Title
Sensitivity and variability of soil health indicators in a California cropping system

Journal
Soil Science Society of America Journal, 85(5)

ISSN
0361-5995

Authors
Lazicki, Patricia
Rodrigues, Jorge L Mazza
Geisseler, Daniel

Publication Date
2021-09-01

DOI
10.1002/saj2.20278

Peer reviewed

Received: 22 January 2021 | Accepted: 11 May 2021

DOI: 10.1002/saj2.20278

Soil Science Society of America Journal

NUTRIENT MANAGEMENT & SOIL & PLANT ANALYSIS

Sensitivity and variability of soil health indicators in a California cropping system

Patricia Lazicki | **Jorge L. Mazza Rodrigues** | **Daniel Geisseler**

Dep. of Land, Air, and Water Resources, Univ. of California, Davis, 1 Shields Ave., Davis, CA 95616, USA

Correspondence
Patricia Lazicki, Dep. of Land, Air, and Water Resources, Univ. of California, Davis, 1 Shields Ave., Davis, CA 95616, USA.
Email: palazicki@ucdavis.edu

Assigned to Associate Editor Carl Bolster.

Abstract

An indicator that is used to monitor whether a management practice is improving soil health must be sensitive to management changes. However, it should not be overly influenced by variations in sampling time or location, previous crop, or annual differences in weather or operations timing. In this study, we assessed the sensitivity and variability of several soil health indicators in long-term plots under typical farming practices in a Mediterranean climate. These plots have been conventionally or organically farmed in a corn (*Zea mays* L.)–processing tomato (*Solanum lycopersicum* L.) rotation for 25 yr. We sampled in both crop phases prior to planting and midseason for two consecutive years, analyzing subsamples taken from three adjacent locations per plot. Management was the most significant factor differentiating most indicators, particularly indicators of biological processes and C accumulation. Whereas management differences were consistent across sampling times, average indicator values for a management system often varied significantly between dates and years. Crop phases, conversely, were usually similar. Accounting for soil texture increased management sensitivity for aggregate stability and most C accumulation indicators. Sensitive indicators such as mineral N, particulate organic matter C, and mineralizable C had greater subsample variability than indicators measuring large, stable pools, such as total C. Our results show that indicators relating to organic C and biological processes most strongly differentiated the two systems, and underline the importance of using consistent sampling dates. They also suggest that an indicator dataset including both stable and sensitive indicators may be the most reliable to interpret.

1 | INTRODUCTION

Soil health is promoted as foundational to producing high yields of nutritious crops while protecting the environment and fighting climate change. Accordingly, "soil health building practices" are increasingly being experimented with by farmers and incentivized by government and private entities. Therefore, how best to assess whether soil health is actually improving is a topic of great practical importance and much recent debate (Karlen et al., 2019; Norris et al., 2020; Roper et al., 2019; Stewart et al., 2018; Wander et al., 2019).

Central to this debate are the concepts of indicator sensitivity, variability, and generalizability. A useful indicator must be able to detect improvements in some property of interest within a reasonable timeframe. However, it should not be affected by other factors to the extent that it is difficult to

Abbreviations: CONV, conventional; EC, electrical conductivity; FDA, fluorescein diacetate hydrolysis; ORG, organic; POM, particulate organic matter; POXC, permanganate-oxidizable C; SOC, soil organic carbon.

interpret or cannot be compared with other systems or over time (Hargreaves et al., 2019). For example, soil organic matter is an important component of soil health, but because it may take several years for a change to be detectable it could have limited usefulness for a farmer who wished to evaluate different management practices for increasing soil health (Hurisso et al., 2018). Conversely, plant-available N is necessary for crop growth, but changes so quickly and is so strongly affected by a variety of climactic and edaphic factors that it is also not very useful for assessing whether a soil is improving over time in response to management (Wander et al., 2019). Between these two extremes lie a host of potential soil health indicators that have been found to be more sensitive or less variable measures to monitor factors such as C storage, fertility, or structural stability (Hurisso et al., 2018).

However, an emerging body of work suggests that the "best" metrics, as well as the appropriate sampling protocols and the thresholds used to interpret them, are likely to be affected by cropping systems, climates, and edaphic properties like texture (i.e., Caudle et al., 2020; Chahal & van Eerd, 2019; Hurisso et al., 2016, 2018; Roper et al., 2017; Wade et al., 2016; Zuber et al., 2020). This need for regional specificity stems from the conceptual definition of soil health as "the continued capacity of a soil to function as a vital living ecosystem that sustains plants, animals, and humans" (Norris et al., 2020, p. 3196), as the threats and limitations to such functioning in an agricultural system depend on specific regional characteristics and crop requirements (Andrews et al., 2004). Hence, different indicators and different sensitivity may be appropriate. For example, it will be more important to be able to document very slight incremental changes in soil C in a region with intrinsically low capacity for organic matter accumulation than in a region with a high capacity.

Thus, as progress is being made toward developing standardized soil health assessment protocols (Norris et al., 2020), there is a need for regional studies that explore indicator sensitivity and variability in locally important management systems and soil types (Zuber et al., 2020). As most soil health assessment research has been conducted in rainfed grain and pasture systems, there is a particularly urgent need for more data from areas with the Mediterranean climates and irrigated high-value cropping systems typical of California. For example, in the "North American Project to Evaluate Soil Health Measurements," (https://soilhealthinstitute.org/north-american-project-to-evaluate-soil-health-measurements/), an exciting continent-scale project that aims to standardize soil health assessment by analyzing indicator sensitivity at long-term research sites, fewer than 5% of the sites are located in California or are in vegetable production (Norris et al., 2020). Additionally, that project's scope does not include spatial and temporal variability, which are likely to be important considerations in California given its variable winter precipitation and widespread use of drip irrigation. To ensure soil health

Core Ideas

- Sensitivity and variability of candidate indicators were tested on long-term research plots.
- Organic C and biological indicators were more sensitive than chemical and physical indicators.
- Sampling year and date were more important sources of variability than crop phase.
- Uniform sampling dates should be used if monitoring soil C changes over time.

assessment protocols are appropriate for the distinct climate, management, and spatiotemporal distribution of soil water occurring in California irrigated row crops, in-depth data on the sensitivity and variability of popularly promoted indicators are needed to complement research done across multiple sites and cropping systems.

Our goal in this study was to assess the sensitivity and variability of a wide range of soil health indicators in the Russell Ranch Century Experiment in the Central Valley of California, a long-term experiment in which plots in a corn (*Zea mays* L.) and processing tomato (*Solanum lycopersicum* L.) rotation have been under conventional (CONV) or organic (ORG) management since 1993. Long-term research trials provide stable, controlled systems that can be used to evaluate the benefits of a management practice and to help select and interpret appropriate soil health indicators (i.e., Diederich et al., 2019; Hurisso et al., 2016; Morrow et al., 2016). The uniform conditions allow for the systematic examination of indicators' relative susceptibility to other factors like weather. By sampling both systems in all crop phases prior to planting and during crop growth across 2 yr, we aimed to document how soil health indicators vary with crop phase, sampling timing and year, and how those differences compare in magnitude to those caused by management. In addition, the relatively wide textural gradient that occurs across the Century Experiment allows us to observe the effect of soil texture differences on indicator sensitivity and variability with factors other than management.

We hypothesize that in the Century Experiment: (a) long-term organic management will have resulted in significant differences in all indicators, particularly increasing those relating to soil organic C (SOC) storage and biological function; (b) indicators such as enzyme activity and C mineralization, which measure biological processes, will be more affected by other sources of variability (crop phase, sampling time and year) than indicators that measure large, stable pools, such as total C; (c) indicators with the greatest sensitivity to management will also have the greatest variability with factors other than management; and (d) that adding a textural covariate will

TABLE 1 Summary of crop management operations and soil and plant sampling in 2018 and 2019

Operation	Management dates 2018	Management dates 2019	Operation	Soil and plant sampling dates 2018	Soil and plant sampling dates 2019
Bed disking and listing	Oct. (2017)	Oct. (2018)	Preplant sampling (all)	4 Apr.	12 Apr.
Cover crop seeding	Nov. (2017)	Nov. (2018)	Tomato midseason soil and plant sampling	13 June	13 June
Cover crop mow/disk/list	23 Feb.	23 Feb.	CONV corn midseason soil and plant sampling	27 June	5 July
Compost spreading	20 Apr.	24 Oct. (2018)	ORG corn midseason soil and plant sampling	20 July	30 July
Compost incorporation (2018 only), cultivation, bed rolling	21–26 Apr.	23–25 Apr.	Tomato bulk density sampling	14 Aug.	30 July
CONV corn seeding, starter NPK	21 Apr.	4 May	Corn bulk density sampling	5 Oct.	30 July
CONV tomato starter NPK	26 Apr.	25 Apr.	Tomato fruit hand-harvest	14 Aug.	23 Aug.
Tomato transplanting	1 May	29 Apr.	Corn grain hand-harvest	5 Oct.	3 Oct.
ORG corn seeding	25 May	3 June			
Tomato harvest	30 Aug.	4 Sept.			
Corn harvest (all)	5 Oct.	11–14 Oct.			

Note. CONV, conventional; ORG, organic.

increase the significance of the management effect for most indicators.

2 | MATERIALS AND METHODS

2.1 | Site description

The experiment was conducted in plots with a corn–tomato rotation at the Century Experiment at the Russell Ranch Sustainable Agriculture Facility in northern California (38˚32′24″ N, 121˚52′12″ W). In this experiment, a conventional management system utilizing synthetic fertilizer and winter fallow is contrasted with a certified organic system with yearly application of composted poultry manure and a winter legume cover crop. Pest control in each system is conducted according to local grower practice. Detailed site and management information can be found in Tautges et al. (2019) and Schmidt et al. (2018). Briefly, the experiment was laid out in 1993 as a randomized complete block design with three blocks. Two blocks are placed on Rincon silty clay loam soil (fine, smectitic, thermic Mollic Haploxeralfs) and the third on Yolo silt loam (fine-silty, mixed, superactive, nonacid, thermic Mollic Xerofluvents). The rotation and management were designed to reflect typical local grower practice. Each crop phase (corn or tomato) of each management system is represented in each block in each year on 0.4-ha replicate plots, each consisting of 48 152.4-cm-wide beds. This scale allows plots to be large enough to be managed with commercial scale farm equipment but small enough to be relatively uniform (Denison et al., 2004). In both systems, beds are cultivated and rolled prior to planting. Following harvest, residues are disked in to a depth of 15 cm, using four passes for corn and two passes for tomato. Tomato also receives inseason cultivation. In the ORG systems, additional operations included disking in the cover crop (two passes), bed listing in early spring, and compost incorporation (detailed below). In 2015, a single subsurface drip line was installed at a depth of 25 cm down the center of each bed. Prior to 2015, all plots were furrow irrigated. Tillage operations do not exceed 25 cm in depth.

2.2 | Management during the 2018 and 2019 growing seasons

Timing of management operations for the 2018 and 2019 growing seasons is summarized in Table 1. In the ORG plots, a cover crop consisting of hairy vetch (*Vicia villosa* Roth), faba bean (*Vicia faba* L.), and cereal oat (*Avena sativa* L.) was seeded in November, terminated by mowing in late February, and then incorporated by disking. For the 2018 crops, composted poultry manure was broadcast applied in late April at the rate of 4 t ha^{-1} and incorporated by disking to 15 cm. For the 2019 crop, composted poultry manure at the same rate was broadcast over the top of the corn and tomato residues in fall 2018, which were then chopped and disked to incorporate. In both years, beds were rolled to prepare the seedbeds in late April. In the ORG and CONV systems, tomatoes were transplanted down the center of the bed in early May. Corn was seeded in double rows in late April, early May in the

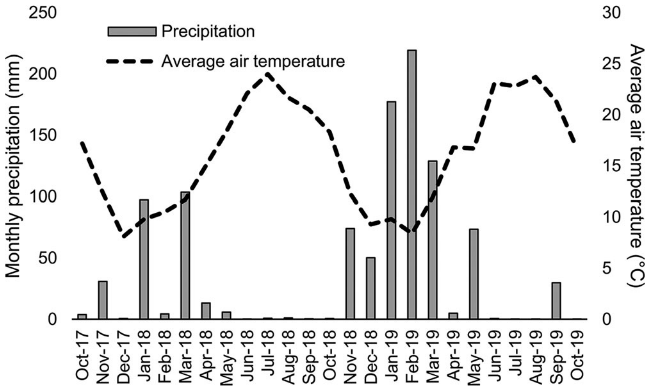

FIGURE 1 Monthly precipitation and average air temperature for Davis, CA, from October 2017 through October 2019

CONV system, and in late May 2018 and early June 2019 in the ORG system. The later seeding date for ORG corn was chosen because insect pressures resulted in a poor stand in 2018, requiring replanting. In the CONV system, both corn and tomato received an application of 56 kg N ha^{-1} as 8–24–6 banded at planting. No other sources of P and K were added. In addition, urea ammonium nitrate (32%) was water-run through the drip lines several times during the growing season, for an annual total of 235 kg N ha^{-1} in corn and 200 kg N ha^{-1} in tomato. In both systems, tomatoes were mechanically harvested in late August and corn in late October.

Total winter (October through March) rainfall prior to the 2018 and 2019 growing seasons was 240 and 650 mm, respectively (California Irrigation Management Information System, Davis, CA station). Average air temperature during both growing seasons (April through September) was 20 ˚C (Figure 1).

2.3 | Soil and plant sampling

For each crop in each year, soil samples were taken prior to planting and in the early reproductive phase (early green fruit in tomato and tasseling in corn; Table 1). At each sampling date, samples were taken from 4.5-m long microplots located 46 m from the southern edge of the plot and at least 7 m from any field edge. Different microplots were sampled in each year. Microplots were sampled rather than the whole plot to avoid sampling from other experiments on the same plots. Samples were taken at three locations within each microplot, spaced 1.8 m apart within a single bed. At each location, several cores were taken 20 cm from the center drip line to a depth of 25 cm, within 0.3 m of a central point, and composited such that three separate subsamples were taken per plot for further analysis. The top 10 cm were discarded, as this portion is not consistently within the wetting zone of these subsurface-drip irrigated beds. Sensitivity and variability estimates reported are, therefore, of microplots rather than whole plots, and are more susceptible to localized random variability. In 2019, the soil was wetter at sampling and a 4.5-cm diameter Edelman combination auger was used to avoid compaction. At least 800 g of soil was obtained from each location, mixed thoroughly, immediately placed on ice, and kept at 4 ˚C until analysis.

Preplant samples were taken on 4 Apr. 2018 and 12 Apr. 2019. Due to the difference in operations timing between years, this was prior to manure application in 2018, but several months subsequent in 2019 (Table 1). Midseason samplings were timed according to crop phenologic stage for each year, crop, and management. Bulk density samples were taken from the top 0–15 cm once per season, just before tomato or corn harvest in 2018, and on July 30th in 2019. Samples were taken 20 cm from the center drip line using a soil corer fitted with a 4.5-cm diameter plastic sleeve.

2.4 | Soil analyses

Field-moist soils were sieved to 4.75 mm. Gravimetric moisture content was determined by drying a 20-g subsample of homogenized soil at 105 ˚C for 24 h. Within 1 wk of sampling, two duplicate samples of 6-g moist soil were weighed into glass 40-ml vials. One set was extracted for 1 h with 30 ml 0.5 M potassium sulfate (Mulvaney, 1996), filtered through medium retention filter paper, and analyzed for ammonium (NH_4–N) and nitrate (NO_3–N) using colorimetry based on the Berthelot reaction (Forster, 1995; Verdouw et al., 1978) and a single reagent method (Doane & Horwath, 2003), respectively. The sum of these was used to calculate the total mineral N. The duplicate sample was fumigated with chloroform for 24 h, and then extracted in the same manner. Organic C was measured in the extracts of fumigated and unfumigated soils, and microbial biomass C was calculated as the difference between C concentrations in fumigated and unfumigated extracts, divided by an adjustment factor k_e of 0.35 (Horwath & Paul, 1996). The remainder of the soil was air-dried and ground to pass through a 2-mm sieve for analyses of additional indicators relating to soil chemistry, organic C pools, biological processes, and physical structure.

Chemical analyses were electrical conductivity (EC) and pH, measured in a 2:1 water/soil slurry (Thomas, 1996), bicarbonate-extractable P analyzed colorimetrically (Olsen P; Watanabe & Olsen, 1965), and base cations Ca, Mg, K, and Na, extracted with ammonium acetate at pH 7 (Helmke & Sparks, 1996; Suarez, 1996). For indicators relating to organic matter pools, total C and N were measured by dry combustion on soils ground to a fine powder (Nelson & Sommers, 1996). Effervescence tests with HCl suggested no appreciable carbonates were present. The particulate organic matter (POM) C and N were assessed using size fractionation to 53 μm followed by dry combustion as described by Cambardella and Elliott (1992). Permanganate-oxidizable C (POXC) was assessed on duplicate 2.5-g samples using the protocol described by Weil et al. (2003), with modifications described by Culman et al. (2012).

Two different indicators of biological processes were measured. Mineralizable C was measured as the CO_2–C mineralized over 72 h from two duplicate centrifuge tubes each containing 6 g of dried soil rewet to 60% water holding capacity, where water-holding capacity was defined as the water concentration of a saturated soil sample after 1 h free draining in a filter-paper lined funnel (Wade et al., 2016). Water adjustments were made from the top, using a pipet (Wade et al., 2018) and both samples were placed in a single sealed jar fitted with a rubber septum for gas sampling and incubated in the dark at 25 ˚C for 72 h. Headspace CO_2–C was measured on an infrared gas analyzer (IRGA; Qubit Systems). As an index of heterotrophic enzyme activity, fluorescein diacetate hydrolysis (FDA) was measured using the method described by Green et al. (2006), with modifications proposed by Prosser et al. (2011). Briefly, three replicate 1-g samples were weighed into 50-ml centrifuge tubes and 30 ml of tris(hydroxymethyl)aminomethane buffer (pH 7.6) was added. A volume of 0.30 ml of FDA solution were added to two of the replicates, and an equivalent amount of acetone was added to the third as a control. Samples were shaken for 3 h on a reciprocal shaker at room temperature. After shaking, the reaction was paused by the addition of 1.2 ml of acetone and briefly vortexing. Samples were allowed to settle for 10 min, after which 1.5 ml was pipetted into 1.5-ml centrifuge tubes and centrifuged at 8,800 *g* for 5 min. One ml of supernatant was pipetted into cuvettes and measured against a fluorescein standard curve at a wavelength of 490 nm.

Water-stable aggregates and bulk density represented the physical indicators. We determined aggregate stability using a method slightly modified from that of Kemper & Roseneau (1986). This method was chosen as preliminary tests in our lab showed it to be the least affected by differences in moisture content at sampling and differences in sampling equipment. About 30 g of air-dried soil sieved to 2 mm was placed on top of a 1-mm sieve, and gently shaken to remove the fraction <1 mm. Then, 4 g of 1-to 2-mm aggregates were poured in a thin layer on the surface of a 250-μm sieve with a diameter of 6 cm and placed in a clean, dry, 10-cm-diameter soil tin. The soil was gradually wet by capillarity by adding deionized water down the side of the tin until the water level reached the level of the sieve. The sample was allowed to equilibrate for 15 min, after which the water level was brought up and the sample was raised and lowered by hand to a height of 1.3 cm for 3 min at a rate of 35 oscillations per minute, ensuring that the sieve mesh did not rise above the surface of the water. The can was placed in a 105 ˚C oven for 24 h to determine the weight of the unstable fraction. The sieve containing the stable fraction was transferred to another can containing a solution of 2% sodium hexametaphosphate, in which the remaining aggregates were completely dispersed. The residual sand particles remaining on the sieve were washed with deionized water and transferred to aluminum weigh boats and dried for 24 h at 105 ˚C. The stable aggregate fraction was calculated by subtracting the weight of the unstable and coarse fractions from the initial 4-g sample. To determine bulk density, the height of each 4.5-cm diameter core was measured and the soil within the core was weighed after drying at 105 ˚C for 24 h.

As texture varies substantially across the site, particle size distribution values for the top 30 cm of each plot were obtained from the Russell Ranch records (Kate Scow and Jessica Chiartas, Davis, CA).

2.5 | Statistical analyses

Descriptive statistics for each indicator were generated using PROC UNIVARIATE in SAS (SAS corporation). The mean, median, SD, skewness, and coefficients of variation (CVs) were assessed for each indicator for the full dataset, and in each management system across all subsample locations, replicate blocks, crops, dates, and years. As many of the variables showed moderate to extreme skewness, we reported the medians and the CVs calculated based on the medians rather than the means (Wade et al., 2018). In most cases the difference between mean- and median-based values was minimal. Outliers were not removed unless values were physically impossible (such as negative numbers), or there was documented evidence of error during the analysis. Overall CV values of 10% or less were categorized as low, 11–30% as intermediate, and above 30% as high.

Locations within plots were combined to calculate means and CVs for each plot × block × crop × management × date × year combination ($n = 48$; three blocks × two crops × two mgt × two dates × 2 yr). The plot mean values were used to assess the main and interactive effects of management, crop, date and year using ANOVA in PROC GLIMMIX in SAS. The experiment was analyzed as a crossover design with repeated measures, as detailed in Tao et al. (2015). "Plot" was considered to be the subject, and the two crop sequences (tomato–corn, corn–tomato) considered as sequences of treatments administered over two periods (2018 and 2019), and "date" was treated as a repeated measure. This approach accounts for the fact that a single subject (plot) received two crop phases. In each phase, two measurements were taken in close enough proximity as to be dependent (Tao et al., 2015). The crossover design's assumption that carryover effect of crop phase between the two periods is minimal was supported by preliminary exploratory analysis using spaghetti plots for individual plots across dates and years. A compound symmetry autocorrelation structure was used, as it is often more appropriate for small sample sizes than more complex designs

(Schaalje et al., 2002), and preliminary exploration showed it generally gave similar or lower Akaike Information Criterion values than more complex designs for most indicators, and was more likely to converge. Management, crop phase, sampling date, and year were regarded as fixed. "Year" was treated as a fixed effect, because in the context of this experiment, the inference space is the combination of weather, operations timing, sampling equipment, and technician variability which contributed to the difference between the 2018 and 2019 growing seasons rather than among all growing seasons. Three random terms were used. In the first, date was specified as a repeated measure with compound symmetry autocorrelation structure, and the subject defined as "plot*crop" (Tao et al., 2015). The second term was "plot within sequence," which specifies that the same subject is sampled in two different periods (years). The third term, "block," allows it to be analyzed as a randomized complete block design. Denominator degrees of freedom were adjusted using the Kenward–Roger method that considers small sample sizes and the potential for unbalanced data due to missing samples (Shaalje et al., 2002). The method is considered to be conservative in comparison to other approaches. Mean separation was performed using Tukey's HSD test with the LINES option in PROC GLIMMIX. The relative strength of a fixed effect was assessed using the adjusted p values and F statistics. The assumptions of homogeneity of variance and normally distributed residuals around a mean of zero were checked using Levene's test and visual assessment of residual plots. Indicators were log-transformed as needed to meet assumptions. As the blocking does not entirely account for texture variation across the site, we also performed the same test with sand or clay concentration as a covariate. Standard errors were back-transformed using the delta transformation where necessary.

To assess whether variability among samples taken in close proximity differed between management systems, crop phase, dates, or years, the same procedure was carried out on the CVs of the three locations within each sampled plot. Descriptive statistics were generated for the in-plot CVs using PROC UNIVARIATE. These CVs (hereafter "subsample variability") represent the combined variability introduced by short-range spatial variation as well as routine differences in sampling, processing, and analytical technique.

3 | RESULTS

3.1 | Mean differences and variability associated with organic and conventional management

Averaged across locations, plots, crops, dates, and years, values for ORG plots were greater than those for CONV plots in most indicators. Exceptions were pH, extractable Mg, and soil C/N ratio, where CONV exceeded ORG plots, and the POM C/N ratio, where they did not differ (Tables 2 and 3). The indicator with the largest proportional difference between CONV and ORG systems was mineral N, which was almost four times greater in the ORG than CONV plots (Table 2). The rate of FDA hydrolysis, MBC, and EC were all more than twice as high in the ORG than CONV plots. The aggregate stability index was only about 20% greater in the ORG than CONV plots, but this difference was still significant (Tables 2 and 3). Bulk density did not differ between the two systems.

The SD for most indicators was generally larger for ORG than CONV plots (Table 2). However, as median values were also greater in the organic systems, the CVs did not consistently differ between the two systems.

Among the chemical indicators, the variability within both ORG and CONV management systems across dates, crops, and years was lowest for extractable Ca and Mg, intermediate for K and Na, and very high for mineral N and Olsen P (Table 2). Among the different indicators of organic matter pools, POXC and SOC, total N, and C/N ratio varied the least. Microbial biomass C and POM C, N, and C/N ratio had intermediate CVs. The CVs for both biological process indicators were relatively high for both management systems. For the physical indicators, variability was lowest for bulk density and highest for the aggregate stability index.

3.2 | Other sources of variability and their interactions with management

As assessed by the size of the F value, for most indicators, the main effect of management was much larger than the effects of the other sources of variability (Table 3). Exceptions were extractable K and total C (which had stronger main effects for sampling date than for management) and mineral N, Olsen P, and aggregate stability (which all had stronger main effects for year than for management). POM C/N, which did not differ between management systems, had larger effects for all three other sources of variability than for management (Table 3). In every instance where a significant interaction with management occurred, differences between management systems remained significant across both levels of the other factors (crops, dates, years) (Table 3; Supplemental Table S1).

Few indicators had a significant main effect for crop phase (Table 3). These included a strong effect ($p < .001$) for EC and mineral N, and less significant effects ($p < .05$) for Na, SOC, and FDA hydrolysis. For all these variables, values for soils sampled in the corn phase (that is, prior to corn planting and during corn growth) were greater than soils sampled in the tomato phase (Supplemental Table S1). The crop effect was constant across sampling dates and years (Supplemental Table S2). In all cases, the main effect for crop

TABLE 2 Descriptive statistics for soil quality indicators measured in conventional (CONV) and organic (ORG) systems at Russell Ranch

Indicator	Median		Minimum		Maximum		SD		CV[a]	
	CONV	ORG	CONV	ORG	CONV	ORG	CONV	ORG	CONV	ORG
									%	
	Chemical indicators									
EC, μS cm^{-1}	96	197	57	121	141	405	19	43	20	22
pH	7.7	7.3	7.3	6.9	8.1	7.8	0.19	0.23	2.5	3.1
Mineral N, mg kg^{-1}	4.9	18	0.68	4.2	9.8	39	2.2	10	45	55
Olsen P, mg kg^{-1}	34	50	6.7	25	131	96	28	20	83	41
Ca, mg kg^{-1}	2,124	2,485	1,872	2,198	2,508	3,171	164	190	8	8
Mg, mg kg^{-1}	1,911	1,744	1,559	1,485	2,287	2,270	183	160	10	9
K, mg kg^{-1}	171	287	106	173	268	596	36	85	21	30
Na, mg kg^{-1}	44	85	23	54	77	200	14	20	31	24
	Organic matter pool indicators									
Total N, %	0.10	0.15	0.08	0.12	0.11	0.22	0.01	0.02	7.4	10
Total C, %	0.93	1.3	0.73	1.0	1.1	1.6	0.09	0.12	10	8.9
Total C/N	9.5	8.5	8.2	6.8	11.1	9.1	0.45	0.31	4.7	3.6
MBC, mg kg^{-1}	171	368	90	184	262	617	41	90	24	25
POXC, mg kg^{-1}	309	465	223	341	391	582	37	56	12	12
POM N, mg kg^{-1}	94	181	49	103	158	264	24	32	25	18
POM C, mg kg^{-1}	925	1,828	468	983	1,709	2,969	237	364	26	20
POM C/N	10.3	10.1	7.1	7.9	18.9	13	1.8	0.9	17	9
	Biological process indicators									
FDA, mg kg^{-1} h^{-1}	8.8	20.9	4.4	11	20	60	2.5	7.9	28	38
Mineralizable C, mg kg^{-1}	48	83	16	36	76	169	15	29	32	34
	Physical indicators									
Aggregate stability	0.38	0.46	0.11	0.24	0.63	0.68	0.13	0.10	34	21
Bulk density, g cm^{-3}	1.35	1.32	1.2	1.1	1.5	1.6	0.07	0.12	5.3	8.8

Note. Values calculated for three subsamples from three replicate blocks, across corn and tomato crops, preplant and midseason sampling dates, and 2018 and 2019 growing seasons ($n = 72$). CV, coefficient of variation; EC, electrical conductivity; FDA, fluorescein diacetate hydrolysis rate; MBC, microbial biomass C; Mineralizable C, CO_2–C respired from rewet soil after 72 h of incubation; Mineral N, NH_4–N+NO_3–N ; POM, particulate organic matter (>53 μm); POXC, permanganate oxidizable C.

[a]CV calculated on a median basis given the skewed distribution of many of the parameters.

phase was small compared with that for management. The only interaction of crop with management was for Na, likely due to a single high outlier in an organic tomato plot in spring of 2019.

The means, medians, and variability for selected indicators are shown in Figures 2 and 3. Because crop phase was usually only a minor source of variability, the two crop phases are combined for each management by date–year combination.

Most indicators differed significantly between sampling dates (Table 3). Only Olsen P, Na, the soil C/N ratio, POM C, POM N, and aggregate stability showed no significant main effect for date (Table 3). For the chemical indicators, the extractable K was always greater prior to planting than during plant growth for both management systems and years, whereas the date effect for the other chemical indicators tended to vary depending on the year (Figure 2). Date did not significantly interact with management for any of the chemical indicators (Table 3). Mean values for most of the organic matter and biological indicators declined significantly between preplant and midseason for both systems in both years (Figure 3). For POM C and POM N, which did not have a significant main effect for date but had significant date by management interactions, values were greater in midseason than prior to planting for CONV plots, but tended to be greater at preplant than midseason for ORG plots. Soil total C and N, POXC, and the mineralizable C all had significant management by date interactions, in which the differences between dates tended to be larger in the ORG plots than in the CONV plots, and the differences between management systems were generally larger prior to planting than during midseason (Figure 3).

TABLE 3 *F* values and Tukey-adjusted *p* values for the main effects of organic or conventional management (Mgt), crop phase (corn or tomato), sample date (preplant or midseason), period within sampling sequence (year 2018 or 2019), and for the interactions of all other effects with management, as well as the effect of Mgt when sand concentration was included as a covariate in the model and for the sand covariate

Indicator	Mgt	Crop phase	Sample date	Period (year)	Mgt × crop	Mgt × date	Mgt × year	Mgt [a] (sand covariate)	Sand covariate
	Chemical indicators								
EC	339***	47.9***	12.7***	18.8**	0.46	0.88	4.39	297***	0.01
Ph	129***	1.72	65.1***	34.2***	1.95	0.75	0.91	124***	0.74
Mineral N	375***	29.6***	38.5***	425***	0.07	0.51	56.5***	364***	0.06
Olsen P	5.45a	0.29	0.09	35.5***	0.05	0.09	0.54	5.14a	0.45
Ca	43.1***	0.41	9.10**	33.9***	0.06	2.38	2.82	165***	20.8**
Mg	8.9*	0.20	9.35**	0.02	0.09	2.75	3.41	34.1**	16.5**
K	50.8***	2.12	345***	42.7**	2.25	2.47	1.62	66.5***	3.67
Na	251***	9.17*	3.78	0.30	14.7*	2.41	4.56	208***	0.50
	Organic matter indicators								
Total N	164***	5.27	109***	0.0	0.03	11.3*	0.85	382***	11.9*
Total C	70.6***	6.51*	128***	3.5	0.32	19.9***	0.49	268***	27.7**
Total C/N	127***	1.49	3.96	13.3**	2.24	0.01	5.94*	168***	4.07
MBC	222***	3.68	97.7	12.5**	1.83	2.03	4.91	443***	9.17*
POXC	193	4.90	165***	71.6***	0.51	22.1***	16.5*	384***	9.13*
POM N	111***	1.21	0.74	8.00*	0.07	12.9*	0.10	218***	10.5
POM C	439***	2.78	2.09	11.2**	0.12	13.7*	0.96	434***	0.80
POM C/N	0.02	0.12	14.6**	0.81	0.00	2.20	0.33	0.12	14.2**
	Biological activity indicators								
FDA	235***	6.60*	123***	55.3***	2.15	2.91	0.44	319***	1.61
Mineralizable C	268***	0.08	38.8***	128***	0.03	15.5*	12.4*	263***	0.65
	Physical indicators								
Bulk density	0.67	0.55	–	12.66*	0.00	–	0.03*	0.61	0.71
ASI	9.47*	1.50	2.0	63.9***	1.99	4.49*	5.25	71.8***	42.8***

Note. ASI, aggregate stability index; EC, electrical conductivity; FDA, fluorescein diacetate hydrolysis rate; MBC, microbial biomass C; Mineralizable C, CO_2–C respired from rewet soil after 72 h of incubation; Mineral N, NH_4–N+NO_3–N; POM, particulate organic matter (>53 µm); POXC, permanganate oxidizable C.

[a] *F* value and significance level for management when the percent sand was included as a covariate.

*Significant at the .05 probability level.

**Significant at the .01 probability level.

***Significant at the .001 probability level.

The bundle of effects encompassed by the period or "year" effect (including weather, operations timing, crop health, and the sampling equipment and technician) was also highly significant for most indicators (Table 3). The only factors that did not differ significantly between years were extractable Mg and Na, soil total C and N, and the POM C/N ratio. For the other indicators, greater values were observed in 2019 than in 2018, except for pH and Ca, which were greater in 2018 (Figures 2 and 3). In general, these trends were consistent across management systems. Notable exceptions were mineral N and mineralizable C, in which differences between years were much more significant for ORG than CONV plots (Figures 2 and 3), and for POXC and POM C/N, in which 2019 values were greater than 2018 values in CONV but not ORG systems (Figure 3).

3.3 | Effect of a textural covariate on management

To determine the extent to which accounting for the texture differences helped differentiate management systems, all ANOVAs were also run with sand or clay as a covariate. The two particle sizes yielded similar results. However, sand differed more widely across the site (8.6–35.5%), and adding it as a covariate yielded greater improvements in sensitivity than clay. In addition, as the sum of sand, silt, and clay must equal 100%, sand can also be used as a proxy for the silt–clay fraction, which is known to be important in SOC accumulation (Hassink, 1997). Thus, only the results with the sand covariate are presented. Because blocking ensured that the two management systems were represented approximately

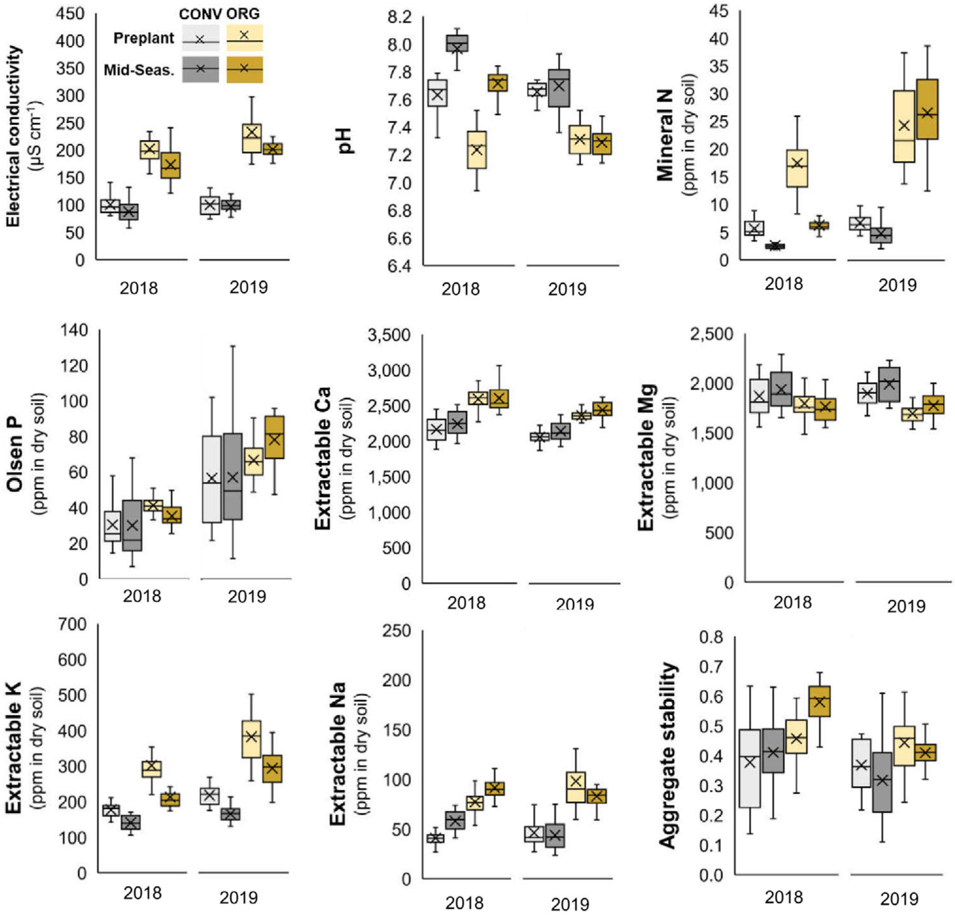

FIGURE 2 Distribution of chemical and physical indicators measured in conventional (gray boxes) and organic (brown boxes) systems, at preplant (lighter boxes) and midseason (darker boxes), in 2018 and 2019. Boxes represent three subsamples taken from two crop phases in three replicate blocks ($n = 18$). Horizontal lines represent medians and X represents means

equally across the texture gradient, adding the sand covariate improved sensitivity by decreasing within-group variability rather than by increasing differences between means. Adding the sand covariate had the greatest effect on aggregate stability, reducing the SE of the means by 64%, increasing the F value sevenfold and lowering the p value from .02 to .0004 (Table 3). Adding sand also increased F values for most of the organic matter pool indicators. The largest effect was for total C, for which the SE was decreased by 54% and the F value was increased fourfold. Sand was also a significant ($p < .01$) covariate for extractable Ca and Mg among the chemical indicators. The sand covariate was not significant for either of the biological process indicators.

3.4 | Subsample variability

As indicated by the median CVs of the three subsamples taken within each microplot at each sampling, the indicators with the most variability were mineral N, Olsen P, and the mineralizable C (Supplemental Table S3). These three indicators had median subsample CVs of 15% or greater, and maximum CVs reaching 40–80%. The EC, Na, aggregate stability, FDA hydrolysis, and the organic matter fractions MBC, POM C, and POM N also had high subsample variability, with median CVs of 10% or greater. Total C and N were the least variable organic matter pool indicators, and Ca, Mg, and pH were the least variable chemical indicators, with median subsample CVs of less than 5%. Analysis of variance performed on the subsample CVs found that Olsen P and POM C/N ratio were significantly more variable in the CONV system than in the ORG system, and that the ORG system had significantly more variability in extractable K ($p < .05$; data not shown). Otherwise, subsample variability did not appear to differ between management systems.

4 | DISCUSSION

Developing a useful minimum data set and appropriate sampling protocols for soil health assessment requires regional data about how sensitive different indicators are to

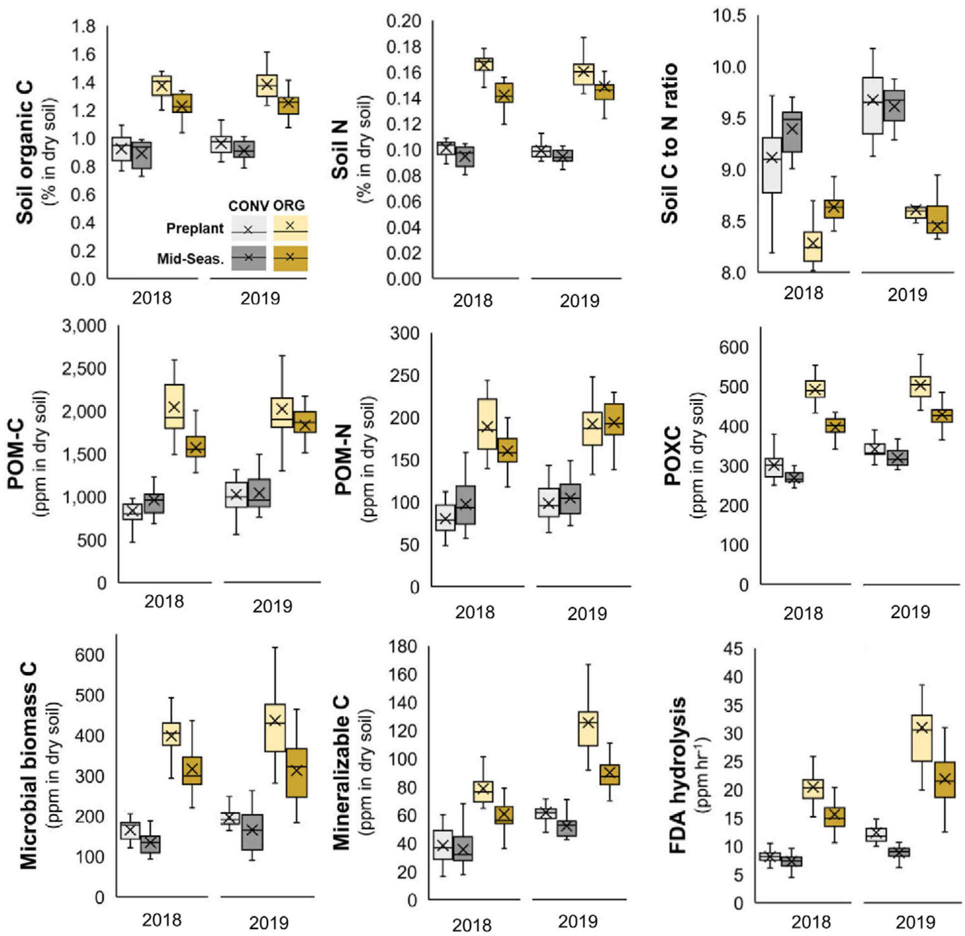

FIGURE 3 Distribution of organic matter and biological process indicators measured in conventional (gray boxes) and organic (brown boxes) systems, at preplant (lighter boxes) and midseason (darker boxes), in 2018 and 2019. Boxes represent three subsamples taken from two crop phases in three replicate blocks ($n = 18$). POM C and N = C and N in particulate organic matter >0.53 µm. POXC = permanganate oxidizable C. Mineralizable C = CO_2 measured after 72-h incubation from dried soil rewet to 60% water-holding capacity. FDA = fluorescein diacetate hydrolysis rate. Horizontal lines represent medians and X represents means

management practices and the degree to which they are affected by spatial or temporal variability. In this study, we assessed how strongly several soil health indicators were affected by an organic or conventional management system, and the strength of that signal compared with differences associated with the crop phase within a rotation, sampling date, and year of sampling. We also measured the intrinsic variability of each indicator by analyzing closely spaced subsamples within each plot, and observing whether adding a texture covariate improved an indicator's ability to differentiate management systems.

4.1 | Differences between conventionally and organically managed systems

As predicted by our first hypothesis, 25 yr under different management had resulted in pronounced differences in soils under ORG and CONV management practices. This is likely due to greater C inputs in the ORG system. A recent analysis of the Russell Ranch Century Experiment found that C inputs to the ORG system exceeded those to the CONV system by an average of 67% between 1993 and 2012 (Tautges et al., 2019). Comparison with C stocks measured in the 0–15-cm depth for these sites in 1993, 2003, and 2012 (Kong et al., 2005; Tautges et al., 2019) suggests that C in the plough layer of the ORG system has steadily increased over the 25 yr of the experiment's duration, whereas in the CONV system it has remained relatively stable.

Judging by the magnitude of the system differences as well as the F value, the most sensitive indicators to management were the chemical indicators EC, mineral N, and extractable Na, the labile organic matter indicator POM C, and both biological process indicators. The increases in labile C pools, biological activity, and somewhat improved physical structure in the ORG compared with the CONV plots, are all expected changes in systems that use organic amendments and cover crops (Jian et al., 2020; Lori et al., 2017). These variables are interpreted as "more is better," and thus suggest improved soil health in the ORG system (Andrews et al., 2004). The limited improvement in physical structure despite their long history of cover crop and compost addition may be due to their additional tillage and to the use of subsurface drip irrigation, which has been shown to reduce macroaggregate stability in these soils (Schmidt et al., 2018). Although bulk density samples were taken only from the surface 0–15 cm once during the season, the lack of difference between the ORG and CONV systems suggests surface compaction was similar between the two systems. Increases in EC and Na are also sometimes observed when manure is used (Miller et al., 2005). However, although on average EC and Na levels in the ORG system were twice as high as in the CONV, they were an order of magnitude below thresholds at which either corn or tomato yields might be reduced (Maas & Grattan, 1999), and thus do not yet constitute threats to soil function (Andrews et al., 2004). Similarly, increases in available P are also characteristic of manure-fertilized systems (Maltais-Landry et al., 2015). The significantly lower pH in the ORG than CONV systems, however, is unusual (Lori et al., 2017). Soil pH is often found to be lower in conventional systems, likely because of the acidification of ammonium fertilizers (Clark et al., 1998; Lori et al., 2017). We do not have a plausible explanation for this small but significant decrease in pH. However, in these slightly alkaline soils, a change of pH toward neutral constitutes a soil health improvement (Andrews et al., 2004).

4.2 | Indicator sensitivity to crop phase

It is interesting that crop phase affected so few indicators, as prior to planting the crop phase determines the quality and quantity of last season's residue additions. During the growing season, operations timing, irrigation management, and root growth are different between the two crops. Two recent California-wide surveys that measured a variety of soil properties as potential indicators for potential N mineralization concluded that residues may have a strong effect on the relevant pools and processes, especially in soils with a low soil organic matter content such as those at the study site (Miller et al., 2019; Wade et al., 2016). Corn contributes considerably more residue C than tomato (Tautges et al., 2019), and has a greater C/N ratio. The general similarity of most organic matter indicators between the two crop phases suggests that any influence of residue quantity and quality was eclipsed by the stronger effects of management history, sampling date, and year, or that the responses were too transient to be captured by our sampling scheme. The lower EC and mineral N in the tomato phase could both be due to immobilization of mineral N by corn residues, as high NO_3–N can contribute to EC. Although tomato received a lower N rate than corn in the CONV system, the consistent difference across management systems suggests it was not merely a function of differences in N additions, as corn and tomato received similar fertility in the ORG system. In addition, processing tomatoes are generally poorer N scavengers than corn (Hills et al., 1983), which may have resulted in greater residual N in the following crop. Interestingly, most of the indicators that were affected by crop phase (mineral N, EC, and FDA hydrolysis) were among those which differed most in magnitude between the ORG and CONV plots. This partially supports our hypothesis that the most sensitive indicators will also tend to vary more with nonmanagement sources. However, the fact that in all cases management differences were much stronger than those of either previous or current crop suggests that crop type does not need to be an important consideration in indicator selection in California systems with annual crops where residues are incorporated in fall and soil samples are taken the following spring.

4.3 | Indicator sensitivity to sampling date

Sampling date was a more important source of variability than crop phase. Between the preplant sampling and the midseason sampling, many changes occur, including the decomposition of labile organic material and release of nutrients as the weather warms, root growth and exudation, changes in moisture content and distribution as the soil dries and irrigation begins, and eventually depletion of nutrients in the root zone and shading of the soil surface as the crop grows. In addition, in 2018, composted manure was applied between the preplant and midseason sampling. The tendency of the organic matter pool and biological process indicators to decline in both years between preplant and midseason in spite of the compost application timing suggests that the rapid decomposition of labile soil organic material and root nutrient uptake were the most important processes contributing to between-date variability. The idea of a rapid decomposition pulse is also suggested by the lack of crop by date interactions, because the midseason corn sampling occurred later than that of tomato. The tendency for the organic matter pool indicators to decline more sharply in the ORG system than the CONV system also suggests they are reflecting cycles of organic matter inputs and decomposition. Interestingly, SOC, total N, and POXC, which are thought to represent processed, stable pools (Hurisso et al., 2016), showed very significant date effects. Indeed, when the effect of texture was not accounted for, the date effect was stronger for total C and N than was management.

4.4 | Indicator consistency across years

Almost all of the indicators were significantly greater in 2019 than 2018. The two factors which differed most strikingly between the two years were the much higher rainfall prior to the 2019 growing season (nearly triple that prior to 2018), and the fact that preplant sampling was done prior to compost application in 2018 but after compost application in 2019. Additional factors that varied between years included a larger cover crop biomass in 2019, the specific locations sampled in each plot, the probe type (necessitated by the wetter soil in 2019, and which may have influenced sampling depth), and minor differences in operations timing. Because the year effect was generally significant for both management systems and few indicators had significant management by year interactions, it appears that the changes in compost application timing and the larger cover crop biomass (which only affected the ORG system) had unexpectedly little additional effect on most indicators. Only mineral N and mineralizable C increased much more strongly in the ORG system than the CONV system between years, suggesting a more direct response to the management changes. This finding is in line with other studies that conclude that mineralizable C is an indicator that quickly responds to practices that increase fertility. Thus it is often related to N mineralization potential, especially in systems receiving organic N sources (Franzluebbers, 2020b; Hurisso et al., 2016). However, it is possible that other indicators responded more strongly to the between-year management changes at some point not captured by our sampling scheme.

Notably, SOC and total N concentrations were among the only indicators that did not change between the years. This is intriguing, given the relatively strong changes that occurred between years. It is important to note that as this data is derived from a limited number of samples taken within a small microplot, the results are more susceptible to random local variability than a greater number of samples taken to represent the whole plot. However, the uniform plot management, the strength of the effect ($p < .001$ for SOC differences between dates and $p = .10$ for differences between years), and the fact that dates were significantly different in both years despite a shift in microplot locations, suggest that the low variability between years in comparison with dates is not a sampling artifact. Together with the variability observed in SOC and total N between dates, our results are consistent with the idea of a large, stable pool of organic matter that, while subject to seasonal fluctuations as material is added and decays, overall is likely to change only slowly (Stott, 2019; Wuest, 2014). Similar seasonal fluctuations of SOC around a comparatively stable mean have been observed under wheat in Eastern Oregon (Wuest, 2014) and under cotton in the Texas High Plains (Burke et al., 2019). These authors attributed their findings to the presence of the growing plants, and to the fluctuations of ephemeral C pools and of bulk density. Our results strongly support the suggestion that uniform sampling dates be used for samples taken to monitor SOC changes over time (Hurisso et al., 2018).

4.5 | Effect of texture on indicator sensitivity

We hypothesized that adding sand as a covariate would decrease the within-group variability for most indicators and increase the significance of the management effect, as texture was not entirely addressed in the blocking. Adding a covariate greatly improved the sensitivity of aggregate stability and many of the organic C pool indicators to management, especially total C. This was expected, as finer textured soils are thought to form more stable aggregates and provide more surface area for the long-term protection and stabilization of SOC (Six et al., 2002). However, the only chemical indicators for which adding a texture covariate improved sensitivity to management were extractable Ca and Mg. In this young alluvial landscape, extractable Ca and Mg concentrations are both greater in older, more weathered, and finely textured soils. However, the ORG system also receives Ca inputs through the manure compost, whereas the CONV receives no supplemental Ca. Competition for cation exchange sites probably led to preferential leaching of Mg in the ORG system, and the observed result that ORG had greater Ca but less Mg than CONV.

Contrary to our hypothesis, however, sand concentration was not significantly related to most of the chemical indicators, the more labile C pools, or either biological process indicator. Many of these properties were likely directly affected by inputs such as fresh organic matter, fertilizer, or irrigation water, or by exports such as crop uptake. Therefore, the lack of relationship to texture may be because their dynamics are governed mostly by external supply. Another possibility is that their relationship with sand is nonlinear. For example, Franzluebbers and Haney (2018), testing rewet C and N mineralization across a range of soil textures, observed the highest rates of rewet respiration in medium textured soils. Whereas examination of our raw data shows that indicators like POM C and FDA hydrolysis are, in fact, greatest in some of the medium textured soils, adding a quadratic covariate did not improve the model fit, suggesting this was not an important mechanism.

4.6 | Indicator variability among subsamples

The variability between samples taken close together within the same plot on the same date reflects the variation of an indicator over short distances, and how sensitive it is to differences in sampling or analytic techniques. It is important as it determines how many samples must be taken or analyzed to get a representative value for an area. As predicted by our third hypothesis, the smallest subsample variability was seen in indicators such as total C and N and extractable Ca and Mg, which represent large pedogenic or stable pools. It is worth noting that bulk density was more variable than total C, and thus C stocks, which are calculated using both values, would be more variable than either.

The large median CV of mineralizable C compared with other soil tests has also been noted by several other studies (Hurisso et al., 2018; Morrow et al., 2016; Wade et al., 2018). Different protocols are used to measure mineralizable C and many factors contribute to its analytical variability, such as the method of rewetting and of CO_2 measurement (Wade et al., 2018), as well as sample size and volume (Franzluebbers, 2020a). The variability in our dataset may have been lower had a larger sample size been used, as our sample size of 12 g is smaller than is common. For example, in the studies reviewed by Wade et al. (2018), sample sizes ranged from 10 to 40 g, and work by Franzluebbers (2020a) suggests variability is minimized when sample sizes exceed 70 g. Using 75-g samples and analyzing CO_2 with the base trap method, that study observed an average CV of 8%. However, the range of CVs for mineralizable C at Russell Ranch, 1.7–39% with a median of 15%, is comparable to many reported in the literature for this analysis. Wade et al. (2018), using gel paddles to measure CO_2, report interlaboratory CVs ranging from 4 to 53% (median, 16%). Examining analytical variability for 10-g samples using the IRGA method to measure CO_2, Hurisso et al. (2018) obtained CVs in the range of 13–23%. Similarly, Franzluebbers et al. (2018) report a mean CV of 14% for 50-g samples analyzed using the base trap method.

The relatively low variability of POXC, another emerging soil health test that is thought to represent a more processed portion of labile C, is also consistent with recent findings that the POXC method is both more repeatable in the field and is subject to less analytical error than mineralizable C (Hargreaves et al., 2019; Hurisso et al., 2018; Wade et al., 2020). The median in-plot CV of 5.4% is comparable to the mean CV of 7–8% reported by Wade et al. (2020) for samples of different soil orders analyzed in triplicate by several different labs, and is lower than the analytical variability of 9–21% observed by Hurisso et al. (2018).

Aggregate stability, which had a relatively high median variability of 13%, is another indicator for which several alternate protocols exist (e.g., wet-sieving, slaking, and raindrop impact approaches). These rely on different physical principles and the sensitivity and analytical variability likely differ based on the approach used. Important considerations include sampling method, the use of dried or moist samples, sieving, storage time, and method of wetting or disruption (Kemper & Roseneau, 1986; Stott, 2019). In this study, in which we took samples with a soil probe and air-dried them prior to analysis due to practical considerations, we found in preliminary tests that the slaking and raindrop impact approaches were too influenced by differences in soil moisture at sampling to be useful. However, the small aggregate size fraction (1–2 mm) necessitated by the wet-sieving method may have resulted in greater analytical variability than if the sampling and processing methods had allowed a different analytical approach to be used.

The very great variability in available P is surprising, particularly in the CONV system. This contrasts with Hurisso et al. (2018), who observed great analytical precision and spatial correlation on another available P measure, Mehlich P. Subsurface drip-irrigated systems may have greater spatial variability than rainfed systems, as the targeted delivery of small quantities of water and nutrients to a limited area may result in zones of depletion in the wetting zone and high concentrations near the bed edges (Lazcano et al., 2015). The fact that the variability was particularly great in the CONV system suggests that the variability observed may also partly be due to the fact that P is applied in a starter band. As P is relatively immobile in the soil, banding may lead to zones of high and low concentrations (Beegle, 2005), and increased sampling density and specialized sampling schemes are recommended in fields that receive banded (or fertigated) fertilizers (Geisseler & Miyao, 2016; Lazcano et al., 2015).

5 | CONCLUSIONS

After 25 yr of contrasting management, the ORG and CONV managed plots in these drip-irrigated Mediterranean annual cropping systems had developed strong differences in almost every indicator tested, particularly in indicators relating to organic matter pools and biological processes. Our results have implications for minimum data set selection and protocol standardization in California and beyond. Given the great variability associated with management, soil type, and year observed, even at a single site in two consecutive years, the results underline the necessity of calibrating robust regional thresholds using data from many soil types and management systems and over several years. The significant covariation of SOC and aggregate stability with soil texture over such a relatively short distance confirms the utility of always performing a soil texture analysis at sites where these indicators are being monitored. The highly significant variability between dates, particularly for biology and fertility-related indicators, confirm the call by Hurisso et al. (2018) for making an effort to sample at the same time each year. As differences between the systems were often greater at the preplant than midseason sampling, preplant sampling appears to be more sensitive as well as more convenient. The crop type at this site had very little effect on most indicators compared with other sources of variability, suggesting that in similar systems the previous or current crop would not greatly influence indicator selection or interpretation. However, further work is necessary before applying these conclusions to other crop types or systems. Whereas the intrinsic variability of mineralizable C and POXC appear comparable to those observed elsewhere, the greater variability in chemical indicators like P suggest that specialized sampling schemes that take into account additional variability introduced over time by practices like fertilizer banding and subsurface drip irrigation may be appropriate. Finally, given the potential for tradeoffs between sensitivity and consistency, including both stable measurements like total C and sensitive measurements like mineralizable C in an minimum data set may be the best strategy.

ACKNOWLEDGMENTS

Many thanks to Israel Herrera, Nicole Tautges, Kate Scow, and the field staff of Russell Ranch. Thanks as well to the Geisseler, Rodrigues, and Melotto labs for the many hours of skilled lab and field assistance, Jessica Chiartas for the use of her soil texture data, and to Mohamed Sahtout for statistical consultation. Special thanks to Carl Bolster and three anonymous reviewers for their excellent critiques and suggestions, which greatly improved this manuscript. This work is supported by the Agriculture and Food Research Initiative – Foundational Program (award no. 2018-67019-27798) from the USDA National Institute of Food and Agriculture.

AUTHOR CONTRIBUTIONS

Patricia Lazicki, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing-original draft, Writing-review & editing; Jorge L. Mazza Rodrigues, Funding acquisition, Project administration, Writing-review & editing; Daniel Geisseler, Conceptualization, Funding acquisition, Project administration, Supervision, Writing-review & editing.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ORCID

Patricia Lazicki https://orcid.org/0000-0002-2481-3116
Jorge L. Mazza Rodrigues https://orcid.org/0000-0002-6446-6462
Daniel Geisseler https://orcid.org/0000-0002-5483-9982

REFERENCES

Andrews, S. S., Karlen, D. L., & Cambardella, C. A. (2004). The soil management assessment framework: A quantitative soil quality evaluation method. *Soil Science Society of America Journal*, *68*, 1945–1962. https://doi.org/10.2136/sssaj2004.1945

Beegle, D. (2005). Assessing soil phosphorus for crop production by soil testing. In J. T. Sims & A. N. Sharpley (Eds.), *Phosphorus: Agriculture and the environment* (pp. 123–143). AgronomyMonographs 46. ASA, CSSA, and SSSA.

Burke, J. A., Lewis, K. L., Ritchie, G. L., Moore-Kucera, J., Delaune, P. B., & Keeling, J. W. (2019). Temporal variability of soil carbon and nitrogen in cotton production on the Texas High Plains. *Agronomy Journal*, *111*(5), 2218–2225. https://doi.org/10.2134/agronj2019.02.0066

Cambardella, C. A., & Elliott, E. T. (1992). Particulate soil organic-matter changes across a grassland cultivation sequence. *Soil Science Society of America Journal*, *56*, 777–783. https://doi.org/10.2136/sssaj1992.03615995005600030017x

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

Chahal, I., & Van Eerd, L. L. (2019). Quantifying soil quality in a horticultural-cover cropping system. *Geoderma*, *352*, 38–48. https://doi.org/10.1016/j.geoderma.2019.05.039

Clark, M. S., Horwath, W. R., Shennan, C., & Scow, K. M. (1998). Changes in soil chemical properties resulting from organic and low-input farming practices. *Agronomy Journal*, *90*(5), 662–671. https://doi.org/10.2134/agronj1998.00021962009000050016x

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

Denison, R. F., Bryant, D. C., & Kearney, T. E. (2004). Crop yields over the first nine years of LTRAS, a long-term comparison of field crop systems in a Mediterranean climate. *Field Crops Research*, *86*(2), 267–277. https://doi.org/10.1016/j.fcr.2003.08.014

Diederich, K. M., Ruark, M. D., Krishnan, K., Arriaga, F. J., & Silva, E. M. (2019). Increasing labile soil carbon and nitrogen fractions require a change in system, rather than practice. *Soil Science Society of America Journal*, *83*(6), 1733–1745. https://doi.org/10.2136/sssaj2018.11.0458

Doane, T. A., & Horwáth, W. R. (2003). Spectrophotometric determination of nitrate with a single reagent. *Analytical Letters*, *36*(12), 2713–2722. https://doi.org/10.1081/AL-120024647

Forster, J. C. (1995). Soil nitrogen. In K. Alef & P. Nannipieri (Eds.), *Methods in applied soil microbiology and biochemistry* (pp. 79–87). Academic Press.

Franzluebbers, A. J. (2020a). Soil mass and volume affect soil-test biological activity estimates. *Soil Science Society of America Journal*, *84*(2), 502–511. https://doi.org/10.1002/saj2.20038

Franzluebbers, A. J. (2020b). Soil-test biological activity with the flush of CO_2: V. Validation of nitrogen prediction for corn production. *Agronomy Journal*, *112*(3), 2188–2204. https://doi.org/10.1002/agj2.20094

Franzluebbers, A. J., & Haney, R. L. (2018). Evaluation of soil processing conditions on mineralizable C and N across a textural gradient. *Soil Science Society of America Journal*, *82*(2), 354–361. https://doi.org/10.2136/sssaj2017.08.0275

Franzluebbers, A. J., Pershing, M. R., Crozier, C., Osmond, D., & Schroeder-Moreno, M. (2018). Soil-test biological activity with the flush of CO2: I. C and N characteristics of soils in corn production. *Soil Science Society of America Journal*, *82*(3), 685–695. https://doi.org/10.2136/sssaj2017.12.0433

Geisseler, D., & Miyao, G. (2016). Soil testing for P and K has value in nutrient management for annual crops. *California Agriculture*, *70*(3), 152–159. https://doi.org/10.3733/ca.2016a0007

Green, V. S., Stott, D. E., & Diack, M. (2006). Assay for fluorescein diacetate hydrolytic activity: Optimization for soil samples. *Soil Biology* *and Biochemistry*, *38*(4), 693–701. https://doi.org/10.1016/j.soilbio.2005.06.020

Hargreaves, S. K., Dejong, P., Laing, K., Mcquail, T., & Van Eerd, L. L. (2019). Management sensitivity, repeatability, and consistency of interpretation of soil health indicators on organic farms in southwestern Ontario. *Canadian Journal of Soil Science*, https://doi.org/10.1139/cjss-2019-0062

Hassink, J. (1997). The capacity of soils to preserve organic C and N by their association with clay and silt particles. *Plant and Soil*, *191*(1), 77–87. https://doi.org/10.1023/A:1004213929699

Helmke, P. A., & Sparks, D. L. (1996). Lithium, sodium, potassium, rubidium and cesium. In D. L. Sparks, A. L. Page, P. A. Helmke, R. H. Loeppert, P. N. Soltanpour, M. A. Tabatabai, C. T. Johnston, & M. E. Sumner (Eds.), *Methods of soil analysis: Part 3. Chemical methods* (2nd ed., pp. 551–574). ASA and SSSA.

Hills, F. J., Broadbent, F. E., & Lorenz, O. A. (1983). Fertilizer nitrogen utilization by corn, tomato, and sugarbeet. *Agronomy Journal*, *75*(3), 423–426. https://doi.org/10.2134/agronj1983.00021962007500030002x

Horwath, W. R., & Paul, E. A. (1996). Microbial biomass. In R. W. Weaver, S. Angle, P. Bottomley, & D. Bezdiecek (Eds.), *Methods of soil analysis: Part 2. Microbiological and biochemical properties* (2nd ed., pp. 753–773). ASA and SSSA.

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

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*(4), 939–948. https://doi.org/10.2136/sssaj2018.03.0098

Jian, J., Du, X., & Stewart, R. D. (2020). Quantifying cover crop effects on soil health and productivity. *Data in Brief*, *29*, 105376. https://doi.org/10.1016/j.dib.2020.105376

Karlen, D. L., Veum, K. S., Sudduth, K. A., Obrycki, J. F., & Nunes, M. R. (2019). Soil health assessment: Past accomplishments, current activities, and future opportunities. *Soil and Tillage Research*, *195*, 104365. https://doi.org/10.1016/j.still.2019.104365

Kemper, W. D., & Rosenau, R. C. (1986). Aggregate stability and size distribution. In A. Klute (Ed.), *Methods of soil analysis: Part 1. Physical and Mineralogical Methods* (2nd ed., pp. 425–442). ASA and SSSA.

Kong, A. Y. Y., Six, J., Bryant, D. C., Denison, R. F., & Van Kessel, C. (2005). The relationship between carbon input, aggregation, and soil organic carbon stabilization in sustainable cropping systems. *Soil Science Society of America Journal*, *69*(4), 1078–1085. https://doi.org/10.2136/sssaj2004.0215

Lazcano, C., Wade, J., Horwath, W. R., & Burger, M. (2015). Soil sampling protocol reliably estimates preplant NO_3^- in SDI tomatoes. *California Agriculture*, *69*(4), 222–229. https://doi.org/10.3733/ca.v069n04p222

Lori, M., Symnaczik, S., Mäder, P., De Deyn, G., & Gattinger, A. (2017). Organic farming enhances soil microbial abundance and activity—A meta-analysis and meta-regression. *PLOS ONE*, *12*(7), e0180442. https://doi.org/10.1371/journal.pone.0180442

Maas, E. V., & Grattan, S. R. (1999). Crop yields as affected by salinity. In R. W. Skaggs & J. van Schilfgaarde (Eds.), *Agricultural Drainage* (pp. 55–108). ASA, CSSA, and SSSA.

Maltais-Landry, G., Scow, K., Brennan, E., & Vitousek, P. (2015). Long-term effects of compost and cover crops on soil phosphorus in two California agroecosystems. *Soil Science Society of America Journal*, *79*(2), 688–697. https://doi.org/10.2136/sssaj2014.09.0369

Miller, J. J., Beasley, B. W., Larney, F. J., & Olson, B. M. (2005). Soil salinity and sodicity after application of fresh and composted manure with straw or wood-chips. *Canadian Journal of Soil Science*, https://doi.org/10.4141/S04-066

Miller, K., Aegerter, B. J., Clark, N. E., Leinfelder-Miles, M., Miyao, E. M., Smith, R., Wilson, R., & Geisseler, D. (2019). Relationship between soil properties and nitrogen mineralization in undisturbed soil cores from California agroecosystems. *Communications in Soil Science and Plant Analysis*, *50*(1), 77–92. https://doi.org/10.1080/00103624.2018.1554668

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

Mulvaney, R. L. (1996). Nitrogen-inorganic forms. In D. L. Sparks, A. L. Page, P. A. Helmke, R. H. Loeppert, P. N. Soltanpour, M. A. Tabatabai, C. T. Johnston, & M. E. Sumner (Eds.), *Methods of soil analysis: Part 3. Chemical methods* (2nd ed., pp. 1123–1184). ASA and SSSA.

Nelson, D. W., & Sommers, L. E.(1996). Total carbon, organic carbon, and organic matter. In D. L. Sparks, A. L. Page, P. A. Helmke, R. H. Loeppert, P. N. Soltanpour, M. A. Tabatabai, C. T. Johnston, & M. E. Sumner (Eds.), *Methods of soil analysis: Part 3. Chemical methods* (2nd ed., pp. 971–1010). ASA and SSSA.

Norris, C. E., Bean, G. M., Cappellazzi, S. B., Cope, M., Greub, K. L. H., Liptzin, D., Rieke, E. L., Tracy, P. W., Morgan, C. L. S., & Honeycutt, C. W. (2020). Introducing the North American project to evaluate soil health measurements. *Agronomy Journal*, *112*(4), 3195–3215. https://doi.org/10.1002/agj2.20234

Prosser, J. A., Speir, T. W., & Stott, D. E. (2011). Soil oxidoreductases and FDA hydrolysis. In R. P. Dick (Ed.), *Methods of Soil Enzymology* (pp. 103–124). SSSA. https://doi.org/10.2136/sssabookser9.c6

Roper, W. R., Osmond, D. L., & Heitman, J. L. (2019). A response to "Reanalysis validates soil health indicator sensitivity and correlation with long-term crop yields." *Soil Science Society of America Journal*, *83*(6), 1842–1845. https://doi.org/10.2136/sssaj2019.06.0198

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

Schaalje, G. B., Mcbride, J. B., & Fellingham, G. W. (2002). Adequacy of approximations to distributions of test statistics in complex mixed linear models. *Journal of Agricultural, Biological, and Environmental Statistics*, *7*(4), 512. https://doi.org/10.1198/108571102726

Schmidt, J. E., Peterson, C., Wang, D., Scow, K. M., & Gaudin, A. C. M. (2018). Agroecosystem tradeoffs associated with conversion to subsurface drip irrigation in organic systems. *Agricultural Water Management*, *202*, 1–8. https://doi.org/10.1016/j.agwat.2018.02.005

Six, J., Conant, R. T., Paul, E. A., & Paustian, K. (2002). Stabilization mechanisms of soil organic matter: Implications for C-saturation of soils. *Plant and Soil*, *241*(2), 155–176. https://doi.org/10.1023/A:1016125726789

Stewart, R. D., Jian, J., Gyawali, A. J., Thomason, W. E., Badgley, B. D., Reiter, M. S., & Strickland, M. S. (2018). What we talk about when we talk about soil health. *Agricultural & Environmental Letters*, *3*(1), 180033. https://doi.org/10.2134/ael2018.06.0033

Stott, D. E. (2019). *Recommended soil health indicators and associated laboratory procedures*. *Soil Health Technical Note No. 450-03*. USDA–NRCS.

Suarez, D. L. (1996). Beryllium, magnesium, calcium, strontium, and barium. In D. L. Sparks, A. L. Page, P. A. Helmke, R. H. Loeppert, P. N. Soltanpour, M. A. Tabatabai, C. T. Johnston, & M. E. Sumner (Eds.), *Methods of soil analysis: Part 3. Chemical methods* (2nd ed., pp. 575–601). ASA and SSSA.

Tao, J., Kiernan, K., & Gibbs, P. (2015). Advanced techniques for fitting mixed models using SAS/STAT® software. Paper SAS1919-2015.

Tautges, N. E., Chiartas, J. L., Gaudin, A. C. M., O'Geen, A. T., Herrera, I., & Scow, K. M. (2019). Deep soil inventories reveal that impacts of cover crops and compost on soil carbon sequestration differ in surface and subsurface soils. *Global Change Biology*, *25*(11), 3753–3766. https://doi.org/10.1111/gcb.14762

Thomas, G. W. (1996). Soil pH and soil acidity. In D. L. Sparks, A. L. Page, P. A. Helmke, R. H. Loeppert, P. N. Soltanpour, M. A. Tabatabai, C. T. Johnston, & M. E. Sumner (Eds.), *Methods of soil analysis: Part 3. Chemical methods* (2nd ed., pp. 475–490). ASA and SSSA.

Verdouw, H., Van Echteld, C. J. A., & Dekkers, E. M. J. (1978). Ammonia determination based on indophenol formation with sodium salicylate. *Water Research*, *12*(6), 399–402. https://doi.org/10.1016/0043-1354(78)90107-0

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., Horwath, W. R., & Burger, M. B. (2016). Integrating soil biological and chemical indices to predict net nitrogen mineralization across California agricultural systems. *Soil Science Society of America Journal*, *80*(6), 1675–1687. https://doi.org/10.2136/sssaj2016.07.0228

Wade, J., Maltais-Landry, G., Lucas, D. E., Bongiorno, G., Bowles, T. M., Calderón, F. J., Culman, S. W., Daughtridge, R., Ernakovich, J. G., Fonte, S. J., Giang, D., Herman, B. L., Guan, L., Jastrow, J. D., Loh, B. H. H., Kelly, C., Mann, M. E., Matamala, R., Miernicki, E. A., … Margenot, A. J. (2020). Assessing the sensitivity and repeatability of permanganate oxidizable carbon as a soil health metric: An interlab comparison across soils. *Geoderma*, *366*, 114235. https://doi.org/10.1016/j.geoderma.2020.114235

Wander, M. M., Cihacek, L. J., Coyne, M., Drijber, R. A., Grossman, J. M., Gutknecht, J. L. M., 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

Watanabe, F. S., & Olsen, S. R. (1965). Test of an ascorbic acid method for determining phosphorus in water and $NaHCO_3$ extracts from soil. *Soil Science Society of America Journal*, *29*, 677–678. https://doi.org/10.2136/sssaj1965.03615995002900060025x

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*(1), 3–17. https://doi.org/10.1079/AJAA200228

Wuest, S. (2014). Seasonal variation in soil organic carbon. *Soil Science Society of America Journal*, *78*(4), 1442–1447. https://doi.org/10.2136/sssaj2013.10.0447

Zuber, S. M., Veum, K. S., Myers, R. L., Kitchen, N. R., & Anderson, S. H. (2020). Role of inherent soil characteristics in assessing soil health across Missouri. *Agricultural & Environmental Letters*, *5*(1), e20021. https://doi.org/10.1002/ael2.20021

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: Lazicki P, Mazza Rodrigues JL, & Geisseler, D. Sensitivity and variability of soil health indicators in a California cropping system. *Soil Sci Soc Am J*. 2021;1–16.
https://doi.org/10.1002/saj2.20278