine et al., 2017). For indicators where greater values reflect decreased soil health (surface and subsurface hardness), we calculated 1 − CND for the score. For parameters with optimum values, we followed guidelines from MSU's Soil and Plant Nutrient Laboratory (https://www.canr.msu.edu/spnl/). Specifically, soil pH values between 6.0 and 6.8 were rated as optimum and received a score of 100, and values ≤5.5 or ≥7.75 received a score of 0 with linear interpolation of intermediate values between optimums and extremes. For soil P, values between 20 and 30 mg kg^{-1} received scores of 100, with scores falling at concentrations above or below this optimum range determined by linear interpolation to MSU recommendations. For soil K, a CND was calculated for increasing soil health scores with higher concentrations of K^+, with ≥100 mg kg^{-1} K^+ set to a score of 100. Virtually all concentrations of soil Ca^{2+} and Mg^{2+} met optimum values based on state recommendations; thus, these cations were omitted from soil health scoring and analysis. We used ANOVA in R to compare differences in percent sand in soils from different regions and to test parameter score differences across field types within each region. Where an effect was significant, Tukey's HSD was used to compare treatments. Paired t tests in R were used to compare results between field types (i.e., between Best and Worst and between each cropped field and the NRC field) across all farms for both raw indicator values and for soil health scores determined from Equation 1.

3 | RESULTS

3.1 | Soil texture and farmers' field descriptions

The regions sampled (Figure 1) represent distinct zones of Michigan in terms of climate and soil series (Table 1). Soils in the North region experience frigid temperature regimes and are characterized by greater alkalinity, whereas soils in the Central region are poorly drained and semiactive in cation exchange. Soils in the South region are well-drained outwash plains or beach ridges with slight acidity. On paired cropped field types, soil classifications differed on 4 out of 13 farms sampled (Farms 6, 8, 9, 10), although at least two of three field types shared a common soil classification on all farms (Table 1). The sand content differed significantly by region ($P < .001$). Fields in the Central region had a mean (± SD) sand content of 64 ± 4%; farms in the North and South regions had nearly the same mean sand contents of 79 ± 2% and 80 ± 2%, respectively.

In comparing field types, the Best, Worst, and NRC fields had mean (± SD) sand contents of 72 ± 13%, 72 ± 14%, and 76 ± 13%, respectively, with paired t test comparisons (not shown) showing no difference in sand contents between field types across sampled farms.

Farmers were asked why they designated a field as either a Best or Worst field. For the Best field, 10 of 13 farmers stated this field had high crop yields, and five farmers commented on both how the soil "worked" and their efforts to take care of this field; and four farmers commented on their field's "reliability" and field drainage (Figure 2; Supplemental Table S2). In designation of the Worst field, 9 out of 13 farmers stated both that yields were lower and that the soil "worked poorly." Other reasons included poor field drainage, low reliability, soil compaction (e.g., describing a field that is "hard to work" or stating a field requires occasional deep tillage), known poor management history (e.g., stating a field had excessive tillage or many seasons in a single crop), poor soil "chemistry," and disease problems (Figure 2; Supplemental Table S2). Nearly all Best and Worst fields on each farm experienced the same crop rotation, although tillage practice (no-till or chisel plow)

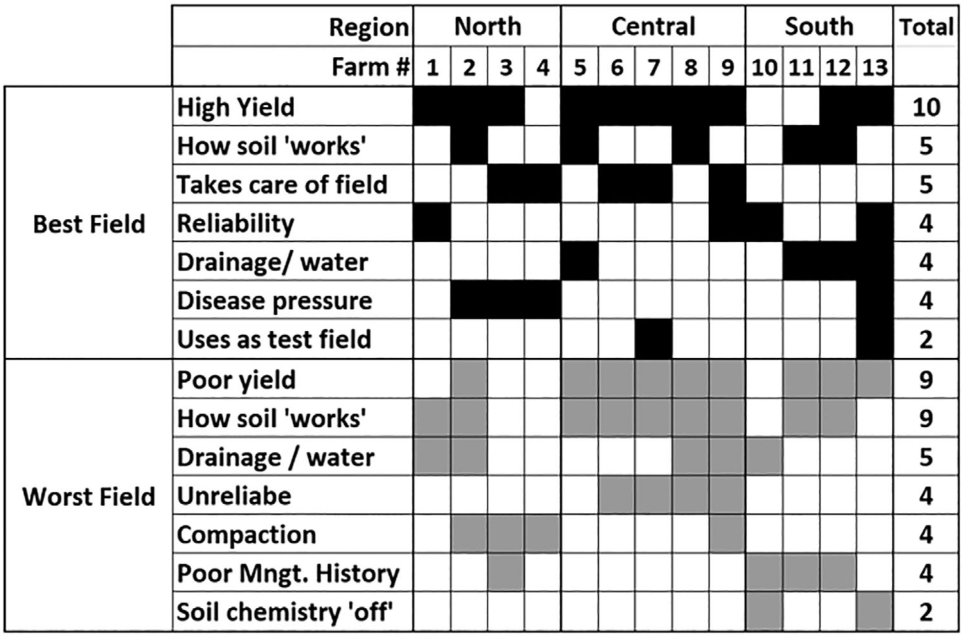

FIGURE 2 Rationale stated by farmers characterizing Best fields (top, in black) and Worst fields (bottom, in gray) in each region and the number of farmers (right column) who assessed each field type based on each select criterion.

and manure input tended to differ more between field types (Figure 3).

3.2 | Soil health test results by field type

The overall soil health score for Best fields was significantly higher than for Worst fields, with a mean difference of 6.9 units (Table 2). Overall, physical and biological soil health parameters had significantly higher scores on Best fields as compared to Worst fields (Table 2). For chemical soil health, Best fields rated higher on 7 of 13 farms, but there were no significant differences between the means of aggregated chemical parameters by each field comparison (Supplemental Figure S1; Table 2).

The mean soil health scores by parameter category and the overall scores were driven by clear patterns in individual soil health metrics. Best fields had a significantly higher rating for AS and AWC compared with Worst fields (Table 2). Although measures of surface resistance (SR) and subsurface resistance (SSR) were generally more favorable on Best fields (Supplemental Figure S1), they did not differ significantly when comparing Best and Worst fields (Table 2). The significantly higher mean biological soil health score on Best compared with Worst fields reflected significantly greater values for SOM and MinC on Best fields (Table 2). Both PMN and POXC were not significantly different in cropped field comparisons.

For the chemical category of soil health, no significant differences were observed between Best and Worst fields (Table 2), and generally these parameters had greater variability among field types and by region. For example, soil pH did not strongly differentiate between Best and Worst field types in the South region relative to other parameters, whereas in the Central region the large magnitude in differences between cropped fields was due to higher-than-optimal pH values on Worst fields compared with Best fields (Supplemental Figure S1). Scores for soil inorganic P and K^+ did not differ significantly between any paired field types (Table 2) and were not limiting on most fields; instead, they were often well in excess of optimal concentrations based on MSU testing guidelines. Excess P inputs to cropped fields were evidenced by higher concentrations compared with NRC fields (Supplemental Table S1), resulting in lower soil health P scores on cropped fields (Table 2) and contributing to lower overall soil health scores on Best fields compared with Worst fields for Farms 8 and 9 in the Central region (Supplemental Figure S1).

Overall soil health scores for NRC fields when compared with Best and Worst were numerically higher and significantly higher, respectively (Table 2). This was driven chiefly by soil biological parameters, especially significantly higher SOM, POXC, and PMN, which scored higher on NRC fields compared with cropped fields (Table 2). High levels of significance were found between Worst vs. NRC field types for all biological measures, with NRC fields having higher scores. Among physical soil health parameters, NRC fields also scored significantly greater in AWC than Best and Worst fields. Although NRC fields scored numerically lower for SSR and significantly lower for SR compared with Best fields (Table 2), the NRC fields had living plant material and dense

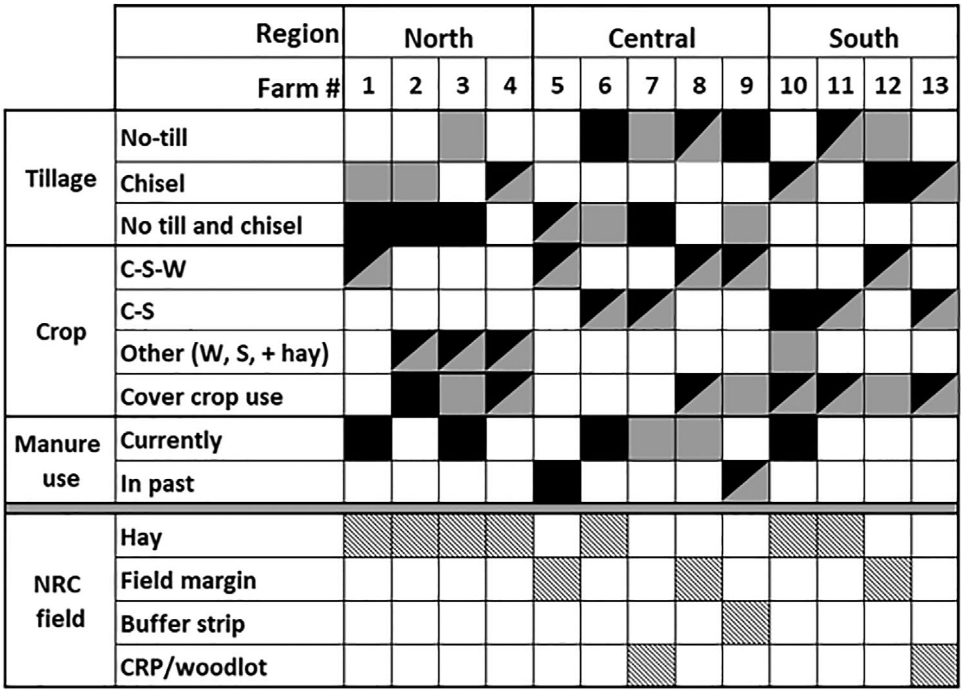

FIGURE 3 Management practices for Best fields (in black) and Worst fields (in gray) for each farm and region as stated by farmers. Where cells are split diagonally, both field types received the same management. Farmers used either no-till practices or chisel plow or some combination depending on the crop and year (C, corn; S, soybean; W, wheat). Cover crop use indicates regular use or some prior use of cover crops during recent management. Manure use indicates regularly used in current management or a known history of inputs. Hatched squares identify land use of non-row crop (NRC) field comparisons for each farm.

TABLE 2 Mean differences in soil health scores from paired *t* tests between all combinations of field types (Best, Worst, non-row crop [NRC]) for all soil health parameters and means of overall physical, biological (including C), and inorganic chemical categories and overall soil health

	Field comparison		
Parameter[a]	Best vs. Worst	Best vs. NRC	Worst vs. NRC
AS	17.7*	1.5	−17.1
AWC	34.3**	−23.5*	−39.5**
SR	8.8	18.3*	8.8
SSR	3.3	16.4	7.3
SOM	11.8**	−15.1*	−27.9**
POXC	0.6	−30.2*	−34.0**
MinC	20.4*	−3.8	−25.5**
PMN	11.0	−30.2*	−38.5**
pH	4.0	3.9	3.9
P	−10.2	−5.2	6.6
K	−2.9	−12.5	−8.1
CEC	4.4	−8.5	−15.2*
Physical	10.9**	3.2	−10.7
Biological	11.0*	−19.8***	−31.5***
Chemical	−1.2	−5.6	−3.2
Overall health	6.9*	−7.4	−14.9***

[a]AS, aggregate stability; AWC, available water capacity; SR, surface resistance; SSR, subsurface resistance; SOM, soil organic matter; POXC, permanganate oxidizable carbon; MinC, mineralizable carbon; PMN, potentially mineralizable nitrogen CEC; cation exchange capacity.

*Significant at the .05 probability level.

**Significant at the .01 probability level.

***Significant at the .001 probability level.

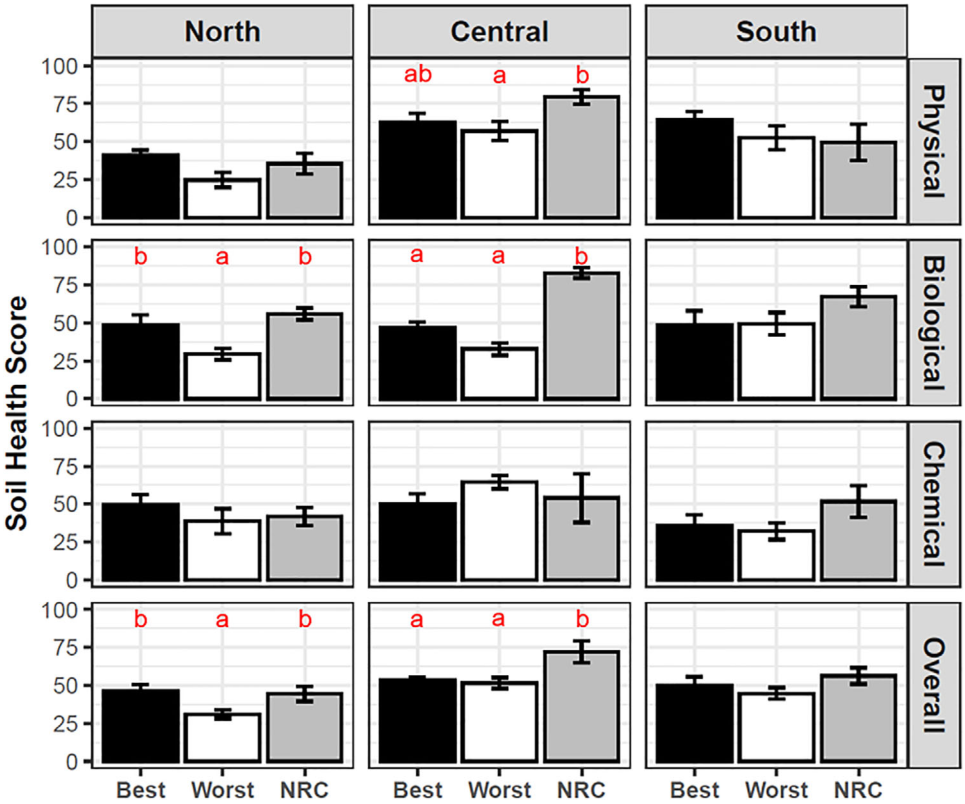

FIGURE 4 Soil health parameter means with standard errors for field type and region, shown separately for biological, physical, chemical parameters, and the overall soil health score. Where field type was significant in ANOVA ($P < .05$), different letters indicate significant differences between treatments using Tukey's HSD.

roots at sampling, making comparisons to cropped fields inappropriate. For chemical soil health parameters, no clear trend distinguished NRC fields from cropped fields, except higher CEC, which was only significantly different from the Worst fields (Table 2).

3.3 | Soil health test results by region

Patterns in soil health scores differed by region (Figure 4). The overall mean and the mean biological soil health scores among the three field types differed significantly in the North and Central regions. Means of physical parameters also differed significantly in the Central region. In the North and Central regions, differences in overall soil health reflected those found in biological and physical categories. No differences occurred in overall or category means in the South region (Figure 4).

In the North region, Best fields scored significantly higher in biological and overall soil health than Worst fields (Figure 4). Among individual parameters, only PMN scored significantly higher on the Best field type compared with the Worst field type (Figure 5). However, the significantly higher overall score of Best compared with Worst fields resulted from numerically higher means for all other parameters for Best fields in the North compared with Worst fields (Figure 5).

In the Central region between cropped field types, Best fields scored significantly higher for PMN and numerically higher for AS, AWC, SOM, and MinC (Figure 5); however, overall soil health scores between these two field types were similar (Figure 4). In South fields, no differences occurred between means of cropped field types for any individual soil health parameter, parameter category, or overall soil health score (Figures 4 and 5). Between cropped fields, AS, SSR, SR, and MinC were numerically greater on Best fields, whereas most biological and chemical parameters showed a less consistent contrast between these two field types.

Across all regions, only PMN was significantly higher on Best fields compared with Worst fields (Table 2). This difference in PMN scores was present in the North and Central regions but not in the South region. Contrasts between PMN on paired cropped fields for individual farms followed farmers' field assessments except in the South (Supplemental Figure S1), leading to no significant difference in the paired contrast for PMN overall (Table 2).

The lowest-scoring fields for overall soil health across all regions were the Worst fields in the North region (Figure 4).

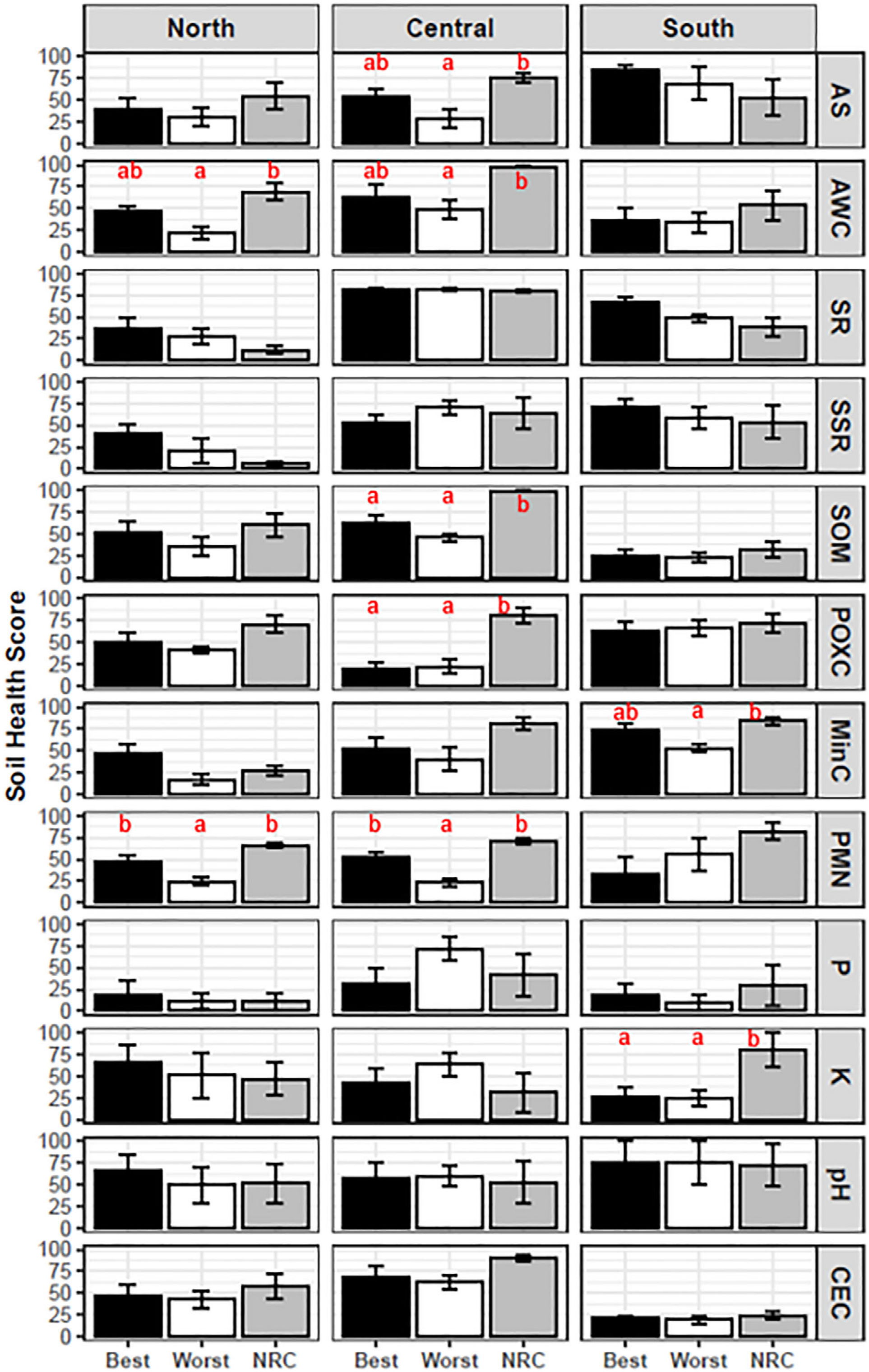

FIGURE 5 Soil health parameter means with standard error for field type and region by individual parameter (listed on right of panels). Where field type was significant in ANOVA ($P < .05$), different letters indicate significant differences between treatments using Tukey's HSD. AS, aggregate stability; AWC, available water capacity; CEC, cation exchange capacity; MinC, mineralizable C; PMN, potentially mineralizable N; POXC, permanganate oxidizable C; SOM, soil organic matter; SR, surface resistance; SSR, subsurface resistance.

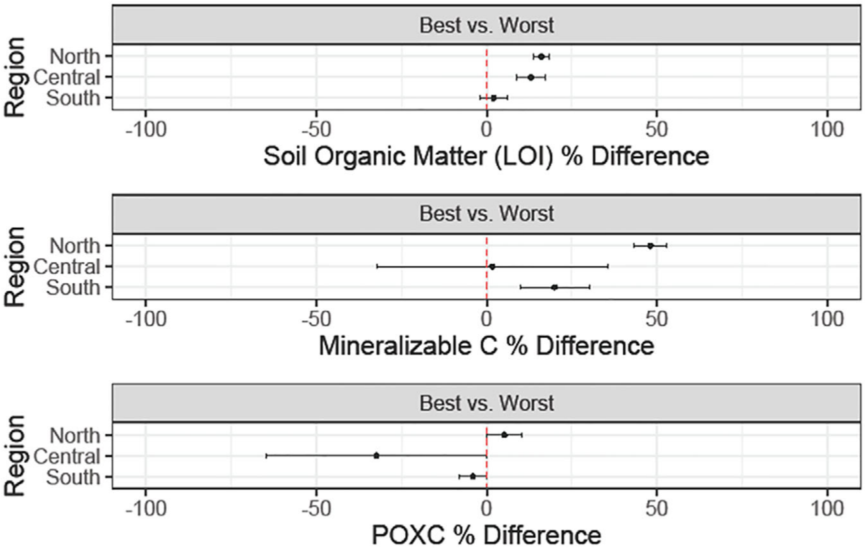

FIGURE 6 Mean percent difference in soil C indicators (percent soil organic matter, mineralizable C, and permanganate oxidizable carbon [POXC]) between Best and Worst fields selected by farmers. LOI, loss on ignition.

All Worst fields of the North region scored in the bottom 25th percentile for AWC, SR, SSR, MinC, and PMN (Figure 5). Across all regions, AS, AWC, SOM, and MinC were significantly higher on Best fields compared with Worst fields (Table 2), and among these parameters MinC reflected the most consistent contrast between these field types (Figure 5; Supplemental Figure S1).

Across all regions, the NRC fields scored significantly higher in paired comparisons with Best and Worst fields for AWC, SOM, POXC, and PMN, and the NRC field also scored significantly higher than Worst fields for MinC (Table 2). The NRC fields in the Central region had the highest overall soil health scores among all fields, which was mirrored in physical and biological categories (Figure 4). In this region, NRC fields scored significantly higher than Worst fields for physical parameters AS and AWC (Figure 5). Biological metrics were numerically higher on NRC fields, with SOM and POXC being significantly higher than both cropped field types (Figure 5). In the North region, NRC fields scores were similar to Best fields in overall soil health but were significantly higher than Worst fields (Figure 4), with physical measurements of AS and AWC and biological parameters SOM, POXC, and PMN following this pattern (Figure 5). In the South region, overall soil health scores of NRC fields were not different from Best or Worst fields (Figure 4).

3.4 | Soil C metrics

Paired *t* test analyses revealed that the Best fields had significantly greater SOM contents compared with the Worst fields when all regions were included in the analysis ($p < .01$; Table 2). On average, SOM values in Best fields were 13% greater than in the Worst fields (not shown). Similarly, MinC was significantly greater in Best fields compared with Worst fields when all regions were considered ($p < .05$; Table 2), with MinC being 45% greater in Best fields compared with Worst fields (not shown). In contrast, POXC values were similar between Best and Worst fields in all regions. Thus, SOM and MinC results coincided with farmer-defined Best and Worst fields, whereas POXC did not distinguish between the two types of fields (Table 2).

We calculated percent difference between Best and Worst fields to further compare the response of three different soil C tests (SOM, MinC, POXC) among the three regions (Figure 6). Indicators with positive values matched farmers' perceptions, based on their Best versus Worst field assessments. For a negative value, the indicator differed from the farmers' field designation. Of the three metrics, SOM had the smallest mean percent difference between Best and Worst fields, where mean (± SD) differences ranged from 2.1 ± 4.0% to 16.0 ± 2.3% across the three regions. The most sensitive test appeared to be MinC, with mean percent differences ranging from 1.6 ± 33.9% to 48.1 ± 4.9%. The small mean percent difference and large standard error in the Central region for MinC are the result of a large negative percent difference at Farm 8 (Supplemental Figure S2), whereas positive percent differences were reported at the other farms in the region. The poorest match with farmers' perceptions occurred with POXC, with generally negative and small mean percent differences.

4 | DISCUSSION

The development and validation of soil health metrics have occurred primarily in controlled field studies, with a focus on metrics' sensitivities to different soil management practices. To serve as tools for farmers to manage for soil health, these tests should also capture meaningful differences across farmers' fields. We sought to understand how soil health test results compared with farmers' field assessments across three regions of Michigan with distinct soil types. We hypothesized that that physical and biological measures of soil health would better align with farmers; field assessments compared with inorganic chemical parameters.

In general, on-farm soil health scores corresponded well with farmers' assessments of their soil's characteristics and performance, though specific soil health parameters varied in their capacity to distinguish between contrasting fields. Physical soil health indicators, particularly AS and AWC, successfully distinguished between farmers' assessments of their cropped fields, with Best fields having significantly higher scores (Table 2; Supplemental Table S1). Biological indicators, which include soil C, also supported farmers' assessments in discriminating among contrasting cropped fields, significantly so for measures of SOM and MinC. Inorganic chemical parameters related little if any to farmers' field assessments. Our results show how specific sets of soil health metrics align with farmer knowledge, demonstrating how testing implementation and interpretation can better guide soil health management.

The degree to which different soil health metrics followed farmers' field assessments varied by region and thus by soil type. Soil texture has a strong influence on the magnitude of some soil health parameters and thus how they are scored. For instance, soils with 3% SOM may score near 100 if they are coarse but below 50 if fine textured (Fine et al., 2017). In this survey, 92% of soils were coarse textured (Table 1), classified as sands, sandy loams, or loamy sands (Soil Survey Division Staff, 1993); however, differences in soil texture and soil type still influenced soil health scores across regions. For example, regional differences in soil type affected the sensitivity of some parameters, such as compaction. On finer-textured soils in the Central region, cropped fields scored in the top half of the distribution (Figure 5) for soil compaction (i.e., SR and SSR reflected low compaction), but these two parameters poorly reflected farmers' field assessments. By contrast, in the North and South regions, soils were more compacted, but SR and SSR corresponded better with farmers' field assessments (Supplemental Figure S1). Therefore, the usefulness of compaction scores to assess soil health differed by region.

In addition, biological indicators differed in their ability to discriminate between farmers' fields in different regions. For soils in the North and Central regions, which had higher SOM (Figure 5), this parameter better corresponded with cropped field assessments compared with the South region with lower SOM. Similarly, PMN scores aligned with farmers' field assessments in the North and Central regions, which had soils higher in SOM, but not in the South region. Furthermore, even though PMN scores differed by field type in all three regions, it only reflected farmers' field designations in the North and Central regions. This underscores the need to identify the specific parameters that are most useful for assessing soil health based on regional conditions.

Comparing paired fields within a farm minimized variability due to soil type and to some extent variable management of cropped fields across farms. Often, soil health parameter comparisons are made among explicitly tested management factors (e.g., tillage practice or rotation) within one site. Paired contrasts of cropped fields across widely varying sites revealed the relative ability of parameters to distinguish soil health and correspond to farmers' assessments of field performance. The NRC field served as a reference for characterizing soil health parameters of background soils because these farms had no previous soil health testing and experienced a range of soil management practices across farms (Figure 3). For example, the NRC paired comparisons indicated highly significant differences in soil health compared with cropped fields for AWC, SOM, POXC, and PMN and less clear differences for AS and MinC (Table 2). Even without the power of paired comparisons across farms, the NRC fields also reflected magnitude differences in parameters scores across regions (Figure 5).

For soil health assessments to be meaningful, they must reflect farmers' understandings of field performance. Our results suggest that chemical soil health metrics do not align with farmers' perceptions of field performance in that P, K, pH, and CEC did not significantly differ between Best and Worst fields (Figure 5). One explanation is that these farmers already typically test and directly manage inputs to adjust soil pH, P, and K levels. In fact, on two farms excess P inputs contributed strongly to poorer overall soil health scores (Supplemental Figure 1). In contrast, physical and biological metrics significantly differed between Best and Worst fields and thus strongly aligned with farmers' field assessments (Table 2). Every farmer in this study used some aspect of physical soil health, such as "how the soil works," drainage, or soil compaction, to describe either favorable characteristics of Best fields or problematic conditions of Worst fields (Figure 2). Across all regions, our results indicated that AWC best distinguished between cropped fields for physical soil health (Table 2), and in two regions, measures of soil compaction (SR and SSR) closely followed farmers' assessments. Thus, in contrast to chemical metrics, physical soil health parameters offer commonalities between farmers' experience and soil health testing by accurately distinguishing cropped fields, even when both fields scored relatively poorly.

Biological indicators of soil health also strongly reflected farmers' assessments of cropped fields (Table 2). The differences in field performance noted by farmers in our study support considerable research that has highlighted the importance of biological indicators for defining soil health (Culman et al., 2013; Veum et al., 2014; Wander et al., 2019). Our results demonstrate that some biological soil health parameters provide sufficient sensitivity to distinguish between field types and align with farmers' experience.

Soil organic matter corresponded significantly with farmers' field assessments, with differences of up to 16% between cropped field comparisons (Figure 6), though less so in the South region. Contrasts between cropped fields were even greater for MinC, with differences of up to 48% between the Best and Worst fields (Figure 6), including on cropped fields in the South region. Indeed, MinC on Best fields did not differ from NRC fields in paired field comparisons but was significantly lower on Worst fields (Table 2). Notably, POXC was poorest at distinguishing between paired crop fields, even though paired comparisons with NRC fields suggested highly significant sensitivity to contrasting management (Table 2).

The greater contrast in MinC values across cropped field types could reflect its sensitivity to management practices, which can increase MinC (Caudle et al., 2020). Practices such as the addition of composted material and conservation tillage can favor C stabilization and higher POXC, whereas increased tillage, cover cropping, and manure addition favor MinC (Hurisso et al., 2016). During in-depth interviews, farmers noted their use of a variety of these practices, with some trends by region; for instance, manure addition and reduced tillage were present in the Central region, whereas there was comparatively more tillage and use of cover crops in the North and South regions (Figure 3). Indeed, variable practices occurred within farms on different field types, indicating distinct management decisions for separate fields. Complex interactions between field management decisions, soil type, and different indicators of soil C highlight the need to increase the precision of MinC by standardizing measurement protocols (Wade et al., 2018). Our results indicate that the alignment of soil health metrics such as MinC with farmers' assessments of field performance make this an important soil health indicator on farms, especially in coarse-textured soils.

To our knowledge, no prior study has compared the sensitivity of POXC and MinC with farmers' assessments of field performance. Of the three metrics related to soil C, POXC did not reveal significant differences between the Best and Worst fields and also had the least accurate correspondence among biological parameters for reflecting farmers' field assessments. In contrast, MinC best captured field variability and was well aligned with farmers' characterizations of their fields, suggesting that MinC is a more meaningful metric for assessing field management decisions in the regions sampled. This is not surprising considering that recent research has demonstrated that MinC and POXC are indicators of different soil C processes (Hurisso et al., 2016; Morrow et al., 2016; Sprunger et al., 2019). Although both are considered indicators of different labile C fractions, MinC reflects microbial stimulation of CO_2 production following the re-wetting of soils (Franzluebbers et al., 2000), and thus it is a strong indicator of nutrient release and is a potential key predictor of agronomic performance (Culman et al., 2013; Sprunger et al., 2019). In contrast, although POXC is strongly correlated to SOM, it is associated with smaller and heavier particulate organic C fractions (Culman et al., 2012), which are often physically protected from microbial decomposition, and could demonstrate early indications of soil C stabilization.

5 | CONCLUSIONS

Soil health testing assesses biological, physical, and chemical attributes to ultimately guide the sustainable management of farm fields. Whether soil health tests align with farmers' own experience of agronomic performance may ultimately influence their impact on farm management practices. Results demonstrate that on-farm soil health testing can effectively distinguish differently performing fields across regions and can inform and strengthen farmers' knowledge of their fields. Although individual soil health parameters varied among regions, patterns of overall soil health scores were consistent with farmers' assessments of Best versus Worst fields. That inorganic chemical test parameters did not track with other metrics of soil health or farmers' assessments of their fields may in part be due to prior application of fertilizers and other inputs that remove most nutrient deficiencies and adjust pH. In contrast, physical and biological soil health parameters better captured variability in soil function and aligned with farmers' perceptions, highlighting an entry point for ecological management strategies through testing.

Although SOM values were consistently greater for the Best fields for all regions, MinC showed a better capacity to distinguish between farmers' field assessments of cropped fields, especially in coarser soils. Measures of POXC did not consistently align with farmers' field designations. Because POXC is an indicator of more stabilized soil C fractions and MinC of nutrient release, these metrics likely differ in their capacity to distinguish between a farmer's Best and Worst fields.

Soil health test results are more meaningful when merged with farmer knowledge. Given that soil health metrics vary by region and soil type, a participatory approach can inform testing protocols and interpretation to improve management practices and target specific constraints on fields. Combining soil health test results and farmer knowledge should facilitate the implementation of soil health management practices as well as guide outreach and on-farm research questions.

ACKNOWLEDGMENTS

We thank the 13 Michigan farmers who participated in this study. We also thank Paul Gross, James DeDecker, Colleen Forestieri, and MSU Extension and Soil and Water Conservation District Staff for help with recruiting farmers and coordinating sampling efforts; Jim Cave and Alessandra Zuniga for help in the field and sample analysis; Dr. Kim Chung and Dr. John Kerr for help designing interviews; and Steve Hamilton and Sieglinde Snapp for extensive feedback on earlier drafts of this work. Finally, we thank three reviewers for excellent feedback. Detailed attention from the editor greatly improved this work. This research was funded by USDA SARE Grants GNC14-196 and GNC14-192 with additional support from the NSF Long-term Ecological Research Program (DEB 1832042) at the Kellogg Biological Station and by Michigan State University AgBioResearch.

AUTHOR CONTRIBUTIONS

Brendan O'Neill, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Writing-original draft, Writing-review & editing; Christine D. Sprunger, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Writing-original draft, Writing-review & editing; G. Philip Robertson, Funding acquisition, Resources, Supervision, Validation, Writing-review & editing.

CONFLICT OF INTEREST

The authors declare no conflict of interest associated with the preparation of this manuscript.

REFERENCES

Andrews, S. S., Flora, C. B., Mitchell, J. P., & Karlen, D. L. (2003). Growers' perceptions and acceptance of soil quality indices. *Geoderma*, *114*(3–4), 187–213. https://doi.org/10.1016/S0016-7061(03)00041-7

Caudle, C., Osmond, D., Heitman, J., Ricker, M., Miller, G., & Wills, S. (2020). Comparison of soil health metrics for a Cecil soil in the North Carolina Piedmont. *Soil Science Society of America Journal*, *84*(3), 978–993. https://doi.org/10.1002/saj2.20075

Chantigny, M. H., Angers, D. A., Prévost, D., Vézina, L.-P., & Chalifour, F.-P. (1997). Soil aggregation and fungal and bacterial biomass under annual and perennial cropping systems. *Soil Science Society of America Journal*, *61*(1), 262–267. https://doi.org/10.2136/sssaj1997.03615995006100010037x

Cherry, K. A., Shepherd, M., Withers, P. J. A., & Mooney, S. J. (2008). Assessing the effectiveness of actions to mitigate nutrient loss from agriculture: A review of methods. *Science of the Total Environment*, *406*(1), 1–23. https://doi.org/10.1016/j.scitotenv.2008.07.015

Culman, S. W., Snapp, S. S., Freeman, M. A., Schipanski, M. E., Beniston, J., Lal, R., Drinkwater, L. E., Franzluebbers, A. J., Glover, J. D., Grandy, A. S., Lee, J., Six, J., Maul, J. E., Mirksy, S. B., Spargo, J. T., & Wander, M. M. (2012). Permanganate oxidizable carbon reflects a processed soil fraction that is sensitive to management. *Soil Science Society of America Journal*, *76*(2), 494–504. https://doi.org/10.2136/sssaj2011.0286

Culman, S. W., Snapp, S. S., Green, J. M., & Gentry, L. E. (2013). Short- and long-term labile soil carbon and nitrogen dynamics reflect management and predict corn agronomic performance. *Agronomy Journal*, *105*(2), 493–502. https://doi.org/10.2134/agronj2012.0382

De, M., Riopel, J. A., Cihacek, L. J., Lawrinenko, M., Baldwin-Kordick, R., Hall, S. J., & McDaniel, M. D. (2020). Soil health recovery after grassland reestablishment on cropland: The effects of time and topographic position. *Soil Science Society of America Journal*, *84*(2), 568–586. https://doi.org/10.1002/saj2.20007

Doran, J. W., & Parkin, T. B. (1994). *Defining and assessing soil quality. Defining soil quality for a sustainable environment* (pp. 1–21). John Wiley & Sons.

Drinkwater, L. E., Cambardella, C. A., Reeder, J. D., & Rice, C. W. (1996). *Potentially mineralizable nitrogen as an indicator of biologically active soil nitrogen. Methods for Assessing Soil Quality* (pp. 217–229). SSSA.

Drinkwater, L., & Snapp, S. (2007). *Understanding and managing the rhizosphere in agroecosystems*. Elsevier.

Fine, A. K., van Es, H. M., & Schindelbeck, R. R. (2017). Statistics, scoring functions, and regional analysis of a comprehensive soil health database. *Soil Science Society of America Journal*, *81*(3), 589–601. https://doi.org/10.2136/sssaj2016.09.0286

Franzluebbers, A. J., Haney, R. L., Honeycutt, C. W., Schomberg, H. H., & Hons, F. M. (2000). Flush of carbon dioxide following rewetting of dried soil relates to active organic pools. *Soil Science Society of America Journal*, *64*(2), 613–623

Gruver, J. B., & Weil, R. R. (2007). Farmer perceptions of soil quality and their relationship to management-sensitive soil parameters. *Renewable Agriculture and Food Systems*, *22*(4), 271–281. https://doi.org/10.1017/S1742170507001834

Hurisso, T. T., Culman, S. W., Horwath, W. R., Wade, J., Cass, D., Beniston, J. W., Bowles, T. M., Grandy, A. S., Franzluebbers, A. J., Schipanski, M. E., Lucas, S. T., & Ugarte, C. M. (2016). Comparison of permanganate-oxidizable carbon and mineralizable carbon for assessment of organic matter stabilization and mineralization. *Soil Science Society of America Journal*, *80*(5), 1352–1364. https://doi.org/10.2136/sssaj2016.04.0106

Idowu, O. J., van Es, H. M., Abawi, G. S., Wolfe, D. W., Schindelbeck, R. R., Moebius-Clune, B. N., & Gugino, B. K. (2009). Use of an integrative soil health test for evaluation of soil management impacts. *Renewable Agriculture and Food Systems*, *24*(3), 214–224. https://doi.org/10.1017/S1742170509990068

Karlen, D. L., Andrews, S. S., Wienhold, B. J., & Zobeck, T. M. (2008). *Soil quality assessment: Past, present and future*. https://digitalcommons.unl.edu/usdaarsfacpub/1203

Karlen, D. L., Goeser, N. J., Veum, K. S., & Yost, M. A. (2017). On-farm soil health evaluations: Challenges and opportunities. *Journal of Soil and Water Conservation*, *72*(2), 26A–31A. https://doi.org/10.2489/jswc.72.2.26A

Karlen, D. L., Hurley, E. G., Andrews, S. S., Cambardella, C. A., Meek, D. W., Duffy, M. D., & Mallarino, A. P. (2006). Crop rotation effects on soil quality at three northern corn/soybean belt locations. *Agronomy Journal*, *98*(3), 484–495. https://doi.org/10.2134/agronj2005.0098

Kettler, T. A., Doran, J. W., & Gilbert, T. L. (2001). Simplified method for soil particle-size determination to accompany soil-quality analyses. *Soil Science Society of America Journal*, *65*(3), 849. https://doi.org/10.2136/sssaj2001.653849x

Lal, R. (2016). Soil health and carbon management. *Food Energy Secur*, *5*(4), 212–222. https://doi.org/10.1002/fes3.96

Minasny, B., Malone, B. P., McBratney, A. B., Angers, D. A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z.-S., Cheng, K., Das, B. S., Field, D. J., Gimona, A., Hedley, C. B., Hong, S.-Y., Mandal, B., Marchant, B. P., Martin, M. P., Mcconkey, B. G., Mulder, V. L., O'Rourke, S. M. … Winowiecki, L. (2017). Soil carbon 4 per mille. *Geoderma*, *292*, 59–86. https://doi.org/10.1016/j.geoderma.2017.01.002

Moebius, B. N., van Es, H. M., Schindelbeck, R. R., Idowu, O. J., Clune, D. J., & Thies, J. E. (2007). Evaluation of laboratory-measured soil properties as indicators of soil physical quality. *Soil Science*, *172*(11), 895–912. https://doi.org/10.1097/ss.0b013e318154b520

Moebius-Clune, B. N., Moebius-Clune, D. J., Gugino, B. K., Idowu, O. J., Schindelbeck, R. R., Ristow, A. J., van Es, H. M., Thies, J. E., Shayler, H. A., McBride, M. B., Kurtz, K. S. M., Wolfe, D. W., & Abawi, G. S. (2016). *Comprehensive assessment of soil health – The cornell framework manual* (3.0 ed.). Cornell University.

Morrow, J. G., Huggins, D. R., Carpenter-Boggs, L. A., & Reganold, J. P. (2016). Evaluating measures to assess soil health in long-term agroecosystem trials. *Soil Science Society of America Journal*, *80*(2), 450–462. https://doi.org/10.2136/sssaj2015.08.0308

O'Neill, B. (2017). *Soil nutrient dynamics and shifting microbial processes with increasing rotational complexity in row crops*. Michigan State University.

Power, A. G. (2010). Ecosystem services and agriculture: Tradeoffs and synergies. *Philosophical Transactions of the Royal Society B*, *365*(1554), 2959–2971.

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Robertson, G. P., Gross, K. L., Hamilton, S. K., Landis, D. A., Schmidt, T. M., Snapp, S. S., & Swinton, S. M. (2014). Farming for ecosystem services: An ecological approach to production agriculture. *Bioscience*, *64*(5), 404–415. https://doi.org/10.1093/biosci/biu037

Robertson, G. P., Wedin, D., Groffman, P. M., Blair, J. M., Holland, E. A., Nadelhoffer, K. J., & Harris, D. (1999). *Soil carbon and nitrogen mineralization, nitrification, and soil respiration potentials. Standard soil methods for long-term ecological research* (pp. 258–71). Oxford University Press.

Roper, W. R., Osmond, D. L., Heitman, J. L., Wagger, M. G., & Reberg-Horton, S. C.. (2017). Soil health indicators do not differentiate among agronomic management systems in North Carolina soils. *Soil Science Society of America Journal*, *81*(4), 828–843. https://doi.org/10.2136/sssaj2016.12.0400

Saldaña, J. (2015). *The coding manual for qualitative researchers*. SAGE.

Schipanski, M. E., MacDonald, G. K., Rosenzweig, S., Chappell, M. J., Bennett, E. M., Kerr, R. B., Blesh, J., Crews, T., Drinkwater, L., Lundgren, J. G., & Schnarr, C. (2016). Realizing resilient food systems. *Bioscience*, *66*(7), 600–610. https://doi.org/10.1093/biosci/biw052

Sinsabaugh, R. L., Reynolds, H., & Long, T. M. (2000). Rapid assay for amidohydrolase (urease) activity in environmental samples. *Soil Biology & Biochemistry*, *32*(14), 2095–2097. https://doi.org/10.1016/S0038-0717(00)00102-4

Soil Survey Division Staff. (1993). *Examination and description of soils. Soil Survey Manual. Soil Conservation Service. United States Department of Agriculture Handbook 18*. USDA-NRCS.

Sprunger, C. D. (2015). *Root production and soil carbon accumulation in annual, perennial, and diverse cropping systems*. Michigan State University.

Sprunger, C. D., Culman, S. W., Palm, C. A., Thuita, M., & Vanlauwe, B. (2019). Long-term application of low C:N residues enhances maize yield and soil nutrient pools across Kenya. *Nutrient Cycling in Agroecosystems*, *114*(3), 261–276. https://doi.org/10.1007/s10705-019-10005-4

Tiemann, L. K., & Grandy, A. S. (2015). Mechanisms of soil carbon accrual and storage in bioenergy cropping systems. *GCB Bioenergy*, *7*(2), 161–174. https://doi.org/10.1111/gcbb.12126

Veum, K. S., Goyne, K. W., Kremer, R. J., Miles, R. J., & Sudduth, K. A. (2014). Biological indicators of soil quality and soil organic matter characteristics in an agricultural management continuum. *Biogeochemistry*, *117*(1), 81–99. https://doi.org/10.1007/s10533-013-9868-7

Wade, J., Culman, S. W., Hurisso, T. T., Miller, R. O., Baker, L., & Horwath, W. R. (2018). Sources of variability that compromise mineralizable carbon as a soil health indicator. *Soil Science Society of America Journal*, *82*(1), 243–252. https://doi.org/10.2136/sssaj2017.03.0105

Wade, J., Culman, S. W., Logan, J. A. R., Poffenbarger, H., Demyan, M. S., Grove, J. H., Mallarino, A. P., McGrath, J. M., Ruark, M., & West, J. R. (2020). Improved soil biological health increases corn grain yield in N fertilized systems across the Corn Belt. *Scientific Reports*, *10*(1), 3917. https://doi.org/10.1038/s41598-020-60987-3

Wander, M. (2004). Soil organic matter fractions and their relevance to soil function. In F. Magdoff, & R. Weil (Eds.), *Soil organic matter in sustainable agriculture*. CRC Press.

Wander, M. M., Cihacek, L. J., Coyne, M., Drijber, R. A., Grossman, J. M., Gutknecht, J. L., Horwath, W. R., Jagadamma, S., Olk, D. C., Ruark, M., Snapp, S. S., Tiemann, L. K., Weil, R., & Turco, R. F. (2019). Developments in agricultural soil quality and health: Reflections by the research committee on soil organic matter management. *Frontiers in Environmental Science*, *7*. https://doi.org/10.3389/fenvs.2019.00109

Weil, R. R., Islam, K. R., Stine, M. A., Gruver, J. B., & Samson-Liebig, S. E. (2003). Estimating active carbon for soil quality assessment: A simplified method for laboratory and field use. *American Journal of Alternative Agriculture*, *18*(01), 3–17. https://doi.org/10.1079/AJAA200228

Williams, H., Colombi, T., & Keller, T. (2020). The influence of soil management on soil health: An on-farm study in southern Sweden. *Geoderma*, *360*, 114010. https://doi.org/10.1016/j.geoderma.2019.114010

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: O'Neill B, Sprunger CD, Robertson GP. Do soil health tests match farmer experience? Assessing biological, physical, and chemical indicators in the Upper Midwest United States. *Soil Sci Soc Am J*. 2021;1-16. https://doi.org/10.1002/saj2.20233