SOIL & WATER MANAGEMENT & CONSERVATION

The soil health assessment protocol and evaluation applied to soil organic carbon

Márcio R. Nunes[1] | **Kristen S. Veum**[2] | **Paul A. Parker**[3] | **Scott H. Holan**[3] | **Douglas L. Karlen**[4] | **Joseph P. Amsili**[5] | **Harold M. van Es**[5] | **Skye A. Wills**[6] | **Cathy A. Seybold**[6] | **Thomas B. Moorman**[1]

[1] USDA-ARS, National Lab. for Agriculture and the Environment, 1015 N. Univ. Boulevard, Ames, IA 50011-3611, USA

[2] USDA-ARS, Cropping Systems and Water Quality Research Unit, 269 Agricultural Engineering Bldg., Univ. of Missouri, Columbia, MO 65211, USA

[3] Dep. of Statistics, Univ. of Missouri, 134 Middlebush Hall, Columbia, MO 65211, USA

[4] USDA-ARS (retired), DLKarlen Consulting, 487 Ashland Ave., St. Paul, MN 55102, USA

[5] Cornell Univ., School of Integrative Plant Science, 233 Emerson Hall, Ithaca, NY 14853, USA

[6] USDA-NRCS, National Soil Survey Center, 100 Centennial Mall North, Room 152, Lincoln, NE 68508, USA

Correspondence
Márcio R. Nunes, USDA-ARS, National Lab. for Agriculture and the Environment, 1015 N. Univ. Boulevard., Ames, IA 50011-3611, USA
Email: marcio_r_nunes@alumni.usp.br

Assigned to Associate Editor DeAnn Presley.

Abstract

The concept of soil health has evolved over the past several decades, recognizing that dynamic soil property response to management and land use is highly dependent on site-specific factors that must be considered when interpreting soil health measurements. Initially, the Soil Management Assessment Framework (SMAF) and Comprehensive Assessment of Soil Health (CASH) were developed and used globally for scoring soil health indicators. However, both SMAF and CASH frameworks were developed using a relatively small dataset and their interpretation curves were not validated at the nationwide scale. Expanding upon these concepts, we propose the Soil Health Assessment Protocol and Evaluation (SHAPE) tool. The SHAPE was developed using 14,680 soil organic C (SOC) observations from across the United States, and accounts for edaphic and climate factors at the continental scale. Data were compiled from the literature, the Cornell Soil Health Laboratory, and the Kellogg Soil Survey Laboratory. In this approach, scoring curves are Bayesian model-based estimates of the conditional cumulative distribution function (CDF) for defined soil peer groups reflecting five soil texture and five soil suborder classes adjusted for mean annual temperature and precipitation. Specifically, SHAPE produces scores between 0 and 1 (0–100%) for measured SOC values that reflect the quantile or position within the conditional CDF along with measures of uncertainty. Herein, we focus on development of the SHAPE scoring curve for SOC with our case studies. SHAPE is a flexible, quantitative tool that provides a regionally relevant interpretation of this key soil health indicator.

1 | INTRODUCTION

Soil health refers to the ability of a specific soil to perform multiple functions that include cycling of nutrients, sustaining productivity, maintaining biodiversity, regulating water dynamics, and moderating climate (Karlen et al., 1997). Current interest in soil health is mostly driven by the recognition

Abbreviations: AI, de Martone aridity index; AS, aggregate stability; CASH, Comprehensive Assessment of Soil Health; CDF, cumulative distribution function; LOI, loss on ignition; MAP, mean annual precipitation; MAT, mean annual temperature; PET, potential evapotranspiration; SHAPE, Soil Health Assessment Protocol and Evaluation; SOC, soil organic carbon; SOM, soil organic matter content; SMAF, Soil Management Assessment Framework; WI, wetness index.

that anthropogenic influences are increasingly superimposed on natural soil quality. Ecosystem services are the direct and indirect contributions of ecosystems to human well-being and are inherently difficult to measure directly due to high cost and spatial and temporal variability; however, a wide range of soil health indicators, including chemical, physical, and biological soil properties (e.g., Doran and Parkin, 1996; Stott, 2019) have been identified as proxy measurements. Among them, soil organic C (SOC) is a keystone indicator, reflecting multiple soil functions and ecosystem services. Although it is widely recognized that climatic and edaphic factors are determinants of soil C content (Bardgett, 2011; Post et al., 2004), land use and management practices have a significant impact on soil C dynamics.

The interactions among inherent and dynamic soil biological, physical, and chemical properties and processes are complex and must be quantified when assessing management effects on soil health. To facilitate such quantifications across land use and management practices, an interpretive framework that provides a wide range of regionally relevant indicator options is needed (Wander et al, 2019), such as a multi-indicator soil health index. This index must also (a) account for inherent site-specific factors, (b) be sensitive to anthropogenic activities, and (c) facilitate broad-scale monitoring to ensure sustainable land management (Doran, 2002).

Many indices have been developed for soil health assessment (e.g., Andrews et al., 2004; De Paul Obade & Lal, 2016; Haney et al., 2018; Idowu et al., 2009; Rinot et al., 2019), but none are universally accepted (Bünemann et al., 2018). One example is the Soil Management Assessment Framework (SMAF), which integrates multiple biological, chemical, and physical soil health indicators by transforming measured values into unitless (0–1) scores (Andrews et al., 2004). The SMAF algorithms were built to be sensitive to several site-specific factors (i.e., soil taxonomy, climate, crop type, topography, sampling season) that influence a soil's potential to support critical functions. The SMAF has been used globally, although the conceptual framework for indicator interpretation was derived using a relatively small dataset representing four U.S. states: Georgia, California, Wisconsin, and Iowa. The Comprehensive Assessment of Soil Health (CASH) is another tool that scores multiple soil health measurements using cumulative normal distributions of a regional dataset from the northeastern United States (Idowu et al., 2009). Recently, CASH algorithms were expanded to include soils from other U.S. regions, including the Midwest and Mid-Atlantic (Fine et al., 2017), but the spatial extent is still limited and the curves do not reflect the national or global distribution of soils.

To further advance science-based soil health assessments, it is essential to develop indexing tools that handle multiple soil attributes and describe relevant soil health status for multiple types of soil (Bünemann et al., 2018). Develop-

> **Core Ideas**
>
> - Response of soil health (SH) indicators to land use and management is site-specific.
> - Soil Health Assessment Protocol and Evaluation (SHAPE) is proposed as a flexible tool.
> - The SHAPE builds upon conceptual frameworks established by the SMAF and CASH protocols.
> - The SHAPE provides SH interpretation for soil peer groups based on edaphic and climatic factors.
> - This tool provides knowledge about the status of soils in response to agronomic practices.

ment of a new assessment framework can be based on statistical approaches (e.g., minimum dataset, principal component analysis, decision trees, and ANOVA), expert opinion, or expert-based frameworks (Andrews et al., 2004; Fine et al., 2017; Ritz et al., 2009). To be useful and interpretable for producers and landowners, the selected soil health indicators must: (a) represent chemical, biological, and physical soil processes; (b) detect variations in soil functions due to management and land-use decisions; (c) be assessable and cost effective; and (d) reflect the connection between soil functions and management targets (i.e., agricultural productivity, ecosystem services).

The SMAF has scoring curves for four biological (SOC, microbial biomass C, potentially mineralizable N, and beta-glucosidase [β-G]); four physical (bulk density, aggregate stability [AS], available water capacity, and water-filled pore space); and five chemical soil properties (pH, electrical conductivity, sodium adsorption ratio, and extractable P and K) (Andrews et al., 2004). The CASH was designed for broader use with lower cost and includes physical (AS, available water capacity, and penetration resistance); biological (organic matter content [SOM], autoclaved citrate-extractable protein, respiration, and active-C); and chemical (pH and nutrient concentration) soil properties (Moebius-Clune et al., 2016). Within CASH and SMAF, biological indicators reflect the amount and quality of SOM, the size and activity of the microbial community, nutrient cycling and storage, soil structure, and provision of food for edaphic organisms, among other soil functions and processes (Lal, 2016; McDaniel et al., 2014; Tiemann et al., 2015). Soil physical properties relate to plant growth, water dynamics, erosion resistance, nutrient retention, rhizosphere processes, resilience to drought or temperature fluctuations, mitigation of nonpoint source pollutants, and agronomic productivity (Lal, 2016; Nunes et al., 2019). The soil chemical indicators reflect acidity and nutrient availability as a guide to soil fertility management and environmental protection.

Following selection, measurement, and interpretation of measured soil health indicators, the unitless values can be used to develop one or more indices (Bünemann et al., 2018) depending on what the assessment is being used for. Typically, the values represent the status of a specific indicator relative to (a) threshold or reference values, (b) one or more specific soil functions, or (c) a dataset of measured soil health indicator values. These relationships are established using nonlinear curves that can be defined and calibrated in several ways. For example, SMAF thresholds were primarily based either on expert opinion or taken from published literature (Andrews et al., 2004), whereas CASH thresholds were empirically derived based on data distributions for each soil health indicator within a regional dataset (Fine et al., 2017). In contrast to the SMAF and CASH approach, some indices (e.g., the Haney test; Haney et al., 2018) use raw measured values or projections from a model-specific outcome, such as biomass production (e.g., De Paul Obade and Lal, 2016) or relative crop yield (Biswas et al., 2017; Lopes et al., 2013) to establish threshold values.

Rinot et al. (2019) stressed that linking soil health indicator values with ecosystem services is a laudable goal that would supply rich and meaningful information from an ecological perspective, but quantitative data to assess those complex relationships is lacking at several spatial and temporal scales. For example, water quality or greenhouse gas emission data are costly and difficult to obtain at scales needed to develop interpretations for those important soil health functions. Also, indicator scores and interpretations can be inconsistent or easily manipulated when based on expert opinion or inappropriate statistical methods (Bünemann et al., 2018). Furthermore, optimal threshold values will often vary or present trade-offs among soil functions or ecosystem services (i.e., crop productivity vs. environmental protection) and are dependent on scale (Simpson, 2016). This can cause significant factors to be overlooked at the continental scale if indices are only developed using regional or site-specific datasets.

Karlen et al. (2019) documented how the soil health concept evolved over the past several decades and emphasized the need to scientifically advance monitoring and assessment protocols by (a) improving indicator scoring tools, (b) developing national monitoring protocols, and (c) identifying new soil biological, chemical, and physical indicators of soil health. They also stressed that producer interest in soil health monitoring and regionally relevant interpretation remains strong. To meet those needs, the USDA–NRCS and USDA–ARS initiated a meta-analysis project focused on indicator interpretation and tool development.

A national database was compiled using soil health data from 456 published studies. Meta-analysis techniques were applied to assess effects of anthropogenic activities on soil health indicators and examine potential interactions among management practices and inherent soil conditions (i.e., soil type, climate condition, soil texture, soil depth, cropping system, duration of the experiment among other factors [Nunes, Karlen, & Moorman, 2020; Nunes, Karlen, Moorman, & Cambardella, 2020; Nunes, Karlen, Veum, Moorman, & Cambardella, 2020]). Results confirmed that the soil health indicators used in SMAF and CASH were sensitive to changes induced by anthropogenic factors, but anomalies emerged at the continental scale, emphasizing the need to review and improve indicator scoring curves (Karlen et al., 2019). The comprehensive database thus provided the foundation for our first step toward development of an improved soil health evaluation (i.e., Soil Health Assessment Protocol and Evaluation [SHAPE]).

The soil genesis concepts outlined by Jenny (1941) illustrate the five dominant soil forming factors (i.e., time, parent material, vegetation, climate, and topography) that interact to determine inherent conditions. These factors influence soil properties and processes and are incorporated into the soil taxonomy (Soil Survey Staff, 1999). For example, in addition to dynamic factors, the SOC content is determined by climate variables, such as moisture and temperature, that are key factors controlling plant growth, decomposition of organic materials, and ultimately SOC content (Jobbagy & Jackson, 2000; Post et al., 2004; Trumbore, 1997). Thus, as with SMAF and CASH, SHAPE development was motivated by the knowledge that the response of dynamic soil properties to management was highly dependent on inherent, site-specific factors (Karlen et al., 1997). Those conclusions were also consistent with recommendations from other meta-analyses (i.e., Baker et al., 2007; Bowles et al., 2016; Graaff et al., 2019; Haddaway et al., 2017; Luo et al., 2010; McDaniel et al., 2014).

Although SMAF and CASH were designed to account for inherent site characteristics, both used relatively small and geographically limited datasets. Furthermore, multiple assessments identified potential scoring curve problems at the regional scale, including both overestimation and underestimation of SOC, β-G, and AS scores (Mbuthia et al., 2015; Nunes, Karlen, Veum, & Moorman, 2020; Stott et al., 2013; Veum et al., 2015; Zobeck et al., 2015) that likely reflected dataset limitations and a need to improve how edaphic and climate information is used to modify the scoring curves. Thus, compared with SMAF and CASH, SHAPE was based on a substantially larger and more representative dataset.

Our ultimate goal continues to be the development of a better, comprehensive soil health index that accounts for continental variation in climate and inherent soil conditions, while remaining sensitive to field-scale land use and management practices. Our broad objectives were to: (a) reevaluate the inherent factors and classes used in SMAF and CASH to convert measured values into scores; (b) develop new scoring curves for soil health indictors including SOC, β-G, AS, autoclaved citrate-extractable protein, active-C, and respiration, among others, based on the peer group approach
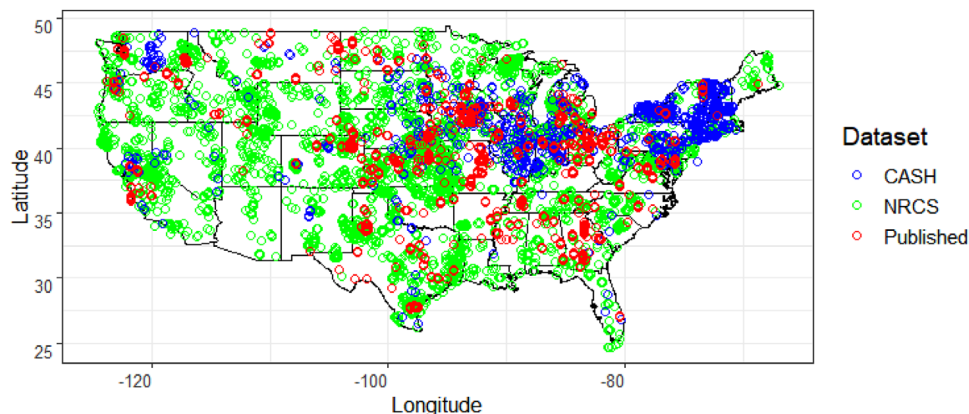
FIGURE 1 Location of sites where soil organic C data for different agricultural production systems were compiled and compared within the conterminous United States for the Soil Health Assessment Protocol Evaluation from the Comprehensive Assessment of Soil Health (CASH), the NRCS Kellogg Soil Survey Lab (NRCS), and the published datasets

(a unique combination of soil suborder group and texture class adjusted for temperature and precipitation); and (c) create a new soil health tool that provides a quantitative and interpretive score with associated measures of uncertainty. Specifically, this manuscript discusses the initial development of SHAPE, focusing on SOC, as the first in a series of indicator scoring curves.

2 | MATERIALS AND METHODS

2.1 | Dataset

A dataset consisting of 14,680 SOC observations from across the conterminous United States was compiled from three sources: (a) published data for studies throughout the United States, (b) CASH data samples analyzed by the Cornell Soil Health Testing Laboratory between 2014 and 2018, and (c) NRCS data from samples analyzed by the NRCS Kellogg Soil Survey Laboratory in Lincoln, NE. A description of each data source is provided below.

2.1.1 | Published data

This portion of the dataset was previously described by Nunes, Karlen, Veum, Moorman and Cambardella (2020). Briefly, the data were compiled using a literature search with keywords (a) soil health or soil quality, plus (b) cropping system, soil tillage, residue management, cover crop, crop rotation, soil fertility or fertilizer. For inclusion in the dataset, publications had to (a) present soil health indicators from perennial ecosystems or studies comparing multiple treatments such as tillage intensity or cropping system diversification, (b) be controlled (i.e., replicated) studies, (c) be written in English, and (d) be conducted in the United States. Duplications, unpub-

lished studies, non–peer-reviewed papers, and studies presenting results only in graphs were excluded. Concentration values (i.e., g kg^{-1}) rather than stocks (i.e., volumetric or areal measurements such as Mg ha^{-1} or kg m^{-2} over fixed depth or soil mass) were used unless BD values were reported that could be used to convert stock values into concentration units. Soil organic C values consisted of total organic C or total C (where soils were shown to be free of carbonates). For data presented as SOM, values were divided by 1.724 (van Bemmelen factor) to convert them to SOC (Cambardella et al., 2001). All values were from the topsoil layer ($\leq$15-cm). In total, 4,129 SOC observations (456 articles) measured across the United States (Figure 1) under several land use and agricultural management practices (i.e., tillage intensity, cropping system, residue management, and addition of amendments [Supplemental Table S1]) were included in the dataset.

2.1.2 | Comprehensive Assessment of Soil Health data

The CASH dataset considered all samples that contained GPS information submitted for analysis to the Cornell Soil Health Laboratory (Ithaca, NY) from 2014 to 2018 (n = 4,183). Based on CASH sampling guidelines, it was assumed each data point represented a composite sample collected from the 0-to-15-cm depth (Moebius-Clune et al., 2016). Only SOC and soil texture data were used for this study. Based on CASH protocols, SOM content had been determined by measuring loss on ignition (LOI) after 2 hr at 500 °C in a muffle furnace. The %LOI was converted to %OM using the equation (%OM = [% LOI × 0.7] − 0.23). Once again, we converted SOM to SOC using the van Bemmelen factor (Cambardella et al., 2001). In contrast with the published data, which comprises primarily plot-scale experimental results, most CASH data originated from commercial farms
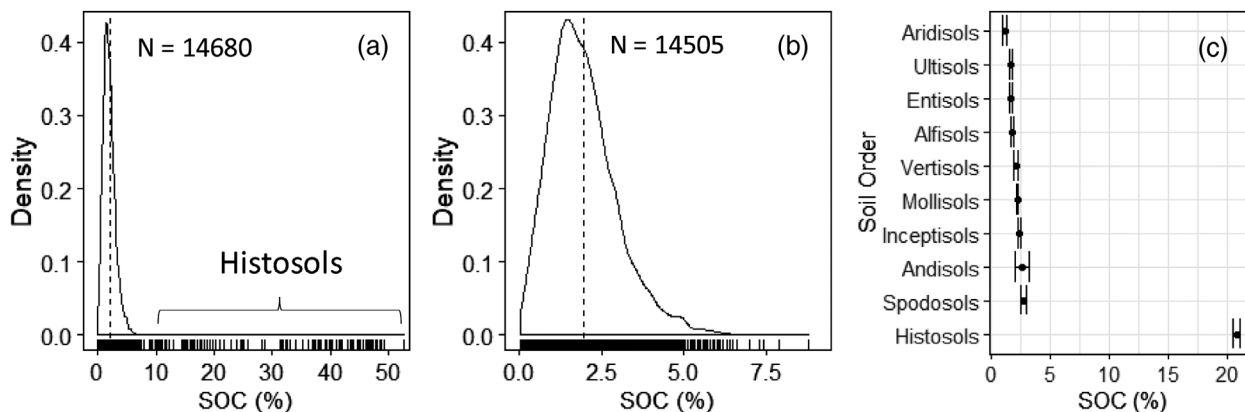
FIGURE 2 Soil organic C (SOC) data distribution with (a) and without (b) Histosols, and mean SOC for each soil order (c) with 0.95% confidence intervals (horizontal bars) across the United States

and represent several tillage and agroecosystems across the United States (Supplemental Table S1).

2.1.3 | Natural Resources Conservation Service data

This dataset was derived from samples submitted for soil characterization to the NRCS Kellogg Soil Survey Laboratory in Lincoln, NE ($n = 6,368$) and contain GPS location information. Natural Resources Conservation Service records were queried for surface and mineral A horizons using the National Cooperative Soil Survey Characterization Database. This database contains over 38,000 pedons with measured chemical and physical properties representing geographically diverse soils from across the conterminous United States, Hawaii, and Alaska. Those samples were generally collected by NRCS soil scientists and cooperating universities, usually for soil survey activities and landscape characterization. Land use and management were not known for all locations, but generally reflected the most common scenarios where the samples were collected. Only SOC and soil texture data were used for this study. The SOC content was determined by difference between the total C (dry combustion) and $CaCO_3$–C (electronic manometer method), and particle size is determined by the pipette method and sieving (Soil Survey Staff, 2014). Soil taxonomic classifications (i.e., soil suborder) from the soil profile descriptions were also used and are defined in Soil Survey Staff (1999).

2.1.4 | Combined SOC dataset

Data from the three sources (Figure 1) were combined to provide 14,680 SOC values ranging from 1 to 530 g kg^{-1} (0.1–53%) from across the United States (Figure 2a). Compared with the original SMAF (Andrews et al., 2004) and CASH

(Fine et al., 2017) development protocols, scoring curves derived from this comprehensive dataset provided a larger, more geographically diverse representation of SOC values from across the continental United States (Figure 1). Furthermore, SOC observations within the dataset also represented a larger range of land uses (i.e., native grassland, forest soils, and agricultural production) and practices reflecting variable tillage intensity, crop residue management, and addition of amendments.

2.2 | Inherent soil and climate variables

Soil and climate variables with potential to affect topsoil SOC were linked to each SOC observation in the dataset (Table 1). Soil variables included order, suborder, texture, and drainage class. Overall, the dataset represented 10 taxonomic orders, 57 suborders, 12 soil texture classes, and eight soil drainage classes (Table 1; Supplemental Table S2). Climate variables included mean annual precipitation (MAP), mean annual temperature (MAT), potential evapotranspiration (PET), de Martone aridity index (AI; Equation 1), and the wetness index (WI; Equation 2).

$$AI = \frac{MAP}{(MAT + 10)} \tag{1}$$

$$WI = \frac{MAP}{PAT} \tag{2}$$

The combined dataset represented a wide range of MAT (from −5.6 to 25.3 °C), MAP (from 42 to 3,671 mm), PET (from 845 to 2,539 mm yr^{-1}), AI (from 1.64 to 305.5), and WI (from 0.03 to 4.16) (Table 1). Several additional rainfall and vegetation variables (e.g., growing degree days) were considered but not selected for evaluation.

TABLE 1 Description of the inherent covariates evaluated for use in the Soil Health Assessment Protocol Evaluation SOC scoring curves

Inherent covariates	Unit	Source	Type	Range	N
Climate					–
Mean annual precipitation	mm	USGS	quantitative	42–3,671	–
Mean annual temperature	°C	USGS	quantitative	−5.6–25.3	–
Potential evapotranspiration	mm yr^{-1}	UMNTSG	quantitative	845–2,539	–
Wetness index	–	calculated	quantitative	0.03–4.16	–
de Martone aridity index	–	calculated	quantitative	1.64–305.5	–
Soil					
Order	–	gSSURGO	categorical	–	10
Suborder	–	gSSURGO	categorical	–	57
Texture class	–	gSSURGO	categorical	–	12
Drainage class	–	gSSURGO	categorical	–	8

Note. USGS, United States Geological Survey; UMNTSG, University of Montana Numerical Terradynamic Simulation Group; gSSURGO, Gridded Soil Survey Geographic Database.

The site-specific variables were obtained using the location (latitude and longitude) from a complement of environmental coverages. Soil variables were assigned using Gridded Soil Survey Geographic Database. Climate variables were from USGS or were supplied directly from articles (specifically for the published dataset) or by calculation (AI and WI). Soil survey information was extracted for the dominant map unit at each given location. In summary, the final database (all three datasets combined) was composed of nine inherent variables (five quantitative and four categorical) with SOC as the target response variable (Table 1), totaling 14,680 topsoil SOC observations from across the continental United States (Figure 1). The number of observations by soil order, suborder, texture, and drainage classes is presented in Supplemental Table S2.

2.3 | Data analyses

All statistical analyses were performed using R software version (R Core Team, 2020). Multiple exploratory analyses were conducted to assess distributional properties of the dataset, including the number of SOC observations within each level (class) of categorical variables (Table 1; Supplemental Table S2), and for each land use and agricultural management practice in the published and CASH datasets (Supplemental Table S1). Linear regression, correlation analysis, exploratory data analysis, and subsequent modeling as described below were pursued.

2.3.1 | Evaluation of edaphic and climate variables

Sensitivity of SOC to inherent factors used in SMAF (five texture classes, four suborder classes, and four climate classes,

as presented in Table 2 [Andrews et al., 2004]) and CASH (three soil texture classes as presented in Supplemental Table S3 [Fine et al., 2017]) were initially evaluated by fitting various linear regression models. Mean SOC values for each class were plotted with the 95% confidence interval assuming normality.

The first linear model (Equation 1) included two categorical soil variables (texture [Table 2], suborder [Table 3]) and five continuous climate variables [Table 1]). Soil suborder class S1 was predominantly represented by the Histosols, which form in organic rather than mineral materials and have unique characteristics including SOC concentrations exceeding 95 g kg^{-1} (9.5%). Since there were only 175 S1 samples out of 14,680 observations, they were removed from the dataset for a separate modeling effort. Thus, for mineral soils the final dataset contained 14,505 observations.

$$\begin{aligned} SOC_i = {} & \beta_0 + MAT_i\beta_1 + MAP_i\beta_2 + PET_i\beta_3 + AI_i\beta_4 \\ & + WI_i\beta_5 + \sum_{t=1}^{T} I\left(\text{Texture}_i = t\right)\beta_{6t} \\ & + \sum_{s=1}^{S} I\left(\text{Suborder}_i = s\right)\beta_{7s} + \varepsilon_i \end{aligned}$$

Equation 1 contained several correlated climate variables that likely represented similar variation in SOC. Therefore, we explored multiple techniques to select the best climate variables for scoring tool development. Initial variable selection for inclusion in the model was done using a combination of exploratory data analysis. A best subset regression with the R package "leaps" (Lumly, 2020) was used to identify the best climate prediction variables for which only five were considered as SOC predictors within the full model. The "leaps" approach considers all possible subsets of covariates before selecting the model that minimizes some criteria, for which we considered adjusted R^2 and Schwarz information criterion.

TABLE 2 Inherent soil and climate factor groups used to determine the soil organic C (SOC) indicator index value, as originally presented in the Soil Management Assessment Framework (Andrews et al., 2004)

Class	Soil texture
Class T1	sand, loamy sand, sandy loam (with <8% clay)
Class T2	sandy loam (with clay >8%), sandy clay loam, loam
Class T3	silt loam, silt
Class T4	sandy clay, clay loam, silty clay loam, silty clay, clay (<60%)
Class T5	clay (>60%)
Suborder class	soil taxonomy suborder
Class S1	Aquands, Histels, Turbels, Fibrists, Folists, Hemists, Saprists, Aquoxs, Aquods
Class S2	Udands, Ustands, Aquepts, Albolls, Aquolls, Borolls, Cryolls, Rendolls, Udolls, Ustolls, Xerolls, Udoxs, Humods, Aquults, Humults, Aquerts, Xererts
Class S3	Aqualfs, Boralfs, Cryalfs, Udalfs, Ustalfs, Xeralfs, Cryands, Vitrands, Aquents, Orthels, Andepts, Anthrepts, Cryods, Cryerts
Class S4	Torrands, Xerands, Argids, Calcids, Cambids, Cryids, Durids, Gypsids, Orthids, Salids, Arents, Fluvents, Orthents, Psamments, Xerents, Cryepts, Ochrepts, Tropepts, Udepts, Umbrepts, Ustepts, Xerepts, Orthoxs, Peroxs, Torroxs, Ustoxs, Orthods, Udults, Ustults, Xerults, Torrerts, Usterts, Uderts

Climate class	Degree days		Precipitation
	˚F (˚C)		mm
Class C1	≥170 (76.7)		≥550
Class C2	≥170 (76.7)		<550
Class C3	<170 (76.7)		≥550
Class C4	<170 (76.7)		<550

Note. Soil texture class 1 represents the texture groups that have the lowest intrinsic potential for sequestering SOC, while class 5 has the highest (Quisenberry et al., 1993). Soil suborder class 1 represents the suborders that are expected to have the highest potential for sequestering SOC, while class 4 has the lowest. Climate class 1 represents the climate class that has the lowest intrinsic SOC, while class 4 has the highest.

TABLE 3 New soil taxonomy suborder classes selected for the Soil Health Assessment Protocol Evaluation soil organic C (SOC) curve

Class	Soil taxonomy suborder
S1	Fribists, Folists, Hemists, Histels, Saprists, Wassists
S2	Aquands, Aquents, Aquepts, Aquods, Aquoxs, Cryods, Humods, Orthels, Peroxs, Torrands, Tropepts, Turbels, Udands, Udoxs, Ustands
S3	Albolls, Andepts, Aquolls, Aquults, Cryands, Cryepts, Cryolls, Gelepts, Gelolls, Humults, Rendolls, Umbrepts, Ustoxs, Vitrands, Wassents, Xerands
S4	Aqualfs, Aquerts, Boralfs, Borolls, Cryalfs, Ochrepts, Orthods, Orthoxs, Udalfs, Udepts, Uderts, Udolls, Usterts, Ustolls, Xeralfs, Xerepts, Xerolls, Xerults
S5	Arents, Argids, Calcids, Cambids, Cryerts, Cryids, Durids, Fluvents, Gypsids, Orthents, Orthids, Psamments, Salids, Torrerts, Torroxs, Udults, Ustalfs, Ustepts, Ustults, Xererts

Note. Class S1 represents suborders that are expected to have the highest potential for sequestering SOC, while Class S5 has the lowest.

We also used random forest modeling and stepwise regression to identify variable importance. Both techniques yielded similar results. The final selection of variables was incorporated into SHAPE, where goodness-of-fit is more formally evaluated. Summarily, exploratory data analysis in conjunction with subject matter expertise was used only for initial stages of model development.

3 | RESULTS AND DISCUSSION

3.1 | Response of SOC to SMAF climate and soil suborder classes

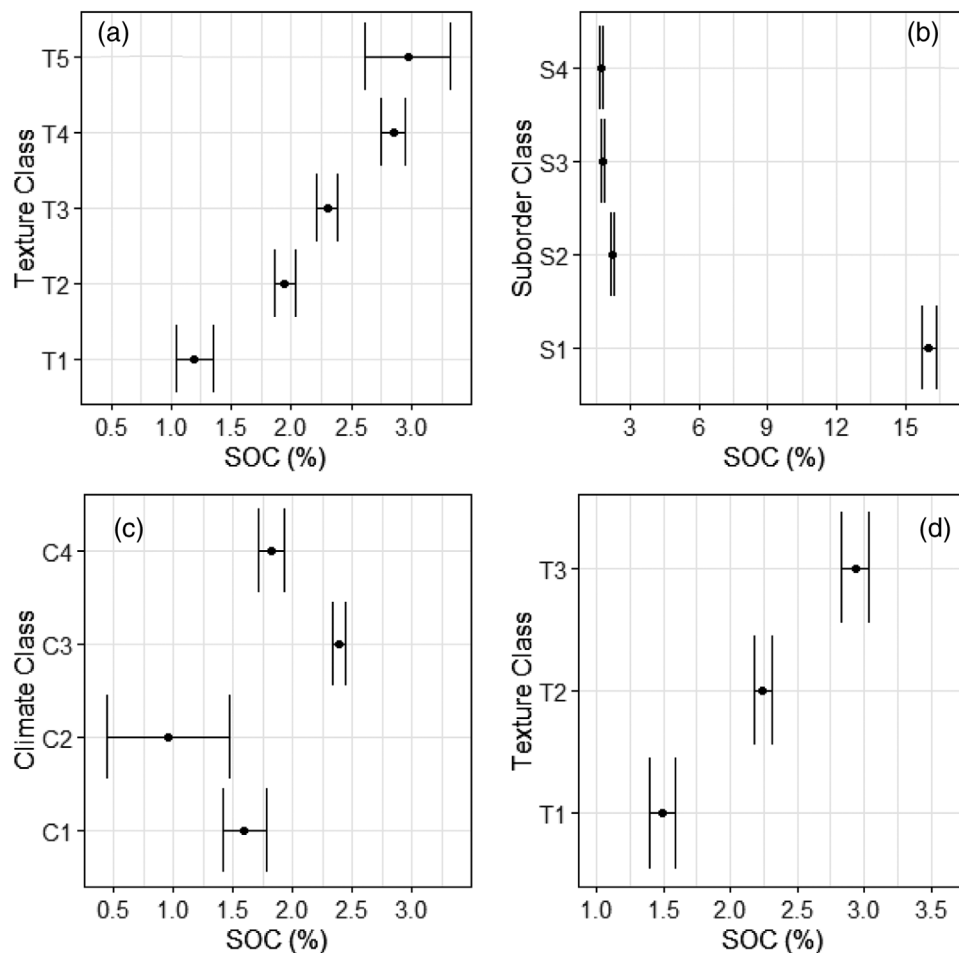The SMAF–SOC scoring curve approach assumed inherent SOC content increased from coarse (T1) to fine (T5)

FIGURE 3 Soil organic C (SOC) as a function of five soil textural classes (a), four soil suborders (b), and four climate classes (c) as used in the Soil Management Assessment Framework (Table 2), and three soil texture classes as used in Comprehensive Assessment of Soil Health ([d]; Supplemental Table S1). Horizontal bars represent the 95% confidence interval (based on normality)

textural classes. Similarly, CASH assumed SOM increases from texture class T1 to T3 (Supplemental Table S3). Those overall trends were confirmed in the new, combined dataset (Figure 3a,d); thus, the five original SMAF textural classes were used for SHAPE–SOC scoring curve development. Similarly, the original SMAF soil taxonomy suborder factor was used, which consisted of four classes (Table 2) based on expert opinion regarding the potential for an individual suborder to sequester C (Andrews et al., 2004). The SMAF procedures presumed that inherent SOC concentration increased from soil suborder class S4 to S1. This assumption was partially supported by the combined dataset, which showed SOC values followed the order S4 = S3 < S2 < S1 (Figure 3b). This likely contributed to the underestimation or overestimation of SOC scores observed by several SMAF assessments (e.g., Mbuthia et al., 2015; Nunes, Karlen, Veum, & Moorman, 2020; Stott et al., 2013; Zobeck et al., 2015). To improve SHAPE, suborder

groupings were reassigned into five classes (Table 3). They now show inherent topsoil SOC values increasing from S5 to S1, as confirmed by the comprehensive dataset (Figure 4).

For climate, SMAF combined MAT and MAP into four broad classes (Table 2) that were parameterized assuming that as temperature and precipitation increased, SOC would decrease due to greater decomposition rates (Andrews et al., 2004). Therefore, it was expected that SOC content would increase from climate class C1 to C4, but again this expectation was not confirmed by the combined dataset (Figure 3c). A significant reason for this discrepancy was a positive relationship between MAP and SOC (Supplemental Figure S1a) in contrast to the negative relationship assumed by Andrews et al. (2004). Indeed, climate is known to be one of the controlling factors for SOC content, but the role varies on a global scale (Bardgett, 2011; Xiong et al, 2014). Thus, the climate factors used in SMAF to score SOC content were reevaluated.
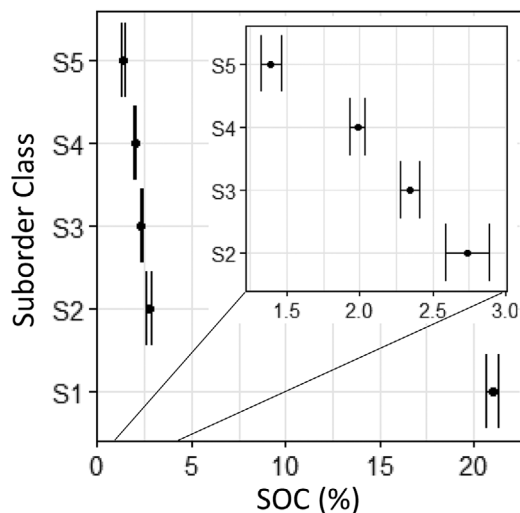
FIGURE 4 Soil organic C (SOC) as a function of the new soil taxonomy suborder classes selected for the Soil Health Assessment Protocol Evaluation SOC curve. Horizontal bars represent the 95% confidence interval (based on normality)

3.2 | Final edaphic and climate variables

The exploratory analysis confirmed that the nine inherent factors presented in Table 1 were likely good predictors of topsoil SOC content (Supplemental Figures S1–S4). For example, topsoil SOC tended to increase with MAP, AI, and WI, and decrease with MAT and PET (Supplemental Figure S1), but due to redundancy it is neither efficient nor desirable to simultaneously include all climate factors. Thus, to simplify the soil health interpretation, a subset of inherent variables was selected using results of exploratory statistical methods.

Regarding edaphic variables, the original five soil texture classes used in the SMAF (Table 2) provided adequate information on soil characteristic trends, and given the size of our dataset, a sufficient number of observations could be maintained in each class (Figure 3a). Also, the size of our dataset enabled us to use the five soil suborder classes presented in Table 3 and Figure 4. Use of soil suborders provided more specific taxonomic and soil property information than soil order (Soil Survey Staff, 2014).

Among climate variables, MAT and MAP were selected over PET, AI, and WI. Soil drainage class, which refers to the frequency and duration of wet periods in soils during their formation (Soil Survey Staff, 2014), was considered but not selected due to its relationship with soil texture, which was already included in the model. Our Random Forest exploratory analyses supported those selections as soil texture class emerged as the best predictor followed by MAT and soil suborder class (Supplemental Figure S5). Collectively, those results confirmed that soil type (i.e., suborder and texture) definitely affect topsoil SOC and should be considered in the interpretation of results. A similar conclusion emerged

when stepwise regression was used to rank the importance of model variables (Supplemental Figure S6). It showed that using both MAT and MAP provided the best prediction of SOC based on the Schwarz information criterion. Therefore, the new SHAPE scoring curves for interpreting topsoil SOC include inherent soil factors, including soil texture (Table 2), suborder class (Table 3), and continuous MAT and MAP values.

3.3 | Bayesian regression model

A Bayesian linear regression model, implemented with standalone code using the R statistical software program, was used to develop the new SHAPE–SOC scoring curves based on the final composite dataset ($n = 14,505$). It provided an estimated conditional cumulative distribution function (CDF) representing each sample's "peer group" as defined by a unique combination of categorical factors (suborder [Table 3] and texture [Table 2]) in combination with continuous MAT and MAP data. This approach assumes the dataset is representative both in range and distribution of SOC for each peer group across the climatic conditions. This assumption is more accurate for some peer groups than others, and further model validation will likely identify soil peer groups where more SOC data are needed. The SHAPE is a significant step forward toward a more precise and accurate assessment of potential SOC and, by association, soil health assessment. Furthermore, model assumptions such as normality of residuals and heteroscedasticity were evaluated, whereby we confirmed differences in variation among soil peer groups, thus confirming the need to account for heteroscedasticity.

A likelihood function accounting for differences in variation among these groups is:

$$y_{i,t,s} \overset{ind}{\sim} N\left(\mathbf{x'}_{i,t,s}\boldsymbol{\beta}, \sigma^2_{t,s}\right), \ i = 1, \ \dots, \ N$$

where $N(\cdot,\cdot)$ represents a normal distribution. The response value (SOC) is represented by $y_{i,t,s}$ for soil sample i with texture category t and suborder s. As stated previously, soil peer groups are comprised of unique combinations of texture and suborder classes, where the response variance differs for each group. The p-dimensional vector of covariates, $\mathbf{x}_{i,t,s}$, contains $T = 5$ texture category indicators, $S - 1 = 3$ suborder category indicators, as well as continuous values of MAT and MAP. Note that one level of soil suborder is dropped from the linear model for identifiability purposes.

Fitting a model as described above provides an estimate of the mean and variance associated with each conditional peer group CDF while accounting for the continuous climate covariates. Thus, for a given combination of inherent
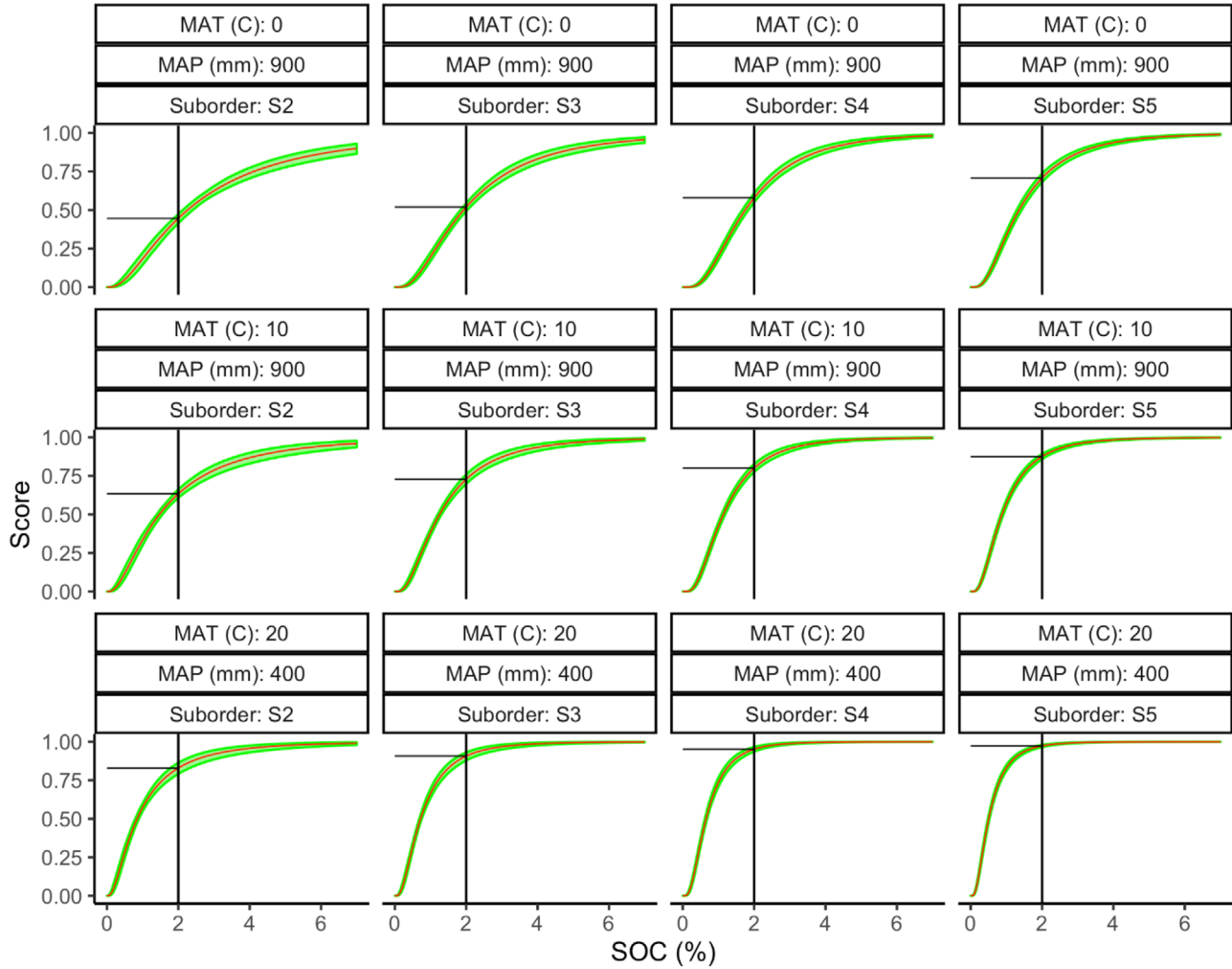
FIGURE 5 Selected SHAPE–SOC scoring curves representing four suborders (Table 3) and three MAT (mean annual temperature) and MAP (mean annual precipitation) combinations for soil texture class T1 (Table 3). Vertical bars represent a hypothetical SOC of 2%, and the horizontal bars represent scores for those SOC concentrations within each peer group. Note that the red line corresponds to the posterior mean and the green envelopes correspond to the pointwise 95% credible intervals

factors, the resulting conditional CDF is evaluated to provide the quantile of a new SOC value within that specific peer group distribution. This is then referred to as the soil health indicator score. In short, the model provides a unitless score representing the probability of being less than or equal to the observed value. The soil health score for a new SOC value of $y_{i,t,s}$ is calculated as the conditional probability of observing SOC values less than or equal to $y_{i,t,s}$ within the given peer group. This may be written as:

$$F_{y_{i,t,s}}\left(y_{i,t,s}\right) = \Phi\left(\frac{y_{i,t,s} - \mathbf{x}'_{i,t,s}\boldsymbol{\beta}}{\sigma_{t,s}}\right)$$

where $\Phi()$ is the CDF of a standard normal distribution. In summary, our primary objective is to identify a condi-

tional CDF of $y_{i,t,s}$, given $\mathbf{x}'_{i,t,s}$, and not necessarily the model parameters themselves.

If the model parameters were known, it would be simple to generate a CDF using the above equation, but in practice, the parameters must be estimated. Doing so is possible by using a weighted least squares approach to maximize the likelihood. While this approach also enables the user to quantify uncertainty associated with the estimated CDF, it is cumbersome and requires a bootstrap procedure. For each iteration of the bootstrap, an estimate of the model parameters is made, and a corresponding conditional CDF is generated.

Alternatively, uncertainty can be quantified as a byproduct of a Bayesian model fitting procedure after placing prior distributions on the model parameters. As this alleviates the need for bootstrap techniques, we propose using this method,
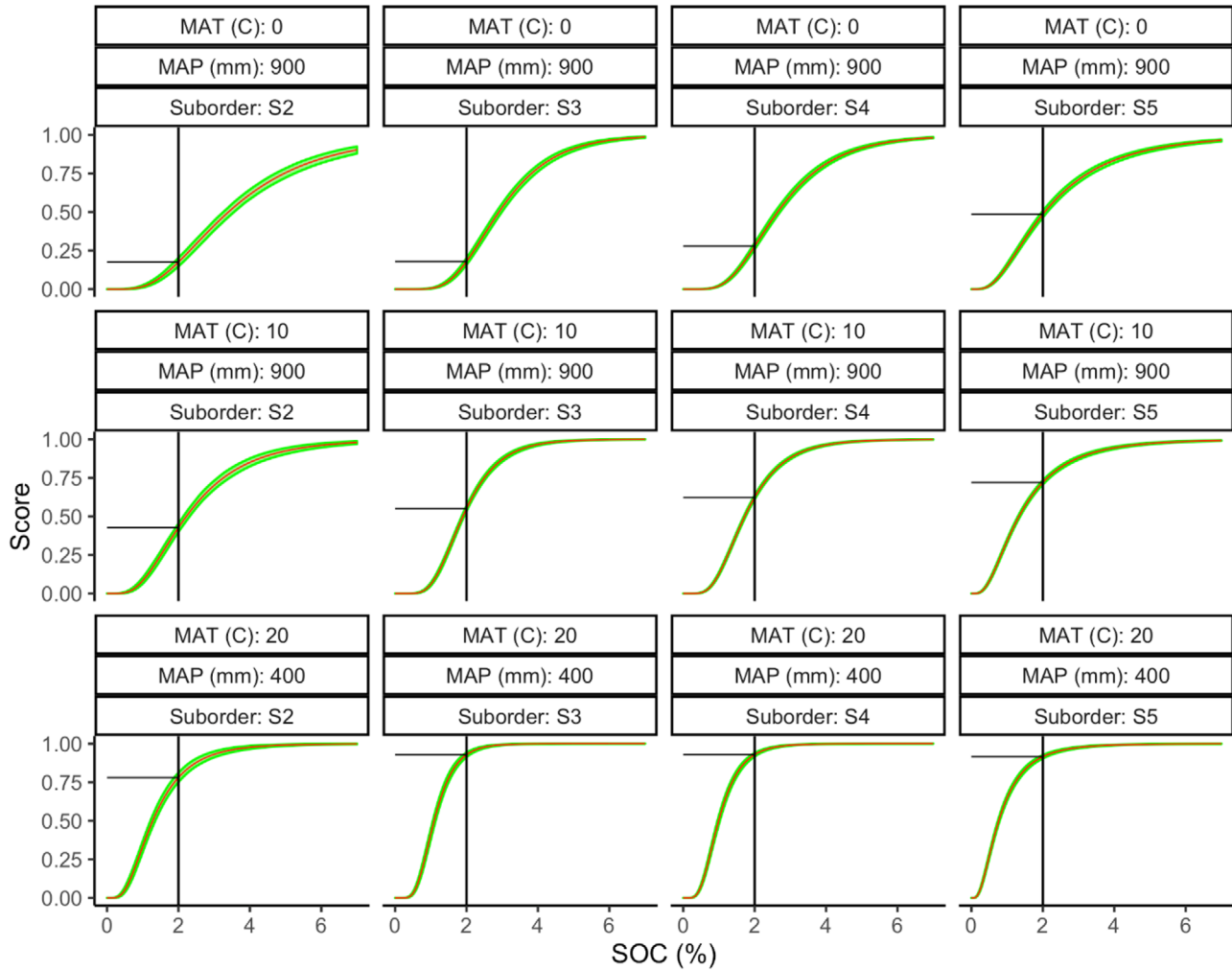
FIGURE 6 Selected SHAPE–SOC scoring curves representing four suborders (Table 3) and three MAT (mean annual temperature) and MAP (mean annual precipitation) combinations for soil texture class T2 (Table 3). Vertical bars represent a hypothetical SOC of 2%, and the horizontal bars represent scores for those SOC concentrations within each peer group. Note that the red line corresponds to the posterior mean and the green envelopes correspond to the pointwise 95% credible intervals

for which the Bayesian scoring model is:

$$y_{i,t,s} \overset{ind}{\sim} N\left(\mathbf{x'}_{i,t,s}\boldsymbol{\beta}, \sigma^2_{t,s}\right), \quad i = 1, \ldots, N$$
$$\boldsymbol{\beta} \sim N\left(0, \sigma^2_{\beta}\mathbf{I}\right)$$
$$\sigma^2_{t,s} \overset{ind}{\sim} IG\left(a, b\right), \quad t = 1, \ldots, T, \quad s = 1, \ldots, S$$

where IG( ), represents an Inverse Gamma distribution. Each element of the p-dimensional vector of coefficients, $\boldsymbol{\beta}$, follows a normal prior distribution with mean zero and variance σ^2_{β}. This approach also provides independent Inverse Gamma priors for each peer group variance term, $\sigma^2_{t,s}$. Finally, we use a vague prior distribution by setting $\sigma^2_{\beta} = 1,000$ and $a = b = 0.1$ which allows the data to dominate the parameter estimates rather than the prior distribution and, thus, imparts little impact on the resulting analysis. For all accompanying analyses, the Bayesian scoring model was fit via Gibbs

sampling with 1,000 iterations, discarding the first 500 as burn-in. Convergence was assessed visually through the use of trace plots of the sample chains, where no lack of convergence was detected.

Model fit was assessed through two methods. The first was a Bayesian posterior predictive p-value based on the discriminant (-2*loglikelihood), where values near zero or one indicate a lack of fit. For this assessment, no lack of fit was detected. Normal QQ plots of the standardized residuals were used to assess model fit visually, in terms of the underlying distributional assumptions. In this case, slight departures from normality were detected when treating SOC as the response.

We also considered the case of transforming the data before fitting the model. Specifically, when taking the logit transformation of SOC as the response, the assumption of normality seemed to hold better as assessed through the
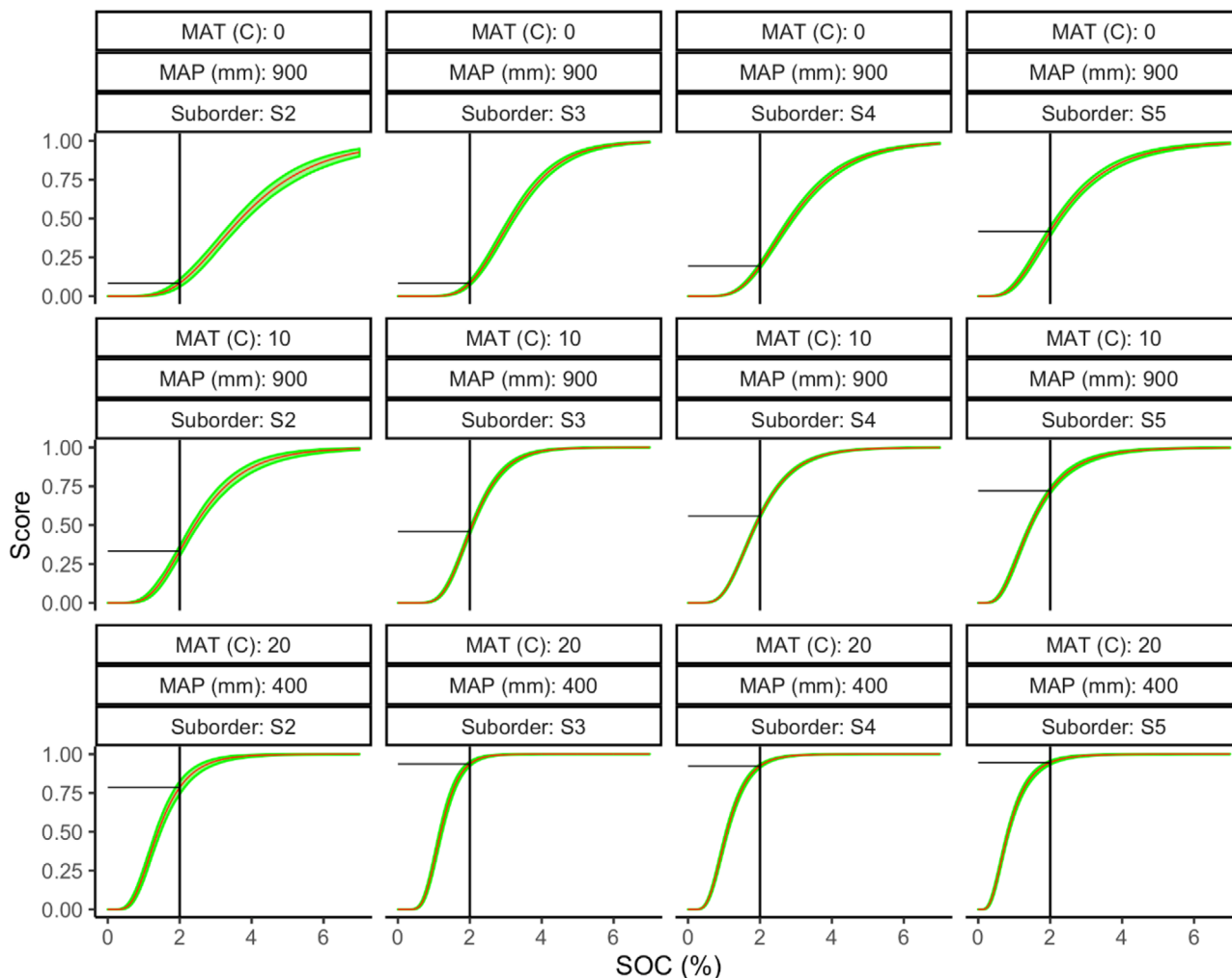
FIGURE 7 Selected SHAPE–SOC scoring curves representing four suborders (Table 3) and three MAT (mean annual temperature) and MAP (mean annual precipitation) combinations for soil texture class T3 (Table 3). Vertical bars represent a hypothetical SOC of 2%, and the horizontal bars represent scores for those SOC concentrations within each peer group. Note that the red line corresponds to the posterior mean and the green envelopes correspond to the pointwise 95% credible intervals

QQ plot. The logit transformation is given by $\text{logit}(x) = \log[x/(1-x)]$ and has the property of mapping values that are bounded between zero and one to the real line. Measurements of SOC are a percentage, and thus are bounded in this way, with many observed values lying close to zero. By fitting the model on the logit transformed data, we ensure the conditional CDF goes to zero as SOC goes to zero. This is in contrast to the model fit on untransformed data, where positive probability of SOC less than zero may be indicated. Thus, we advocate for the use of logit transformed SOC when fitting the Bayesian scoring model in order to establish appropriate bounds on the CDF and to better satisfy the normality assumption. In other words, we now assume that SOC follows a logit normal distribution.

After fitting the model using logit transformed data, it is still desirable to make an interpretation based on the original scale of the data. Since input to the CDF is on the logit transformed scale, a generated CDF curve consists of a collection of points on the transformed scale as well as their corresponding cumulative probabilities $[x_j^*, F(x_j^*)]$, $j = 1, \ldots, J$. The results are then mapped back to the original data scale by replacing x_j^* in the first coordinate with its inverse logit transformed value, which is written as $\text{logit}^{-1}(x) = 1/(1 + e^{-x})$. Thus, the conditional CDF evaluated on the original data scale consists of the points $[\text{logit}^{-1}(x_j^*), F(x_j^*)]$, $j = 1, \ldots, J$. Ultimately, we end up with an estimated conditional CDF (using the pointwise posterior mean CDF), the posterior mean probability of being less than or equal to the observed value, as well as the entire posterior distribution of conditional CDFs. The latter allows us to quantify uncertainty using the pointwise 95% credible intervals obtained from the posterior distribution of the conditional CDF.
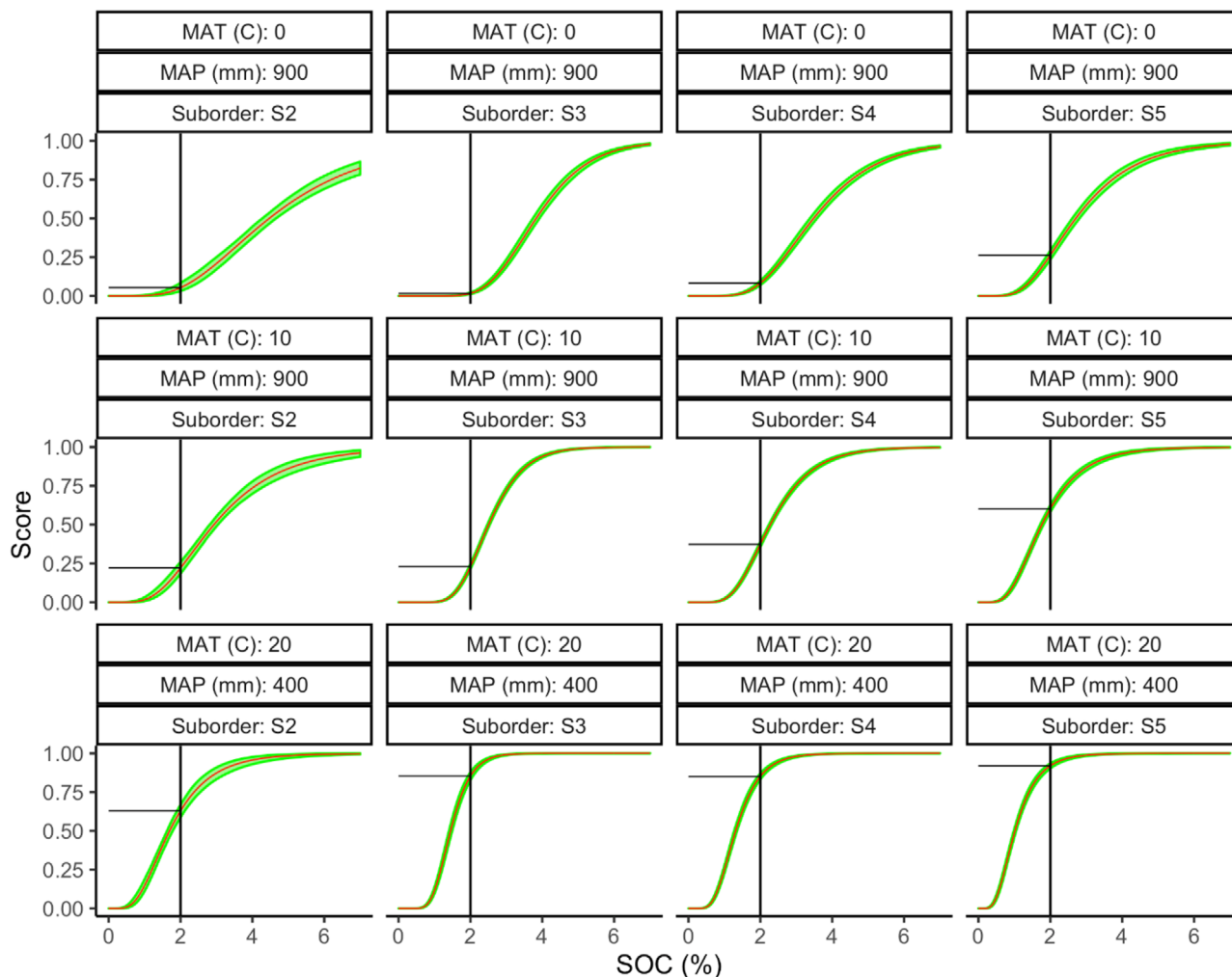
FIGURE 8 Selected SHAPE–SOC scoring curves representing four suborders (Table 3) and three MAT (mean annual temperature) and MAP (mean annual precipitation) combinations for soil texture class T4 (Table 3). Vertical bars represent a hypothetical SOC of 2%, and the horizontal bars represent scores for those SOC concentrations within each peer group. Note that the red line corresponds to the posterior mean and the green envelopes correspond to the pointwise 95% credible intervals

3.4 | Final SOC scoring curves

Figures 5–9 display 60 examples of the SHAPE–SOC scoring curves representing five soil textures, four suborders, and three combinations of MAT and MAP (i.e., 20 ˚C–400 mm, 10 ˚C–900 mm, and 0 ˚C–900 mm). Scoring curves are accompanied by 95% pointwise credible intervals to indicate the level of uncertainty around these estimates. Vertical bars within each graph represent a SOC concentration of 2% to illustrate combined effects of soil suborder and climate. The horizontal bars represent equivalent SOC scores based on the model. Upon examination, it is evident that for the same SOC concentration (2%), scores generally increase from fine-textured soils (T5) to coarse-textured soils (T1), independent of soil suborder, MAT, and MAP (Figures 5–9). This confirms that the new scoring curves are sensitive to soil texture, a dominant inherent factor affecting topsoil SOC content

(Supplemental Figure S5). Also, a comparison of scores within the same texture and climate combination, but varying across soil suborders, shows that for the same SOC concentration value (2%), scores decrease from suborder S5 (lowest inherent SOC concentration) to suborder S2 (highest inherent SOC concentration). Those changes in scores reflect the significant effect of soil taxonomy (suborder class) on topsoil SOC content (Figure 4). Finally, a comparison of scores within the same peer group (suborder and textural combination) demonstrates that scores for the same SOC value decrease from higher to lower temperature and from lower to higher precipitation. Again, those changes reflect the strong influence of climate on inherent topsoil SOC content.

The SHAPE approach expands upon the cumulative normal distribution method used in CASH and provides Bayesian model-based mean and variance estimates for peer groups to produce a conditional CDF. The SHAPE thus gains strength
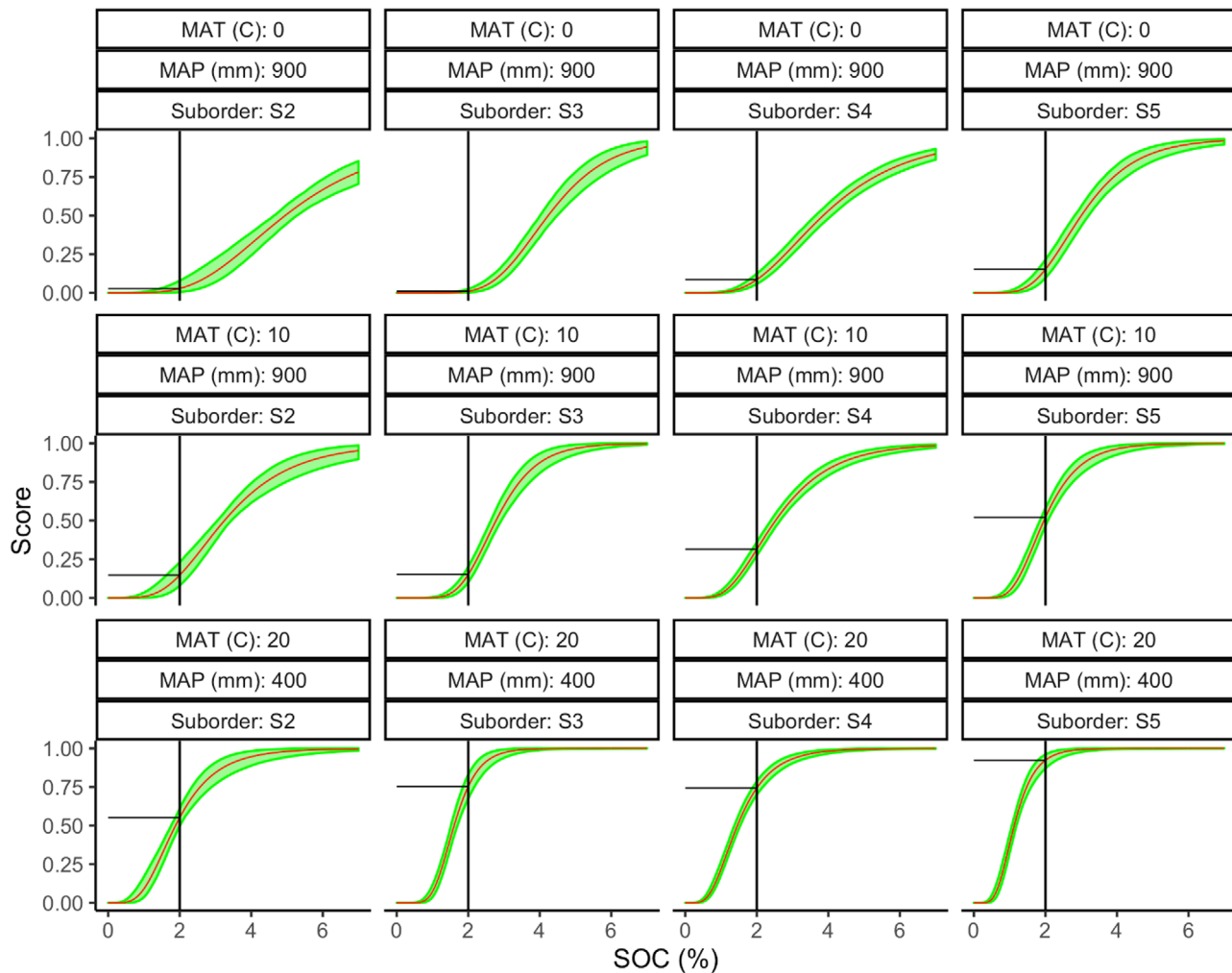
FIGURE 9 Selected SHAPE–SOC scoring curves representing four suborders (Table 3) and three MAT (mean annual temperature) and MAP (mean annual precipitation) combinations for soil texture class T5 (Table 3). Vertical bars represent a hypothetical SOC of 2%, and the horizontal bars represent scores for those SOC concentrations within each peer group. Note that the red line corresponds to the posterior mean and the green envelopes correspond to the pointwise 95% credible intervals

from relationships among all dataset variables, as opposed to calculating a sample mean and SD directly from each peer group subset. Specifically, this method "borrows strength" from the fact that continuous (i.e., climate) variables are estimated with the data points across all groups. Furthermore, the SHAPE provides a natural framework for quantifying uncertainty in the scoring curve and underlying model parameters. The SHAPE also builds on the SMAF framework by accounting for multiple inherent factors as well as continuous climate variables known to influence soil properties. The SMAF accounted for those factors using a series of logic statements to create interpretation algorithms where "more is better," "less is better," or "optimum value" curves were fit to the data using somewhat arbitrary parameters (i.e., inductive interpretations or expert opinion). In contrast to the SMAF, SHAPE does not attempt to adjust scores based on thresholds linked to soil functions or ecosystem services (e.g., the

concept of diminishing returns). That approach requires selection of specific ecosystem services, access to an associated dataset with appropriate scope and scale, as well as definable, quantifiable relationships with measured soil health indicators. For example, the importance of SOC and overall soil health for drought resilience is moderated by complex interactions between weather patterns and crop genetics (Pareek et al, 2020). From its inception, users of the SMAF have acknowledged that there are multiple benefits and complex tradeoffs (e.g., provisioning and environmental protection) reflected by changes in soil health indicator values. Ecosystems provide multiple services, both tangible and intangible, and the value of those services is complex and scale dependent (Simpson, 2016). Thus, SHAPE does not weight scores based on relationships with specific ecosystem services or the valuation of those services at this time. Rather than adjust scores based on thresholds linked to a single soil function such as crop

TABLE 4 An overview of SHAPE, SMAF, and CASH ANOVA results from a reanalysis of previously published case studies from New York (NY) (Nunes et al., 2018) and Missouri (MO) (Veum et al., 2015)

State and texture group	Treatment	SOC%	SHAPE	CASH	SMAF
NY T4 (clay loam)	PT	1.81a	0.25b	0.21b	0.89b
	NT	2.12a	0.37a	0.37a	0.96a
	ANOVA	ns	*	*	*
NY T1 (sandy)	PT	0.73b	0.20b	0.19b	0.23b
	NT	1.23a	0.39a	0.55a	0.61a
	ANOVA	***	***	***	***
NY T3 (silt loam)	PT–NCC	1.49b	0.20b	0.32b	0.81b
	PT–CC	1.55b	0.22b	0.36b	0.84b
	NT–NCC	1.69ab	0.29ab	0.46ab	0.88ab
	NT–CC	2.12a	0.34a	0.54a	0.93a
	ANOVA	***	***	***	***
MO T3 (silt loam)	CSG	1.72a	0.47a	0.48a	0.88a
	HAY	1.60ab	0.41ab	0.40ab	0.85ab
	WSG	1.44bc	0.32bc	0.29abc	0.78abc
	NTCSW	1.39bc	0.29bcd	0.26bc	0.75bc
	NTCS	1.25cd	0.21cd	0.19c	0.67cd
	MTCS	1.22d	0.20d	0.16c	0.61d
	ANOVA	***	***	***	***

Note. CC, cover crops; CSG, cool-season grass; Hay, mowed mixed grass; MTCS, mulch-till corn–soybean; NCC, no cover crop; NT, no-till; NTCS, no-till corn–soybean; NTCSW, no-till corn (*Zea mays* L.)–soybean–wheat; PT, full width plow tillage; WSG, wheat (*Triticum aestivum* L.)–soybean [*Glycine max* (L.) Merr.]–season grass; SOC concentrations and SHAPE, SMAF, and CASH scores followed by the same letter are not significantly different at α =.05 using Tukey's HSD.
*Significant at the .05 probability level. ***Significant at the .001 probability level.

yield or water quality, SHAPE takes into consideration the full range of SOC values for a specific soil group under a specific climate condition. A repository that can be used to batch score soil SOC data using the approach described above (i.e., SHAPE method) is available at https://github.com/paparker/SHAPE.

3.5 | Sensitivity to management (case studies)

It is widely recognized that inherent factors (i.e., climate, texture, and soil suborder) define a soil's potential SOC content. The SHAPE accounts for those dominant factors at the continental scale, but effective soil health indices must also be able to discriminate among broad land use categories and specific, field-scale soil and crop management practices to be useful to policy makers, conservationists, landowners, and producers. We evaluated the utility of SHAPE for meeting those needs by reanalyzing four case studies where SOC data had been scored using the SMAF (Table 4). The first three are from long-term experiments (20+ yr) comparing continuous no-tillage and plow-till management on several soil health indicators and corn (*Zea mays* L.) yield in New York (Nunes et al.,

2018). Inherent conditions associated with the experimental sites were: Site 1 = texture class T4, soil suborder class S4, MAP of 760 mm, and MAT of 7.8 °C; Site 2 = texture class T1, soil suborder class S2, MAP of 760 mm, and MAT of 7.8 °C; and Site 3 = texture class T3, soil suborder class S4, MAP of 1,070 mm, and MAT of 7.9 °C. Within Site 3, tillage effects were assessed with and without an interseeded cover crop. The fourth case study represents an experimental site in the Central Claypan Region of Missouri (Veum et al., 2015), for which soil samples were collected in 2008 from the 0-to-15-cm depth using a RCB design with three blocks. Tillage, crop rotation, and the perennial nature of the cropping system were variables at the Missouri site.

Collectively the results show the SHAPE, SMAF, and CASH discriminated between annual and perennial systems and were sensitive to management practices (Table 4). Overall, SHAPE and CASH scores were similar, while SMAF generated substantially higher scores among the three tools, independent of management (Table 4; Supplemental Figure S7). As previously described, the soil suborder groupings and reliance on a small dataset of primarily agricultural soils in SMAF likely explains this discrepancy and the issues with overestimation identified in other studies. For example, with the exception of the NY T1 sandy case study, SMAF scores for

SOC suggested soils under more than 20 yr of intensive tillage and monoculture were high-functioning (scores >0.81). In contrast, SHAPE indicates that, for the given set of inherent conditions (soil taxonomy, soil texture, and MAT and MAP), SOC concentration could be much higher. Therefore, based on the SHAPE scores, we may conclude that SOC concentration in those four studied soils could improve depending on land use and management. These examples are provided simply to illustrate the utility of the new SHAPE curves and to invite future studies to evaluate the effects of land use and management on SOC using this framework.

4 | CONCLUSION

A robust and principled Bayesian statistical foundation for developing a new soil health assessment interpretation tool (called SHAPE) was developed and illustrated using new soil organic C (SOC) scoring curves. The curves were built using a comprehensive, nationwide dataset with 14,505 SOC observations. The SHAPE builds upon conceptual frameworks established by the SMAF and CASH protocols. A suite of edaphic and climate variables were evaluated to account for the inherent factors that define a soil's site-specific SOC potential across the continental United States. The Bayesian linear regression model provides a score based on estimated conditional CDF of each soil's peer group based on a combination of categorical factors (soil suborder and texture class) with adjustments for continuous climate variables. Reanalysis of published case studies confirmed sensitivity to land use and field-scale management. Future SHAPE developments are anticipated in order to generate scoring curves for multiple soil health indicators, especially as the comprehensive dataset continues to grow. Full development of the SHAPE will help meet the growing demand for an accessible, interpretive, and quantitative scoring curve that provides regionally relevant knowledge regarding the status of soils in response to various agronomic and conservation initiatives.

AUTHOR CONTRIBUTIONS
Márcio R. Nunes, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Visualization, Visualization, Writing-original draft, Writing-review & editing; Kristen S. Veum, Conceptualization, Formal analysis, Investigation, Methodology, Resources, Supervision, Writing-review & editing; Paul A. Parker, Conceptualization, Formal analysis, Methodology, Resources, Visualization, Writing-original draft; Scott H. Holan, Conceptualization, Formal analysis, Methodology, Resources, Writing-review & editing; Douglas L. Karlen, Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing-review & editing; Joseph P. Amsili, Conceptualization, Data curation, Formal analysis, Methodology, Resources, Writing-review & editing; Harold M. van Es, Conceptualization, Data curation, Methodology, Resources, Writing-review & editing; Skye A. Wills, Conceptualization, Data curation, Methodology, Resources, Writing-review & editing; Cathy A. Seybold, Conceptualization, Data curation, Methodology, Resources, Writing-review & editing; Thomas B. Moorman, Conceptualization, Methodology, Project administration, Resources, Supervision, Writing-review & editing.

CONFLICT OF INTEREST
The authors declare no conflicts of interest.

ORCID
Márcio R. Nunes https://orcid.org/0000-0002-3674-279X
Kristen S. Veum https://orcid.org/0000-0002-6492-913X
Joseph P. Amsili https://orcid.org/0000-0002-8293-5824
Harold M. van Es https://orcid.org/0000-0001-9822-9476

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

Baker, J. M., Ochsner, T. E., Venterea, R. T., & Griffis, T. J. (2007). Tillage and soil carbon sequestration–What do we really know? *Agriculture Ecosystem and Environment*, *118*, 1–5. https://doi.org/10.1016/j.agee.2006.05.014

Bardgett, R. D. (2011). Plant-soil interactions in a changing world. *F1000 Biology Reports*, *3*, 16. https://doi.org/10.3410/B3-16

Biswas, S., Hazra, G. C., Purakayastha, T. J., Saha, N., Mitran, T., Roy, S. S., Basak, N., & Mandal, B. (2017). Establishment of critical limits of indicators and indices of soil quality in rice–rice cropping systems under different soil orders. *Geoderma*, *292*, 34–48. https://doi.org/10.1016/j.geoderma.2017.01.003

Bowles, T. M., Jackson, L. E., Loher, M., & Cavagnaro T. R. (2016). Ecological intensification and arbuscular mycorrhizas: A meta-analysis of tillage and cover crop effects. *Journal of Applied Ecology*, *54*, 1785–1793. https://doi.org/10.1111/1365-2664.12815

Bünemann, E. K., Bongiorno, G., Bai, Z., Creamer, R. E., Deyn, G., Goede, R., Fleskens, L., Geissen, V., Kuyper, T. W., Mäder, P., Pulleman, M., Sukkel, W., van Groenigen, J. W., & Brussaard, L. (2018). Soil quality–a critical review. *Soil Biology and Biochemistry*, *120*, 105–125. https://doi.org/10.1016/j.soilbio.2018.01.030

Cambardella, C. A., Gajda, A. M., Doran, J. W., Wienhold, B. J., & Kettler, T. A. (2001). Estimation of particulate and total organic matter by weight loss-on-ignition. In R. Lal, J. M. Kimble, R. J. Follett, & B. A. Stewart (Eds.), *Assessment methods for soil carbon* (pp. 349–359). Lewis Publishers/CRC Press.

De Paul Obade V., & Lal, R. (2016). Towards a standard technique for soil quality assessment. *Geoderma*, *265*, 96–102. https://doi.org/10.1016/j.geoderma.2015.11.023

Doran, J. W. (2002). Soil health and global sustainability: Translating science into practice. *Agriculture Ecosystem and Environment*, *88*, 119–127. https://doi.org/10.1016/S0167-8809(01)00246-8

Doran, J. W., & Parkin, T. B. (1996). Quantitative indicators of soil quality: A minimum data set. In J. W. Doran, & A. J. Jones (Eds.), *Methods for assessing soil quality*. Soil Science Society of America.

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*, 589–601. https://doi.org/10.2136/sssaj2016.09.0286

Graaff, M. A., Hornslein, N., Throop, H., Kardol, P., & van Diepen, L. T. (2019). Effects of agricultural intensification on soil biodiversity and implications for ecosystem functioning: A meta-analysis. *Advances in Agronomy*, *155*, 1–44. https://doi.org/10.1016/bs.agron.2019.01.001

Haddaway, N. R., Hedlund, K., Jackson, L. E., Kätterer, T., Lugato, E., Thomsen, I. K., Jørgensen, H. B., & Isberg, P. E. (2017). How does tillage intensity affect soil organic carbon? A systematic review. *Environmental Evidence*, *6*, 2–48.

Haney, R. L., Haney, E. B., Smith, D. R., Harmel, R. D., & White, M. J. (2018). The soil health tool–Theory and initial broad-scale application. *Applied Soil Ecology*, *125*, 162–168. https://doi.org/10.1186/s13750-017-0108-9

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*, 214–224. https://doi.org/10.1017/S1742170509990068

Jenny, H. (1941). *Factors of soil formation: A system of quantitative pedology*. McGraw Hill.

Jobbágy, E. G., & Jackson, R. B. (2000). The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecological Applications*, *10*, 423–436. https://doi.org/10.1890/1051-0761(2000)0100423:TVDOSO2.0.CO;2

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

Karlen, D. L., Mausbach, M. J., Doran, J. W., Cline, R. G., Harris, R. F., & Schuman, G. E. (1997). Soil quality: A concept, definition, and framework for evaluation. *Soil Science Society of America Journal*, *61*, 4–10. https://doi.org/10.2136/sssaj1997.03615995006100010001x

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

Lopes, A. A. C., Sousa, D. M. G., Chaer, G. M., Reis Junior, F. B., Goedert W. J., & Mendes, I. C. (2013). Microbial soil indicators as a function of crop yield and organic carbon. *Soil Science Society of America Journal*, *77*, 461–472. https://doi.org/10.2136/sssaj2012.0191

Lumly, T. (2020). leaps: Regression Subset Selection. R package version 3.1. https://cran.r-project.org/web/packages/leaps/leaps.pdf

Luo, Z., Wang, E., & Sun, O. J. (2010). Can no-tillage stimulate carbon sequestration in agricultural soils? A meta-analysis of paired experiments. *Agriculture Ecosystem and Environment*, *139*, 224–231. https://doi.org/10.1016/j.agee.2010.08.006

Mbuthia, L. W., Acosta-Martínez, V., DeBruyn, J., Schaeffer, S., Tyler, D., Odoi, E., Mpheshea, M., Walker, F., & Eash N. (2015). Long term tillage, cover crop, and fertilization effects on microbial community structure, activity: Implications for soil quality. *Soil Biology and Biochemistry*, *89*, 24–34. https://doi.org/10.1016/j.soilbio.2015.06.016

McDaniel, M. D., Tiemann, L. K., & Grandy, A. S. (2014). Does agricultural crop diversity enhance soil microbial biomass and organic matter dynamics? A meta-analysis. *Ecological Applications*, *24*, 560–570. https://doi.org/10.1890/13-0616.1

Moebius-Clune, B. N., Moebius-Clune, D. J., Gugino, B. K., Idowu, O. J., Schindelbeck, R. R., Ristow, A. J., van Es, H. M., This, J. E., Shayler, H. A., McBride, M. B., Kurtz, K. S. M., Wolfe, D. W., & Abawi, G. S. (2016). *Comprehensive assessment of soil health* (3rd ed.). bit.ly/SoilHealthTrainingManual

Nunes, M. R., Karlen, D. L., & Moorman, T. B. (2020). Tillage intensity effects on soil structure indicators–A US meta-analysis. *Sustainability*, *12*, 1–17. https://doi.org/10.3390/su12052071

Nunes, M. R., Karlen, D. L., Moorman, T. B., & Cambardella, C. A. (2020). How does tillage intensity affect chemical soil health indicators? A United States meta-analysis. *Agrosystems, Geosciences and Environment*, *3*, e20083. https://doi.org/10.1002/agg2.20083

Nunes, M. R., Karlen, D. L., Veum, K. L., Moorman, T. B., & Cambardella, C. A. (2020). Biological soil health indicators respond to tillage intensity: A US meta-analysis. *Geoderma*, *369*, 114335. https://doi.org/10.1016/j.geoderma.2020.114335

Nunes, M. R., Karlen, D. L., Veum, K. L., & Moorman, T. B. (2020). A SMAF assessment of U.S. tillage and crop management strategies. *Environmental and Sustainability Indicators*, *369*, 100072. https://doi.org/10.1016/j.indic.2020.100072

Nunes, M. R., Karlen, D. L., Denardin, J. E., & Cambardella, C. A. (2019). Corn root and soil health indicator response to no-till production practices. *Agriculture, Ecosystems and Environment*, *285*, 106607. https://doi.org/10.1016/j.agee.2019.106607

Nunes, M. R., van Es, H. M., Schindelbeck, R. R., Ristow, A. J., & Ryan, M. (2018). No-till and cropping system diversification improve soil health and crop yield. *Geoderma*, *328*, 30–43. https://doi.org/10.1016/j.geoderma.2018.04.031

Pareek, A., Dhankher, O. P., & Foyer, C. H. (2020). Mitigating the impact of climate change on plant productivity and ecosystem sustainability. *Journal of Experimental Botany*, *71*, 451–456. https://doi.org/10.1093/jxb/erz518

Post, W. M., Izaurralde, R. C., Jastrow, J. D., McCarl, B. A., Amonette, J. E., Bailey, V. L., & Zhou, J. (2004). Enhancement of carbon sequestration in US soils. *Bioscience*, *54*, 895–908. https://doi.org/10.1641/0006-3568(2004)0540895:EOCSIU2.0.CO;2

Quisenberry, V. L., Smith, B. R., Phillips, R. E., Scott, H. D., & Nortcliff, S. (1993). A soil classification system for describing water and chemical transport. *Soil Science*, *156*, 306–315.

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing

Rinot, O., Levy, G. J., Steinberger, Y., Svoray, T., & Eshel, G. (2019). Soil health assessment: A critical review of current methodologies and a proposed new approach. *Science of The Total Environment*, *648*, 1484–1491. https://doi.org/10.1016/j.scitotenv.2018.08.259

Ritz, K., Black, H. I., Campbell, C. D., Harris, J. A., & Wood, C. (2009). Selecting biological indicators for monitoring soils: A framework for balancing scientific and technical opinion to assist policy development. *Ecological Indicators*, *9*, 1212–1221. https://doi.org/10.1016/j.ecolind.2009.02.009

Simpson, R. D. (2016). Putting a price on ecosystem services. *Issues in Science and Technology*, *32*(4). https://issues.org/putting-a-price-on-ecosystem-services/

Soil Survey Staff (1999). *Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys* (2nd ed.). Agriculture Handbook No. 436, USDA–NRCS. https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcs142p2_051232.pdf

Soil Survey Staff (2014). Kellogg soil survey laboratory methods manual. Soil Survey Investigations Report No. 42, USDA–NRCS. https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1253871.pdf

Soil Survey Staff, NRCS, USDA. *Soil Survey Geographic (SSURGO) Database*. https://sdmdataaccess.sc.egov.usda.gov

Stott, D. E., Karlen, D. L., Cambardella, C. A., & Harmel, R. D. (2013). A soil quality and metabolic activity assessment after fifty-seven years of agricultural management. *Soil Science Society of America Journal*, *77*, 903–913. https://doi.org/10.2136/sssaj2012.0355

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

Tiemann, L. K., Grandy, A. S., Atkinson, E. E., Marin-Spiotta, E., & McDaniel, M. D. (2015). Crop rotational diversity enhances belowground communities and functions in an agroecosystem. *Ecology Letters*, *18*, 761–771. https://doi.org/10.1111/ele.12453

Trumbore, S. E. (1997). Potential responses of soil organic carbon to global environmental change. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(16), 8284–8291.

Veum, K. S., Kremer, R. J., Sudduth, K. A., Kitchen, N. R., Lerch, R. N., Baffaut, C., Stott, D. E., Karlen, D. L., & Sadler, E. J. (2015). Conservation effects on soil quality indicators in the Missouri Salt River Basin. *Journal of Soil and Water Conservation*, *70*, 232–246. https://doi.org/10.2489/jswc.70.4.232

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*(109). https://doi.org/10.3389/fenvs.2019.00109

Xiong, X., Grunwald, S., Myers, D. B., Ross, C. W., Harris, W. G., & Comerford, N. B. (2014). Interaction effects of climate and land use/land cover change on soil organic carbon sequestration. *Science of the Total Environment*, *493*, 974–982. https://doi.org/10.1016/j.scitotenv.2014.06.088

Zobeck, T. M., Steiner, J. L., Stott, D. E., Duke, S. E., Starks, P. J., Moriasi, D. N., & Karlen, D. L. (2015). Soil quality index comparisons using Fort Cobb, Oklahoma, watershed-scale land management data. *Soil Science Society of America Journal*, *79*, 224–238. https://doi.org/10.2136/sssaj2014.06.0257

SUPPORTING INFORMATION

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