

A Comparison of Soil Texture-by-Feel Estimates: Implications for the Citizen Soil Scientist

Shawn W. Salley*

Jeffrey E. Herrick

Jornada Research Unit
USDA-Agricultural Research Service,
MSC 3JER, NMSU Box 30003
Las Cruces, NM 88003-8003

Caitlin V. Holmes

Jornada Research Unit
New Mexico State Univ.
MSC 3JER, NMSU, Box 30003
Las Cruces, NM 88003-8003

Jason W. Karl

Dep. of Forestry, Rangeland, and Fire
Sciences
Univ. of Idaho
875 Perimeter Dr., MS 1135
Moscow, ID 83844-1135

Matthew R. Levi

Dep. of Crop and Soil Sciences
Univ. of Georgia
120 Carlton Street
Athens, GA 30602

Sarah E. McCord

Jornada Research Unit
USDA-Agricultural Research Service,
MSC 3JER, NMSU Box 30003
Las Cruces, NM 88003-8003

Cornelis van der Waal

Agri-Ecological Services
Box 510, Omaruru, 9000, Namibia

Justin W. Van Zee

Jornada Research Unit
USDA-Agricultural Research Service,
MSC 3JER, NMSU Box 30003
Las Cruces, NM 88003-8003

Estimating soil texture is a fundamental practice universally applied by soil scientists to classify and understand the behavior, health, and management of soil systems. While the accuracy of both the soil texture class and the estimates of the percentage of sand and clay is generally accepted when completed by trained soil scientists, similar estimates by “citizen scientists” or less experienced seasonal resource scientists are often questioned. We compared soil texture classes determined by texture-by-feel and laboratory analyses for two groups: professional soil scientists who contributed to the USDA-NRCS National Soil Characterization Database and seasonal field technicians working on rangeland inventory and assessment programs in the Western United States and Namibia. Texture accuracy was compared using a confusion matrix to evaluate classification accuracy based on the assumption that laboratory measurements were correct. Our results show that the professional soil scientists predicted the laboratory-determined texture class for 66% of the samples. Accuracy for seasonal field technicians was between 27 and 41%. When a “correct” prediction was defined to include texture classes adjacent to the laboratory-determined texture based on a standard USDA texture triangle, accuracy increased to 91% for professionals and 71 to 78% for seasonal field technicians. These findings highlight the need to improve options for increasing the accuracy of field-textured estimates for all soil texture observers, with relevance to career soil scientists, seasonal technicians, and citizen scientists. Opportunities for improving soil texture accuracy include training, calibration, and decision support tools that go beyond simple dichotomous keys.

Texture is considered one of the most important properties influencing nearly all soil processes, functions, and properties, and is the most requested analytical procedure of national soil survey laboratories (Soil Survey Staff, 2014a). Soil texture is determined through established laboratory procedures that measure the relative proportion of soil separates (Gee and Bauder, 1986; Zobeck, 2004) or in the field where the sample’s apparent “texture-by-feel” is estimated based on grittiness, cohesiveness, and stickiness (Rowell, 2014; Thien, 1979). Field estimation of soil texture is considered a fundamental practice (Franzmeier and Owens, 2008) universally applied by resource scientists to classify and understand the behavior, health, and management of soil systems. Because laboratory analysis of soil texture is costly and often takes time to analyze, recent work has suggested that texture-by-feel estimates can replace laboratory analysis altogether (Vos et al., 2016). While it is believed experienced soil scientists often estimate texture class and the percentage of separates (clay and sand) of mineral soils with relatively high precision (Post et al., 2006; Vos et al., 2016), past work has shown inexperienced soil observers require training, practice, and calibration before similar texture-by-feel proficiency can be obtained (Levine et al., 1989).

Previous studies have evaluated the ability of experienced or inexperienced soil observers to estimate apparent field textures against laboratory-derived particle-size analysis (Akamigbo, 1984; David, 1999; Foss et al., 1975; Hodgson et

Core Ideas

- Soil scientists estimate soil texture class with higher accuracy than previously reported in the literature.
- Seasonal field scientists and citizen scientists estimate texture-by-feel with similar accuracy to university students with limited training.
- When novice observers misclassify texture class, it is more likely because of errors in estimating ribbon length than estimating grittiness.

Soil Sci. Soc. Am. J. 82:1526–1537

doi:10.2136/sssaj2018.04.0137

Received 6 Apr. 2018.

Accepted 7 Aug. 2018.

*Corresponding author (Shawn.Salley@ars.usda.gov).

© Soil Science Society of America, 5585 Guilford Rd., Madison WI 53711 USA. All Rights reserved.

Table 1. Previous studies comparing soil texture-by-feel to laboratory-derived particle-size analysis.

Location	Sample size	Number of people	Years completed	Soil type	Training	Source
Maryland, U.S.	598	many	1955–1970	127 profiles, 38 soil series	USDA-SCS soil survey	(Foss et al., 1975)
Wales, U.K.	184	4	1973–1975	47 profiles 38 series	England and Wales Soil Survey	(Hodgson et al., 1976)
Arizona, U.S.	25	36	1982	20 series	SCS (27), USFS (7), BLM (2) NCSS soil survey	(Post et al., 1986)
Nigeria	450	Unknown	1977–1982	Multiple pedons	Unknown	(Akamigbo, 1984)
Nigeria	100	2		Multiple pedons	Soil science graduates w/basic training	(Ogunkunle, 1993)
Nigeria	100	2	< 1999	Multiple pedons	Soil science graduates w/basic training	(David, 1999)
Arizona, U.S.	20	129	1997–2004	Multiple pedons	College students, 4 wk	(Post et al., 2006)
Arizona, U.S.	15	115	1983–1987	Multiple pedons	College students, 3 wk and end of semester	(Levine et al., 1989)
Continental, U.S.	1002	~209	< 2002	209 Pedons	NCSS trained	(Rawls and Pachepsky, 2002)
Australia	19,500	many	< 2007	Multiple pedons	Trained soil scientists	ASRIS database (Minasny et al., 2007)
Germany	3896	8	< 2016	728 sites	Unknown	(Vos et al., 2016)
Western U.S.	155	40	2015–2016	75 Pedons	Minimal, 1 d	This Study
Namibia	346	7	2014	175 Pedons	3 d	This Study
U.S./Global	237,068	many	< 2017	> 21,000	NCSS trained soil scientists	This Study

Abbreviations: USDA, United States Department of Agriculture; SCS, Soil Conservation Service (now Natural Resources Conservation Service); USFS, United States Forest Service; BLM, Bureau of Land Management; ASRIS, Australian Soil Resource Information System; NCSS, National Cooperative Soil Survey, U.K., United Kingdom, and U.S., United States.

al., 1976; Levine et al., 1989; Minasny et al., 2007; Ogunkunle, 1993; Post et al., 1986, 2006; Rawls and Pachepsky, 2002; Vos et al., 2016). However, most of this work was based on limited datasets, constrained to highly trained soil scientists or university students, or focused only on the percentage of clay estimation (Table 1). In general, research has shown trained soil scientists have demonstrated approximately 50% absolute accuracy (Minasny et al., 2007), where absolute accuracy was defined as an exact match with the laboratory determination of soil texture class (see Box 1). Studies of student proficiency published prior to 1989 reported that absolute accuracy for less experienced observers averaged approximately 40% (Levine et al., 1989). Previous work (Table 2) has also shown that both novice and skilled observers predict Sandy- and Clayey- classes well, while their absolute accuracy of medium texture samples (Silty- and

Loamy-) was somewhat lower. Novice observers found the predominately Silty- classes to be the most difficult to predict. For example, the Silt Loam texture class showed the greatest separation between novice and professional soil scientists (17 and 59% respectively; Levine et al., 1989).

Box 1. Laboratory analyses as reference for field estimates. As in previous texture-by-feel studies, we used laboratory analyses as the reference for accuracy assessments. Because laboratory methods can be imprecise, and at times inaccurate, the maximum possible accuracy for the field estimates may be much <100%. Reported within-laboratory error rates for sedimentation methodologies (pipette and hydrometer) range from 0 to 6% with higher absolute differences in sand and clay (6 and -4%, respectively) then silt (-1%) texture classes (Kettler et al., 2001). Furthermore, traditional laboratory techniques under es-

Table 2. Results of previous studies (see Table 1) which reported individual soil texture class accuracy of texture-by-feel compared with laboratory-derived particle-size analysis.

Previous Studies	N†	Soil Texture Class Accuracy‡(%)											OA%§
		S	LS	SL	SC	SCL	L	SiL	CL	SiCL	SiC	C	
Akamigbo, 1984	455	95	73	58	71	47	31	21	25	2	46	46	55
Foss et al., 1975	598	74	35	68	-	44	36	63	18	19	-	73	50
Post et al., 1986	25	86	44	83	-	30	26	59	-	36	-	36	46
Levine, 1989- 1 ¶	15	81	-	48	-	33	27	19	50	41	47	48	39
Levine, 1989- 2 ¶	30	81	50	48	-	35	27	17	-	33	31	47	45
Rawls and Pachepsky, 2002	1002	80	34	32	-	58	39	48	28	41	48	48	41
Minasny et al., 2007 #	17,979	75	41	50	14	22	27	44	25	21	17	74	-

† Sample Size.

‡ S, sand; LS, loamy sand; SL, sandy loam; SC, sandy clay; SCL, sandy clay loam; L, loam; SiL, silt loam; Si, silt; CL, clay loam; SiCL, silty clay loam; SiC, silty clay; c, clay.

¶ OA, Overall Accuracy is the percent correctly classified.

§ Levine 1, college course end of week 3; Levine 2, college course end of semester.

Averaged from texture classes of Northcote (1971) based on Australian field texture classes (McDonald et al., 1998).

timate clay in highly weathered oxide-rich tropical and volcanic ash soils which resist dispersion (Silva et al., 2015). No methods for particle-size analysis can serve as a universal benchmark as all methods (sedimentation, sieve, and laser diffraction) suffer from inherent biases based on particle shape, sphericity, density, chemical composition, and pretreatment methods (Eshel et al., 2004).

Because soil science's main objective is to obtain reliable information about the nature, properties, dynamics, and function of the soil (Brady and Weil, 2013; FAO, 2006), a basic requirement for attaining that objective is to reliably assess soil morphology through examining and describing field soil properties. Individuals who describe field soil properties come from a range of backgrounds. We have observed the level and training received by individuals completing field texture-by-feel estimates varies widely, including trained professionals with a degree in soil science, ecologists, or other scientific specialty with limited university soils experience, early career field scientists with no formal soils training, and citizen science observers whose soils background is difficult to evaluate. Considering that surface and subsurface soil texture is often the most useful soil property for interpreting soil potential (Nauman and Duniway, 2016), an important question persists regarding the accuracy of field based texture-by-feel estimates completed by individuals with different levels of training and experience.

There is increasing interest in using crowdsourcing and citizen scientists to collect and interpret natural resource information (McKinley et al., 2017; Rossiter et al., 2015). Citizen science's goal is to engage nonprofessionals in the process of generating and sharing scientific knowledge while expanding access to science-based interpretations. Even though citizen science has the potential to generate a great wealth of information, the reliability of volunteer-produced data is often questioned compared with information collected by professionals (Riesch and Potter, 2014). Furthermore, few projects have evaluated the accuracy and bias of professionally produced data within the same context as volunteer-produced data (Kosmala et al., 2016). While confirming overall education and experience of a citizen soil scientist is difficult, we suggest that they may be comparable with seasonal, early career field scientists (hereafter, simply called technicians) who often lack formal soils training and experience, although we acknowledge that some citizen scientists may possess advanced training and experience (such as a retired soil scientists) or be completely novice with no knowledge of the discipline.

The objectives of this study were to evaluate the accuracy texture-by-feel class estimates of mineral soil by individuals with differing levels of training and experience, and to propose a set of recommendations for improving these estimates. While previous studies of texture-by-feel class accuracy have evaluated professional versus novice soil scientists (e.g., Levine et al., 1989), less is known about the accuracy of citizen-scientists, seasonal field technicians, or professionals from other disciplines that may possess limited training and experience using the texture-by-feel method. Furthermore, there has been no comprehensive study on the relationship between laboratory-derived and field

soil texture using the extensive (over 300,000 samples) U.S. National Cooperative Soil Survey soil characterization database (NCSS-SCD). We conducted analysis of observers by training level, including academically-trained soil scientists (National Cooperative Soil Survey), moderately-trained seasonal field technicians (U.S. Public Lands Assessment), and minimally-trained citizen scientists (Namibia Rangeland Assessment). Taking advantage of recently generated datasets from the U.S. and Namibia, this study compares accuracy of texture-by-feel field estimates of texture class to laboratory-derived texture class to determine the reliability of texture-by-feel estimates by both professional NCSS soil scientists and seasonal technicians who were not academically trained soil scientists.

MATERIALS AND METHODS

Professional Soil Scientists: NCSS Characterization Dataset

The NCSS soil characterization database (NCSS-SCD) is a comprehensive database of soil information in the United States built and maintained by the Kellogg Soil Survey Laboratory (KSSL). The database contains pedon data from soil characterization analyses completed by the KSSL and cooperating laboratories of soil samples collected from the 1920s to the present. Data for this study were queried from a Microsoft Access database (<http://ncsslabdatamart.sc.egov.usda.gov/>, accessed 1 Jul. 2017). While the NCSS-SCD is considered a global database (Fig. 1a), most pedons are located within the United States.

A subset of 243,898 soil horizon samples was selected from over 364,000 samples associated with 21,000 pedons (Fig. 1a). Primary criteria for the selected set included availability of both field texture-by-feel estimates and lab measured particle-size analysis (PSA) of soil separates (sand, silt, and clay contents) based on KSSL protocols (Soil Survey Staff, 2014a). Texture classes were reduced from 24 detailed USDA classes (which includes subdivisions of sand and clay) to 12 soil texture classes commonly referenced in the USDA soil texture triangle (Soil Science Division Staff, 2017). Soils data were further reduced by excluding duplicate information, samples with negatives values, soil depths greater than two meters, non-mineral soil horizons (e.g., O, R, W, and L), and texture data where the sum of sand, silt, and clay percentages did not equal 100% (c.f. Levi, 2017; Sequeira et al., 2014). Soil Order was queried from the NCSS-SCD using the KSSL correlated taxonomy when available or field-determined taxonomy when correlated taxonomy was unavailable. Historically, the standard methodology for PSA at KSSL includes passing through a 2-mm sieve, removal of plant material, and use of the pipet sedimentation method (Soil Survey Staff, 2014a), although multiple techniques are assumed to have been used because of the age and complexity of the dataset. We further considered changes in texture class boundaries over the historical period of the NCSS-SCD, including the modification of the clay fraction from 5 to 2 μm (Knight, 1938) and the addition of the silt and loamy sand classes in 1951 (Soil Survey Staff, 1955) over the life of the dataset; however, analysis of those early samples did not change our overall results.

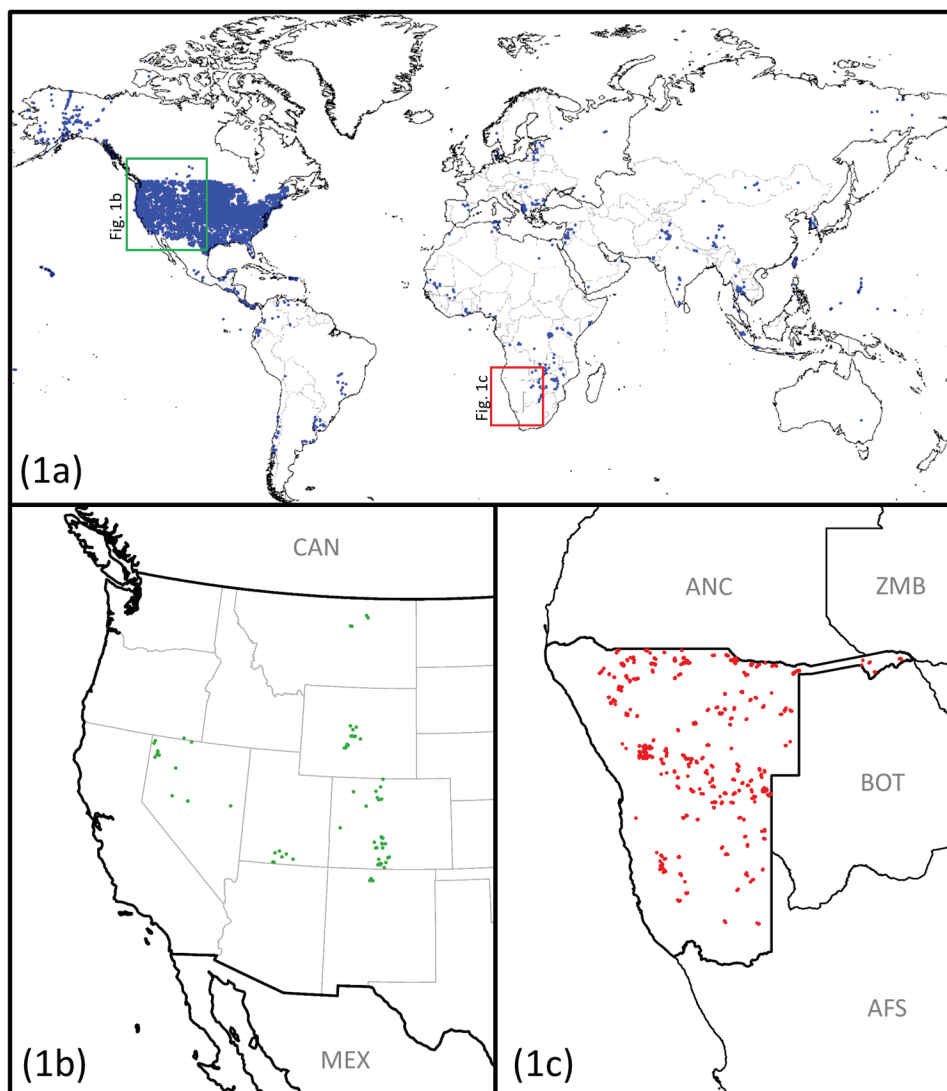


Fig. 1. Spatial distribution of soil samples used in this study, include (a) the National Cooperative Soil Survey–Soil Characterization Database ($n = 243,898$), (b) a select dataset from the U.S. federal lands–assessment, inventory, and monitoring program ($n = 155$), and (c) a citizen science based rangeland assessment study of communal areas in Namibia recorded in the Land Potential Knowledge System ($n = 346$).

Early-Career Seasonal Technicians: U.S. Public Lands Assessment Dataset

Soils were sampled and field-textured across the Western United States from June to October in 2015 and 2016 as part of the U.S. Bureau of Land Management (BLM) assessment, inventory, and monitoring programs (MacKinnon et al., 2011; Taylor et al., 2014). For selected BLM monitoring projects, technicians with no prior soils training received a 1-d soil profile description training texture-by-feel (including instruction following Thien, 1979) using representative soil samples containing a range of soil textures of varying sand, silt, and clay content with known laboratory-determined textures. Two to six weeks after training, these field technicians identified soil horizons and described soil properties for each horizon at locations across the Western United States. Forty individuals completed texture-by-feel at 75 different pedons (Fig. 1b) for a total of 155 soil horizon samples (so that each individual uniquely described approximately 2-5 pedons) from Wyoming, Colorado, Utah, New Mexico, and Nevada. Soil taxonomy was not

determined. Laboratory determinations were completed at the USDA-ARS soil laboratory in Las Cruces, NM using the hydrometer method, with analysis completed in a constant-temperature bath and not pretreated for organics and salts (Bouyoucos, 1962; Soil Survey Staff, 2014a). A few drops of amyl alcohol were added as necessary to limit foaming on high organic matter samples.

Minimally-Trained Citizen Scientists: Namibia Rangeland Assessment Dataset

Seasonal field data collectors determined soil texture-by-feel as part of a rangeland assessment study of communal areas in Namibia and recorded in the citizen science-Land Potential Knowledge System (LandPKS) Data Portal (Herrick et al., 2013; LandPotential.org, 2018). Texture was estimated using the field texture-by-feel method as part of a rangeland assessment study of communal areas in Namibia in August 2014 using the LandPKS mobile application and methodology (Herrick et al., 2017). Training included an approximately 1.5-h field texture-by-feel instruction led by the second author using samples from the area that had not been lab tested. Data collectors were encouraged to compare their estimates and refer to the USDA ribbon

chart (similar to Thien, 1979) throughout the collection period, but no additional formal training was provided. Field estimates were completed for five depths between August and September in 2014. Samples were collected at the 1- to 10- and 50- to 70-cm depths only, except where soil was <50 cm deep. This resulted in a total of 346 soil samples from 175 pedons (Fig. 1c). Soil class and taxonomy were not collected. Particle-size analysis was performed on the soil samples using the pipet methodology with no pretreatment for organics and salts (Day, 1965; Miller and Miller, 1987) by the Analytical Laboratory Services in Windhoek, Namibia. Two of the three Namibian field scientists had received no prior soils training, while one received a brief introduction to soils in an undergraduate college agronomy course.

Classification Assessment and Statistical Approach

Confusion matrix classification accuracy analysis was completed using the *caret* (Kuhn, 2008) and the *psych* (Revelle, 2011)

packages in R version 3.4 (R Core Team, 2016) to validate texture class agreement. The confusion matrix is a cross-tabulation of observed and predicted classes (Congalton and Green, 2008) from which overall accuracy, producer's accuracy, user's accuracy, and Kappa index are generated.

Overall accuracy (OA) represents the absolute number of texture-by-feel observations that match texture classes according to PSA-derived methods. User's accuracy (UA) shows the proportion of texture-by-feel classes that match a given PSA-derived class relative to the total number of estimated points of that texture class (error of commission). Producer's accuracy (PA) is a measure of the proportion of PSA-derived soil texture classes correctly classified by the user relative to the total number of observed points within each PSA-derived texture class (error of omission). These indices were calculated as follows:

$$OA = \frac{\sum_{i=1}^r E_{ii}}{N} \quad [1]$$

$$UA = \frac{X_{ii}}{\sum_{i=1}^r X_{ij}} \quad [2]$$

$$PA = \frac{X_{jj}}{\sum_{i=1}^r X_{ij}} \quad [3]$$

where r is the number of texture classes, E_{ii} is the sum of diagonal elements, N is the number of observations, X_{ii} is the diagonal value for each class in one row, X_{ij} is the sum of values in one row or column, and X_{jj} is the diagonal value for each class in one column. Both UA and PA were not reported for the Namibia and U.S. technician data sets as interpretation of UA and PA are often unreliable with smaller sample sizes and unbalanced class distribution (Congalton and Green, 2008). Kappa (κ) was calculated to account for unbalanced sample class distribution and measures classification accuracy after accounting for probability of chance agreement among the 12 texture classes. The kappa index is calculated with the number of texture classes, number of correctly classified samples, and the total number of classes (Congalton and Green, 2008) by:

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad [4]$$

where P_o is the proportion of correctly classified samples, and P_e is the probability of random agreement. Kappa results can range from -1 to 1, with values <0.0 showing poor agreement, values between 0.0 and 0.2 represent slight agreement, values between 0.2 and 0.4 represent fair agreement, values between 0.4 and 0.6 represent moderate agreement, values between 0.6 and 0.8 represent substantial agreement, and values between 0.8 and 1.0 indicate almost perfect agreement (Landis and Koch, 1977). Lower and upper Kappa (κ_L and κ_U) confidence intervals were defined at $\alpha = 0.05$ (Revelle, 2011). We note that this classification of Kappa values is based on assumptions we are comparing to an absolute reference. As this is not the case for our study (Box 1), the results represent conservative estimates of accuracy.

Results were further evaluated based on near-class accuracy where adjacent texture classes were combined for accuracy of any given class. Because texture class is categorical, we developed a weighted matrix similar to Cohen (1968) based on each adjacent proximity classes of the USDA texture triangle, to find the adjacent-overall accuracy so that:

$$OA_{adj} = \frac{\sum_{i=1}^r E_{ij}}{N} \quad [6]$$

where r is the number of texture classes, E_{ij} is the sum of correct near class elements and N is the total number of observations. Weighted Kappa values (κ_w , κ_{wL} , and κ_{wU}) were similarly calculated using an error matrix where adjacent class estimates were given the same weight as correct classes (Cohen, 1968). As an example, an estimated texture-by-feel of Sandy Loam would be grouped with Sandy Clay Loam, Loam, and Loamy Sand and then compared against Sandy Loam PSA-derived class. For some texture classes, specifically the Loam and Clay Loam classes, our adjacent-overall accuracy analysis expands a correct result to larger portion of the texture triangle compared with other texture classes. While not considered in this study, including clay percentage estimates combined with the *modified-centroid* method (Levi, 2017) would help reduce this disparity.

To explore hand-texturing error among datasets, we further evaluated results by reviewing the texture-by-feel methodology. The technicians (both from Namibia and the U.S.) were trained using Thien's flow diagram (Thein, 1979) which begins with manipulating the soil sample into a cast (ball), followed by a ribbon, and concluded with identifying grittiness. Feedback from public land assessment trainings included comments that novice observers struggle with assessing grittiness level as well as ribbon length. We calculated the conditional probability on soil samples where the probability of an incorrect grittiness class given a correct ribbon length class is:

$$P(\text{Incorrect Grittiness}|\text{Correct ribbon}) = \frac{P(\text{wrong grittiness and correct ribbon})}{N} \quad [7]$$

where N is total number of observations not including the Sand, Loamy Sand, or Silt texture classes. We also calculated the conditional probability on soil samples where the probability of an incorrect ribbon given a correct grittiness class is:

$$P(\text{Incorrect Ribbon}|\text{Correct Grittiness}) = \frac{P(\text{wrong ribbon and correct grittiness})}{N} \quad [8]$$

where N is total number of observations not including the Sand, Loamy Sand, or Silt texture classes.

RESULTS AND DISCUSSION

Summary of Texture Class Accuracy

The overall (absolute) accuracy (OA) of professional soil scientists from the NCSS-SCD database was 66% ($\kappa = 0.61$, $\kappa_L = 0.61$, $\kappa_U = 0.62$) and increased to 91% ($\kappa_w = 0.82$, $\kappa_{wL} = 0.81$, κ_{wU}

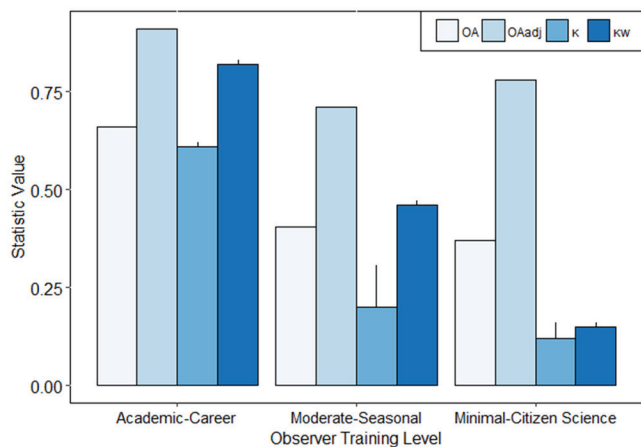


Fig. 2. Accuracy metrics for soil texture-by-feel compared with laboratory-derived particle-size analysis. Analysis of observers by training level include academically-trained career soil scientists (National Cooperative Soil Survey), moderately-trained seasonal field technicians (U.S. Public Lands Assessment, Inventory, and Monitoring program), and minimally-trained citizen scientists (Namibia Rangeland Assessment). OA is the overall accuracy, OAadj is the overall accuracy when combining adjacent texture classes for from the USDA texture triangle (see text), κ is kappa which takes into account of unbalanced class size, and κ_w is the weighted kappa with similar weight as the OAadj (see text). Bars represent upper and lower estimates ($\alpha = 0.05$).

= 0.83) when including adjacent texture-by-feel classes (Fig. 2 and 3; Table 3). Overall accuracy of the Western U.S. dataset was 31.8% ($\kappa = 0.23$, $\kappa_L = 0.18$, $\kappa_U = 0.27$), and the OA for the Namibia dataset was 40.5% ($\kappa = 0.12$, $\kappa_L = 0.08$, $\kappa_U = 0.16$). Only six texture classes were identified by laboratory analysis on the Namibia dataset, thus inflating the OA percentage. Including adjacent texture-by-feel classes, the public land assessment datasets results increased to 78% ($\kappa_w = 0.42$, $\kappa_{wL} = 0.35$, $\kappa_{wU} = 0.49$) for the U.S. set and 72.5% ($\kappa_w = 0.07$, $\kappa_{wL} = 0.17$, $\kappa_{wU} = 0.28$) for the Namibia set (Fig. 2 and 4; Table 4). As explained in Box 1, the maximum possible absolute accuracy is unknown and may be much <100%.

The overall absolute accuracy for professional soil scientists (NCSS-SCD) was highest where one particle-size fraction domi-

Table 4. Confusion matrix for texture-by-feel validation of early career-seasonal field scientists sampling in the Western U.S. public land assessments and citizen scientists sampling in Namibia communal areas.

Dataset	Hand Texture†	Laboratory Texture †								
		S	LS	SL	SCL	L	SiL	CL	SiCL	C
Western U.S.	S	3	1	1	0	0	0	0	0	0
	LS	2	4	5	0	0	0	0	0	0
	SL	2	4	21	0	9	1	0	0	0
	SC	0	0	1	2	4	1	0	0	0
	SCL	0	0	6	1	6	1	0	0	0
	L	0	0	2	0	2	7	0	0	0
	SiL	0	0	0	0	0	8	2	1	0
	CL	0	0	1	1	9	5	2	1	0
	SiCL	0	0	0	0	1	3	1	4	6
	SIC	0	0	1	0	4	3	0	1	0
C	0	0	0	0	5	2	1	2	4	
		OA‡ = 31.8%, κ‡ = 0.23								
Namibia	S	120	0	0	0	0	0	-	-	-
	LS	84	7	3	0	1	0	-	-	-
	SL	48	6	8	0	0	2	-	-	-
	SCL	4	1	4	0	0	0	-	-	-
	L	0	0	0	0	0	0	-	-	-
	SiL	8	6	1	0	0	1	-	-	-
	CL	3	0	8	1	0	2	-	-	-
	SiCL	1	2	8	0	0	4	-	-	-
	SIC	0	0	1	0	0	1	-	-	-
	C	0	0	1	0	0	0	-	-	-
		OA‡ = 40.5%, κ‡ = 0.12								

† S, sand; LS, loamy sand; SL, sandy loam; SC, sandy clay; SCL, sandy clay loam; L, loam; SiL, silt loam; Si, silt; CL, clay loam, SiCL, silty clay loam; SiC, silty clay; C, clay; Si, silt.

‡ OA equals percent correctly classified, and k equals kappa statistic.

nated the sample, such as Clay (74%) and Sand (73%). A similar pattern was seen with UA of Silt (82%), Clay (74%), and Sand (78%). Silt texture class were estimated very poorly (PA = 16%) where 77% of the time Silt PSA was estimated as Silt Loam, with a

Table 3. Confusion matrix for texture-by-feel validation with the NCSS-SCD ($n = 233,808$).

Hand texture†	Laboratory texture†												UA%‡
	S	LS	SL	SC	SCL	L	SiL	CL	SiCL	SIC	C	SI	
S	9746	1405	1093	23	25	60	76	8	5	4	7	6	78
LS	1979	3136	1391	1	44	62	59	9	5	0	5	1	48
SL	1322	2409	18108	29	656	2221	1301	101	67	11	59	39	69
SC	2	3	45	270	236	24	2	67	16	53	114	0	33
SCL	26	79	2191	225	3801	825	80	365	95	30	149	0	50
L	110	187	4757	22	1004	17167	3023	1823	360	72	199	11	60
SiL	49	102	1026	2	96	3986	41652	1165	4118	422	251	1094	77
CL	18	27	668	117	1406	3790	802	10812	1275	411	1383	1	53
SiCL	42	41	174	10	130	1427	4020	2668	21232	2800	1252	16	62
SIC	4	1	24	9	48	105	217	595	2000	8787	2423	3	61
C	9	6	88	232	378	227	118	1663	963	2303	16881	7	74
SI	0	0	1	0	0	3	35	0	4	3	0	187	82
PA%‡	73	42	61	28	49	58	81	57	70	59	74	16	OA‡ = 66.3% κ‡ = 0.61

† S, sand; LS, loamy sand; SL, sandy loam; SC, sandy clay; SCL, sandy clay loam; L, loam; SiL, silt loam; Si, silt; CL, clay loam; SiCL, silty clay loam; SiC, silty clay; C, clay; Si, silt.

‡ UA is user's accuracy, PA is producer's accuracy, OA is percent correctly classified, and κ equals kappa statistic.

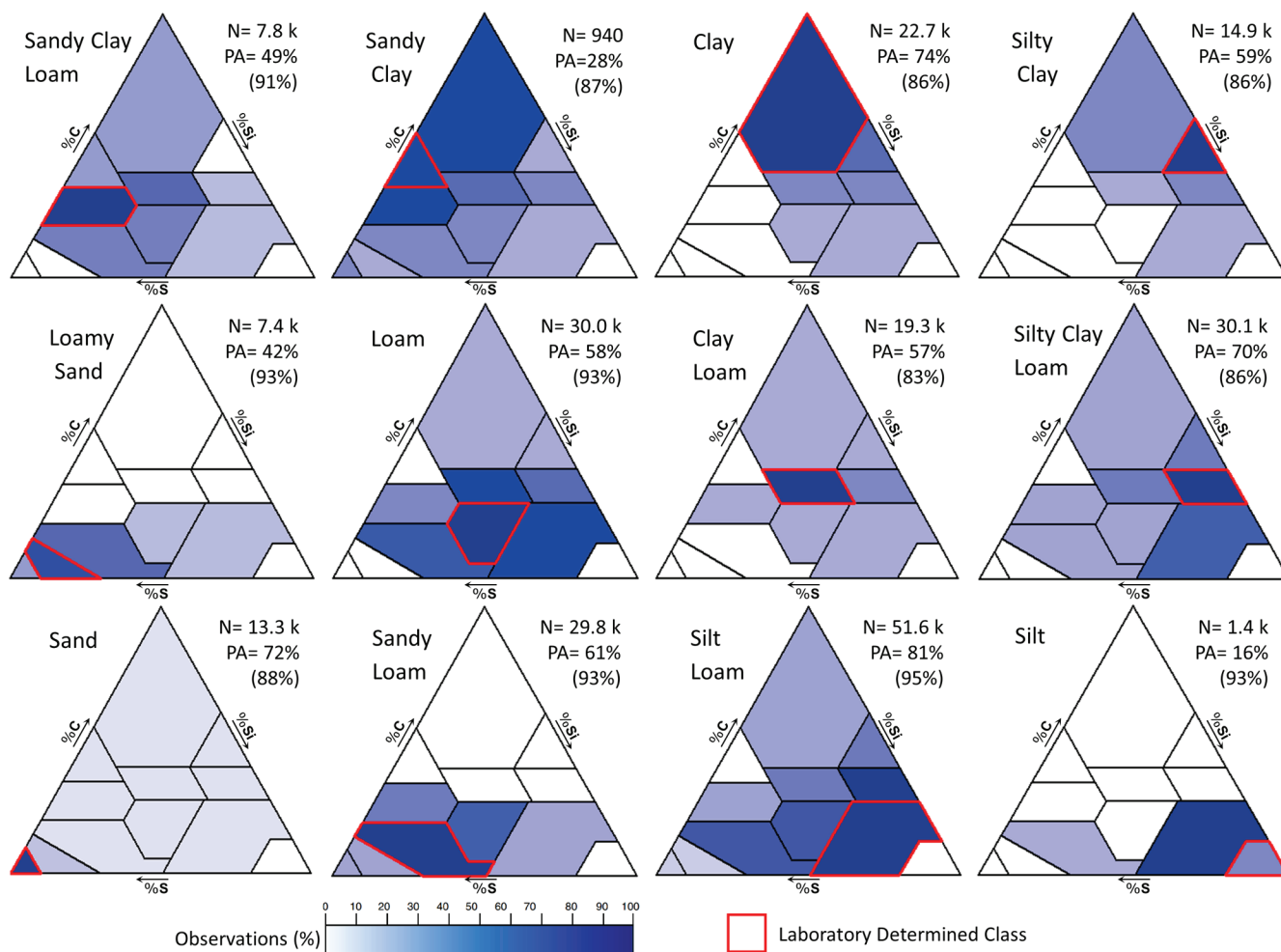


Fig. 3. Classification agreements for individual texture classes for estimates by professional soil scientists illustrating the texture similarity in misclassified observations. Red outline represents the laboratory validation texture group. Plots include samples from National Cooperative Soil Survey. For each texture class, N is the sample size, PA is the producer accuracy, and value in parenthesis is the producer's accuracy-adjacent value.

near-class accuracy of 93% for Silt. Sandy Clay also had low agreement of 28% with similar agreement for Sandy Clay Loam and Clay (24 and 25% respectively). Silt and Sandy Clay classes represented the smallest overall number of samples in the database (0.6 and 0.4% respectively), while Silt Loam represented 22% of the total number of NCSS-SCD samples. Loamy Sand also had relatively high confusion, with 42% PA with Sandy Loam and Sand being the most common misclassification (32 and 19%, respectively).

The Namibia and U.S. technician's assessment samples contained relatively low sample sizes within most texture classes—the exception being the Sand texture class from the Namibia set (268 samples which was 85% of total dataset). The unevenness of texture classes reflected monitoring objectives targeting certain land types and not necessarily because of flaws in data collection. Sand showed the highest PA (45%) followed by sandy loam (40%), loamy sand (35%), and clay loam (33%). Loam texture class was very poorly estimated (5% PA) from the 36 Loamy PSA-derived classes, although loam showed a near-class accuracy of 63%. Loamy Sand and Sandy Loam had a UA of 38 and 30%, respectively and 65 and 60% near-class accuracy. Although we note that UA and PA accuracy metrics are difficult to interpret given these smaller samples size.

Analysis and Interpretation of Misclassification

As the soil texture-by-feel methodology used to train technicians and university students in the U.S. has typically been modeled after the Thien (1979) flow diagram (i.e., Moorberg and Crouse, 2017), we tested the probability of observers matching texture class through the correct ribbon length class and grittiness class when an incorrect texture class was estimated (Fig. 5). This analysis was completed for the NCSS-SCD and a combined Namibia and Western U.S. dataset. As this matrix only contains nine texture classes, any results (including PSA and texture-by-feel) of Sand, Loamy Sand, and Silt classes were omitted from the analysis. We found minimal differences between grittiness and ribbon length from the NCSS-SCD database (41 and 39%) and significant difference from the seasonal technician's database (17 and 37%). This suggests that when seasonal field scientists misclassify texture class, it was more likely because of an error in estimating ribbon length than grittiness. Completing the grittiness check prior to estimating ribbon length could improve these results; however, other methods which help refine the clay-texture classes—such as implementing a worm-and-bend method (FAO, 2006; Landon, 1988)—or requesting percentage of clay estimation could also improve results.

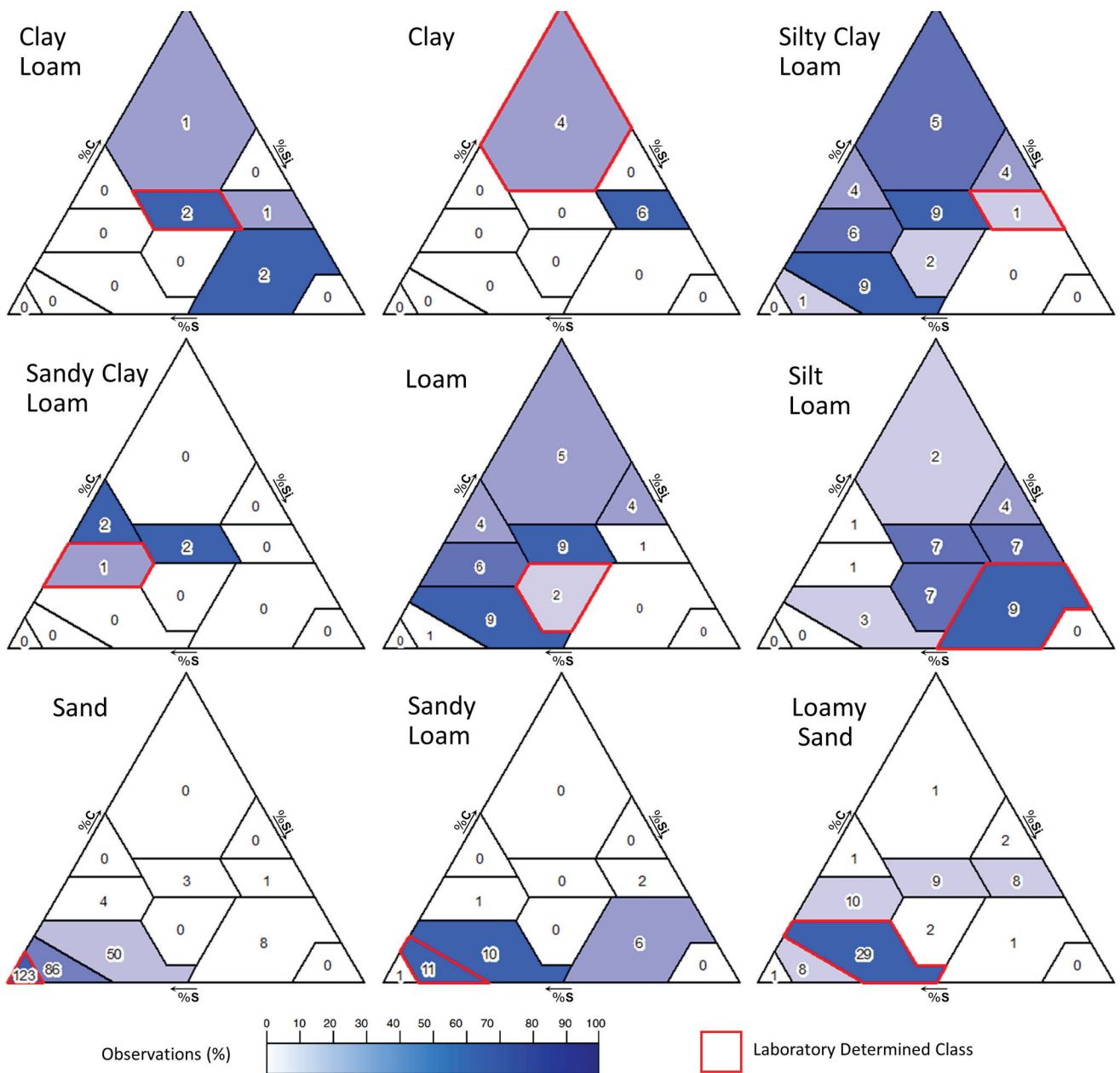


Fig. 4. Classification agreement for individual texture classes for estimates by the Namibia and Western U.S. seasonal field scientists illustrating the texture similarity in miss-classified observations. Red outline represents the laboratory validation texture group.

We also evaluated classification accuracy of soil texture classes of mineral horizons by Soil Order for the professional soil scientist dataset (Table 5). The highest accuracy was found for Alfisols, Entisols, Ultisols, and Vertisols (72, 68, 71, and 70%, respectively) while, Andisols, Aridisols, Gelisols, and Oxisols showed the lowest overall accuracy (44, 46, 34, and 41%). Compared with the overall database, PA was noticeably lower for the Clay texture class from Andisols, Histosols, and Spodosols; the Sand class in Aridisols; the Loam class from Gelisols, Oxisols, and Vertisols; and the Silt Loam from Aridisols, Mollisols, and Vertisols. Overall accuracy was prominently higher for Loamy Sands from Entisols and Ultisols and for the Clay class of Vertisols. The kappa values generally follow overall accuracy for individual Soil Orders (Table 5, Supplemental Tables S1-S12).

Aridisols results confirm previous analysis that soils with excessive salts (such as calcium carbonate, gypsum, or other salts) complicate the soil's feel (Hodgson et al., 1976). This leads to either: underestimation of clay content because they reduce the stickiness of clays and dilute the volume of silicate mineral matter in some cases, or overestimation of clay where carbonate crystals (which are silt- to clay-size) cannot be readily distinguished by feel from clay particles while other salts can disperse clay particles leading to inflated clay estimates. Similarly, soil samples with high organic matter (such as Histosols and some surface horizons) impart a greasy silty feel, reducing coherence of clay and causing an underestimation of clay (Hodgson et al., 1976, Vos et al., 2016). Not properly sorting soil through a 2-mm sieve may also be a source of error because of not removing coarse particles > 2-mm or by not

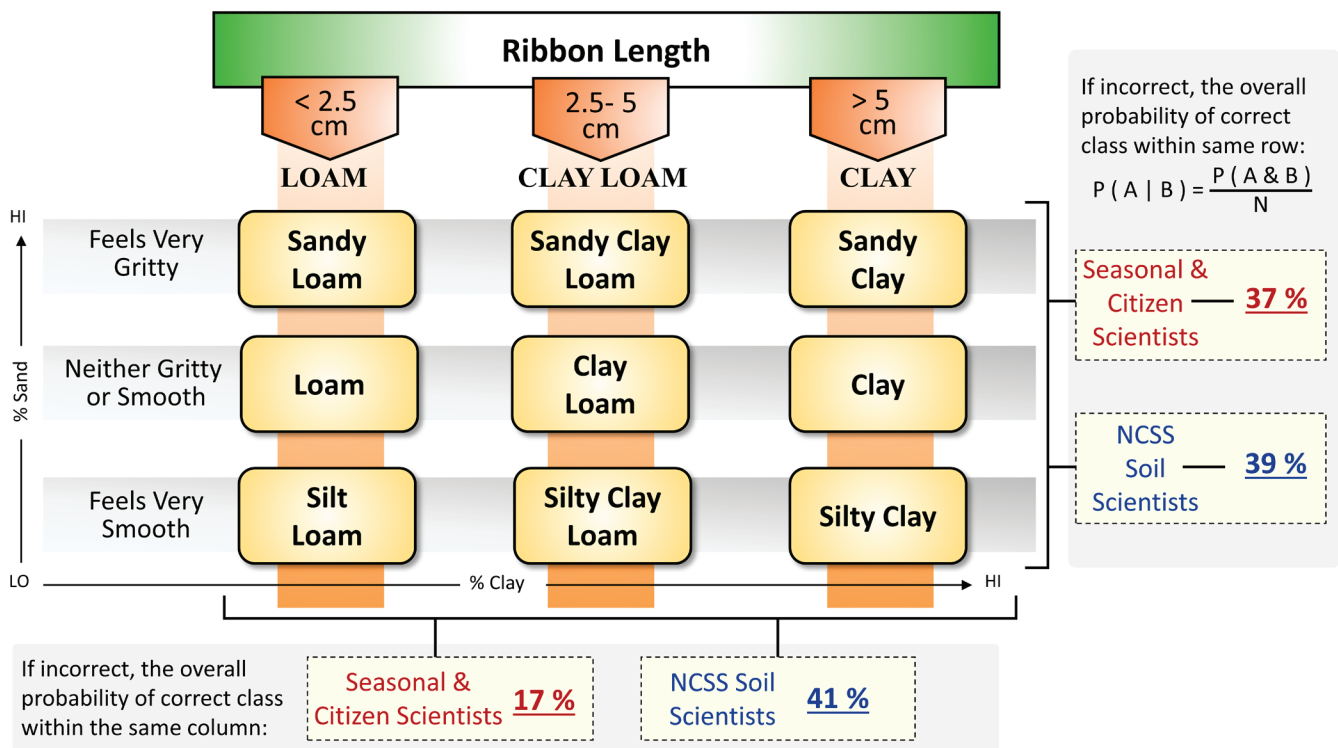


Fig. 5. Probability of either ribbon length or grittiness being correct when texture-by-feel is misclassified. Based on the Thien (1979) flow diagram where soil sample was able to manipulate into a ribbon first and then check grittiness. Samples hand- or lab- textured as Silt, Loamy Sand, and Sand texture classes were omitted from this probability analysis.

properly breaking up clay aggregates which require more vigorous kneading. Often, sticky smectitic clays (which resists deformation and make longer ribbons) can feel different from the less sticky micaceous or kaolinitic clays (which make shorter ribbon).

Lower Kappa values of Oxisols and Andisols highlight a limitation of the dataset as well as issues of texture-by-feel estimates from tropical soils, which comprise 0.5 and 2%, respectively of the NCSS-SCD. Accuracy of these two soil orders is further complicated as even traditional laboratory techniques under estimate clay in the highly weathered oxide-rich tropic and glassy volcanic soils. Clays form stable silt- and sand-sized particles through bonding of oxides and organics, electrostatic attraction of oxides and non-crystalline minerals, or the irreversible drying of amorphous minerals (Silva et al., 2015). Soil samples resist dispersion with standard laboratory techniques as well as prove difficult to break down by hand because of the high degree of micro aggregate stability. Additionally, Andic soils exhibit a smeary feel compared with the more common sticky feel. Even with these difficulties, previous work has shown that field texture-by-feel estimates provide reliable estimates of texture class (Nettleton et al., 1999).

While Vertisols showed relatively high overall accuracy, Kappa values were low because of the Kappa value reflecting an equal probability of each texture class being correct even though Vertisols are mostly clay textures. Errors because of clay mineralogy may be reduced through calibration with locally-sourced soils where mineralogy is relatively uniform, supplemented by an additional (e.g., color- or landscape position-based) indicators of mineralogy (McDonald et al., 1998).

Comparison with Past Studies

Our analysis showed that texture-by-feel estimates by professional soil scientists in the U.S. were improved than those previously reported in the literature (see Tables 2 and 3). This was generally the result of higher texture class PA than from previous studies, including double-digit differences in the Silty- and Loamy-texture classes. The largest proportion of this difference was driven by Silty Clay Loam and Silt Loam texture classes, which comprise a large proportion of samples from the database. One of the largest differences in class agreement was in the Loam class where PA from previously reported accuracy of 26 to 39% (Table 2), while the NCSS-SCD showed a 58% PA. Sandy Clay class accuracy was not previously well reported because of its rarer occurrence in natural systems, as Sandy Clay PSA-derived samples make up only contains 0.4% of the total number of samples in the NCSS-SCD. Silt was also poorly estimated in the NCSS-SCD database, and this may also be a result of its rarity or because it does not appear in the standard Thien (1979) flow chart.

We found that technicians performed similarly to previously reported university student assessments (Levine et al., 1989). Sand texture class agreement was lower than past studies, driven by the overwhelming number of sandy textures in the Namibia dataset. We believe Namibia observers may have tried to reflect real differences in sand content that were insufficient to result in an absolute texture class change as well as a desire to second guess the lack of variability in the rangeland assessment study. This poor agreement suggests that secondary checks could be added to the Thien (1979) flow chart to feel differences between Sand and Loamy Sand texture classes. For example Loamy Sand will usually

Table 5. Summary of the National Cooperative Soil Survey – Soil Characterization Database accuracy of mineral horizon texture-by-feel agreement with laboratory-derived particle-size analysis by Soil Order. Percentages represent user accuracy within each class. Class estimates under 5% of total samples by Soil Order were omitted from table. Detailed confusion matrix for each soil order are provided in supplementary materials, Tables S1-S12

	Soil Texture Class User Accuracy ‡ (%)														
	N†	S	LS	SL	SC	SCL	L	SIL	CL	SICL	SIC	C	Si	OA¶	κ¶
Soil Order §¶															
Alfisols	70237	70	45	65	–	–	61	85	63	74	–	–	–	72	0.67
Andisols	4959	–	–	–	–	13	37	65	13	42	–	27	–	44	0.30
Aridisols	10314	51	34	58	–	–	45	47	41	34	21	48	–	46	0.37
Entisols	13375	82	56	63	–	–	52	78	41	55	53	45	–	68	0.61
Gelisols	393	67	33	33	–	–	21	54	–	35	–	–	–	34	0.15
Histosols	371	75	35	63	–	–	30	72	24	61	47	33	–	54	0.47
Inceptisols	24589	64	38	64	–	44	58	72	42	55	42	56	–	60	0.51
Mollisols	49914	74	37	54	32	36	55	73	52	67	58	55	–	60	0.53
Oxisols	1157	–	–	16	5	6	2	42	10	–	–	58	15	41	0.13
Spodosols	5526	82	32	70	–	17	33	–	–	–	–	27	–	62	0.50
Ultisols	26470	67	54	65	–	–	61	88	56	71	55	75	39	71	0.66
Vertisols	4596	–	–	–	–	32	15	48	28	49	46	87	–	70	0.44
Full Database															
Overall Accuracy	243,898	73	42	61	28	49	58	81	57	70	59	74	16	66	0.61

† Sample size.

‡ S = sand, LS = loamy sand, SL = sandy loam, SC = sandy clay, SCL = sandy clay loam, L = loam, SIL = silt loam, Si = silt, CL = clay loam, SICL = silty clay loam, SIC = silty clay, C = clay, Si = silt.

§ OA = percent correctly classified, and κ = kappa statistic.

¶ USDA soil taxonomy (Soil Survey Staff, 2014b).

stain fingers because of silt and clay components while Sand will not stain fingers when washed with water. This potential source of limited-range bias merits further investigation. Dominance of one size fraction may facilitate accurate estimates of the Sand and Clay texture classes because they are end members of the texture triangle and checks for these fractions are reliably definitive when a sample is comprised of one size fraction. We believe the Silt class continues to be difficult to identify because inexperience, its rarer occurrence in nature, and the absence of Silt on many texture flow charts.

Implications for Citizen Science and Non-Professional Soil Scientists

Results confirm findings from previous studies showing soil texture-by-feel estimates often differ from laboratory estimates, and that the differences are generally more frequent for individuals with less experience. While the results of the Namibia dataset showed lower agreement than for the professional soil scientists (OA = 41%, κ = 0.12 based on assumption of correct laboratory results, see Box 1), the level of accuracy required depends on objectives. Even the least accurate field estimates may be sufficient for some applications, such as determining whether the soil is too sandy for a rain-fed home garden or for making adobe bricks.

While absolute accuracy was low (31.8%) with only slight agreement (κ = 0.23) for the U.S. public lands inventory and monitoring dataset, accuracy increased to 78% (OA_{adj}) when combining adjacent texture classes, with moderate agreement (κ_w = 0.42). Here, OA_{adj} of texture class is higher than OA demonstrated by professional soil scientists (66%). For inventory and monitoring programs, soil texture is considered important to verify soil properties to match soil class or soil properties (Brungard

et al., 2015; Young et al., 1991) as well as for identification of the ecological site (Duniway et al., 2010). When texture determinations are used for soil identification, often the *relative* texture differences of surface and subsurface layers (not evaluated here) may be more important than absolute accuracy. Additional research is needed to determine the relative and absolute accuracy required for soil classification in different landscapes.

In traditional laboratory and field research, data quality is typically evaluated based on multi-dimensional measures of completeness, consistency, and accuracy (Pipino et al., 2002; Whitney et al., 1998). In addition to these criteria, challenges remain regarding the interpretation and management of data generated by seasonal field technicians, and other individuals, such as citizen scientists, without extensive training and experience (Newman et al., 2012). Determining how to interpret and integrate data from individuals with different levels of experience into natural resource research presents real challenges (Rossiter et al., 2015), and more research is needed on how to treat this disparity in data quality.

IMPROVING FIELD TEXTURE-BY-FEEL ESTIMATES

Where more accurate estimates are required, our analyses support the idea that accuracy can be improved. In this section, we propose three strategies for improving the accuracy of field texture-by-feel determinations: training, calibration, and decision support tools.

Training

Both our review of the literature and our own results support the hypothesis that training increases accuracy. We found accuracy was highest for the professional soil scientists, inter-

mediate for those with a full day of soil profile description (including texture) training, and lowest for those with just 1.5 h of training. Direct comparison of these three datasets is potentially confounded by the fact that the results are based on laboratory determinations from three different laboratories, and that the 1.5-h training group was also working across landscapes with relatively homogenous (sandy) soils, which may have led to a tendency to assign real, but minor differences in texture to different texture classes. Finally, additional training in estimating the percentage of clay would greatly benefit interpretation of soils occurring near boundaries of texture classes.

Calibration

Accuracy of field texture-by-feel estimates is recognized to depend on both the observers' general experience applying the method and their knowledge of how to adjust for locally encountered conditions that may modify determinations (Landon, 1988). Working with several partners' (Current partners include: the Bureau of Land Management [BLM], Natural Resources Conservation Service [NRCS], US Geological Survey [USGS], USDA-ARS, Great Basin Institute, World Agroforestry Centre [ICRAF], Regional Centre for Mapping of Resources for Development [RCMRD], and the Lewa Conservancy in Kenya.), the authors of this study have begun to establish a set of calibration soils throughout the world. We found that building sets of a dozen diverse soils from the local area have been extremely helpful for increasing observer confidence, and accuracy increases as trainees work through the sample set. Availability of these calibration soils has many advantages. First, these soils reduce reliance on professional soil scientists for training. Second, calibration soils increase accuracy as observers are comparing estimates to laboratory determinations rather than those of another individual. Third, calibrating to known textured soils provide a focal point discussing supplementary cues, and how the relative importance of these cues varies across texture classes and soil mineralogy. Finally, the samples can be made available to individuals who are unable to participate in a formal training.

Decision Support Tools

Despite hand texturing being considered a fundamental skill, there is no generally accepted international field methodology. The *Soil Survey Manual* (Soil Science Division Staff, 2017) simply defines texture classes based on the PSA-derived limits rather than providing a specific field methodology, while the FAO's *Guidelines for Soil Profile Description* (FAO, 2006) present the *worm-and-bend* method following Landon (1988). In practice however, numerous ad hoc field systems are used by organizations involved in soil investigation, and many groups have developed their own idiosyncratic methods. This has led to a general belief that no methodology can be applied globally to all soils and that individual scientists must work out their own ability to reliably estimate texture. As a result, lack of standardized methodology provides challenges for observers who may lack the resources (e.g., training and calibration soils) to develop a personalized soil texture framework.

One group that would benefit from a more robust soil texture decision support framework is citizen scientists. Because many volun-

teer citizen-scientists lack training, calibration, and experience, we suggest further controls and checks could be added to improve their texture-by-feel estimates. For example, if the goal of a citizen scientist goal is to identify a soil unit class (such as a soil series), in many locations an individual would only need to identify relative finer texture with depth to indicate the presence of an argillic horizon, instead of classifying texture class correctly. Furthermore, if the goal is to increase accuracy of texture-by-feel estimates, we propose citizen scientists could be guided with additional optional or redundant soil texture-by-feel methods within a stepwise flow diagram to differentiate between similar classes. We suggest that a suite of tools and technology delivered through mobile apps can generate these iterative manipulative checks to improve texture-by-feel estimates. However, there will probably be no substitution for hand-on training with known textured soils and periodic re-calibration.

CONCLUSIONS

Accurate descriptions of soil properties are important to reliably characterize the nature, properties, dynamics, and function of the soil. Here we have shown that professional soil scientists predict soil texture classes with higher absolute accuracy than previously stated in the literature. Our results demonstrate that seasonal field scientist estimate texture-by-feel with similar accuracy to university students with limited training. These findings underscore the need for practicing hand texture against known samples and calibrating for localized variability in soil texture. We have provided a discussion of options for increasing the accuracy of field soil texture predictions by citizen scientists, including training, calibration, and decision support tools that go beyond simple dichotomous keys. Future work should focus on standardized field techniques as well as additional manipulative checks on Loamy to Silty texture classes. Results support both continued investments in citizen science, and use of tools (such as the Land-Potential Knowledge System) designed to increase observer accuracy with training provided through universally available mobile apps.

ACKNOWLEDGMENTS

This research was based on work supported by the United States Agency for International Development (USAID), USDI-Bureau of Land Management, and the USDA-Agricultural Research Service. We are grateful to the many USDA-NRCS employees who contributed to the Soil Survey database, BLM field staff and seasonal technicians who provided samples for analysis, and to A. Beh for his assistance coordinating the Land-Potential Knowledge System (LandPKS) project. Special thanks to N. Stauffer and S. DiStefano for organizing BLM samples for analysis; to C. Garton for help editing the manuscript; and to two reviewers whose comments helped improve the manuscript. Use of trade, product, industry, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. government.

REFERENCES

- Akamigbo, F. 1984. The accuracy of field textures in a humid tropical environment. *Soil survey and land evaluation* 4: 63-70.
- Bouyoucos, G.J. 1962. Hydrometer method improved for making particle size analyses of soils. *Agron. J.* 54:464-465. doi:10.2134/agronj1962.0002196205400050028x
- Brady, N.C., and R. Weil. 2013. *The Nature and properties of soils*. Pearson New International Edition. Pearson Higher Ed, New York, NY.
- Brungard, C.W., J.L. Boettinger, M.C. Duniway, S.A. Wills, and T.C. Edwards. 2015. Machine learning for predicting soil classes in three semi-arid landscapes.

- Geoderma 239:68–83. doi:10.1016/j.geoderma.2014.09.019
- Cohen, J. 1968. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychol. Bull.* 70:213–220. doi:10.1037/h0026256
- Congalton, R.G. and K. Green. 2008. Assessing the accuracy of remotely sensed data: Principles and practices. CRC Press, Boca Raton, FL.
- David, O.O. 1999. Improvement in field texture accuracy for sustainable agriculture. *J. Sustain. Agric.* 15:61–68. doi:10.1300/J064v15n02_07
- Day, P. R. 1965. Particle Fractionation and Particle-Size Analysis. In: C.A. Black, editor, *Methods of Soil Analysis. Part 1. Physical and Mineralogical Properties, Including Statistics of Measurement and Sampling*, Agron. Monogr. 9.1. ASA and SSSA, Madison, WI, p. 545–567. doi:10.2134/agronmonogr9.1.c43
- Duniway, M.C., B.T. Bestelmeyer, and A. Tugel. 2010. Soil processes and properties that distinguish ecological sites and states. *Rangelands* 32:9–15. doi:10.2111/Rangelands-D-10-00090.1
- Eshel, G., G. Levy, U. Mingelgrin, and M. Singer. 2004. Critical evaluation of the use of laser diffraction for particle-size distribution analysis. *Soil Sci. Soc. Am. J.* 68:736–743. doi:10.2136/sssaj2004.7360
- FAO. 2006. Guidelines for soil description. Food & Agriculture Organization, Rome, Italy.
- Foss, J., W. Wright, and R. Coles. 1975. Testing the accuracy of field textures. *Soil Sci. Soc. Am. J.* 39:800–802. doi:10.2136/sssaj1975.03615995003900040051x
- Franzmeier, D., and P. Owens. 2008. Soil texture estimates: A tool to compare texture-by-feel and lab data. *J. Nat. Resour. Life Sci. Educ.* 37:111–116.
- Gee, G. W., and J. W. Bauder. 1986. Particle-size Analysis. In: A. Klute, editor, *Methods of Soil Analysis: Part 1. SSSA Book Ser. 5.1. SSSA and ASA, Madison, WI*, p. 383–411. doi:10.2136/sssabookser5.1.2ed.c15
- Herrick, J.E., J.W. Karl, S.E. McCord, M. Buenemann, C. Riginos, E. Courtright, J. Van Zee, A. C. Ganguli, J. Angerer, J. R. Brown, D.W. Kimiti, R. Saltzman, A. Beh, and B. Bestelmeyer. 2017. Two new mobile apps for rangeland inventory and monitoring by landowners and land managers. *Rangelands* 39:46–55. doi:10.1016/j.rala.2016.12.003
- Herrick, J.E., K.C. Urama, J.W. Karl, J. Boos, M.-V.V. Johnson, and K.D. Shepherd, J. Hempel, B.T. Bestelmeyer, J. Davies, J. Larson Guerra, C. Kosnik, D.W. Kimiti, A. Losinyen Ekai, K. Muller, L. Norfleet, N. Ozor, T. Reinsch, J. Sarukhan, and Larry T. West. 2013. The global Land-Potential Knowledge System (LandPKS): Supporting evidence-based, site-specific land use and management through cloud computing, mobile applications, and crowdsourcing. *J. Soil Water Conserv.* 68:5A–12A. doi:10.2489/jswc.68.15A
- Hodgson, J., J. Hollis, R. Jones, and R. Palmer. 1976. A comparison of field estimates and laboratory analyses of the silt and clay contents of some West Midland soils. *J. Soil Sci.* 27:411–419. doi:10.1111/j.1365-2389.1976.tb02011.x
- Kettler, T., J.W. Doran, and T. Gilbert. 2001. Simplified method for soil particle-size determination to accompany soil-quality analyses. *Soil Sci. Soc. Am. J.* 65:849–852. doi:10.2136/sssaj2001.653849x
- Knight, H.G. 1938. New size limits for silt and clay. *Soil Sci. Soc. Am. Proc.* 2:592.
- Kosmala, M., A. Wiggins, A. Swanson, and B. Simmons. 2016. Assessing data quality in citizen science. *Front. Ecol. Environ* 14:551–560. doi:10.1002/fee.1436
- Kuhn, M. 2008. Caret package. *J. Stat. Softw.* 28:1–26.
- Landis, J.R. and G.G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics*: 159–174.
- Landon, J. 1988. Towards a standard field assessment of soil texture for mineral soils. *Soil survey and land. Evaluation* 8:161–165.
- LandPotential.org. 2018. Land Potential Knowledge System Data Portal. Available online at <http://landpotential.org/landpks> (accessed 1 Aug. 2018).
- Levi, M.R. 2017. Modified centroid for estimating sand, silt, and clay from soil texture class. *Soil Sci. Soc. Am. J.* 81:578–588. doi:10.2136/sssaj2016.09.0301
- Levine, S., D.F. Post, and T. Ellsworth. 1989. An evaluation of student proficiency in field estimation of soil texture. *J. Agron. Educ.* 18:100–104.
- MacKinnon, W.C., J.W. Karl, G.R. Toews, J.J. Taylor, S. Karl, C.S. Spurrier, and J.E. Herrick. 2011. BLM core terrestrial indicators and methods. U.S. Dep. of the Interior, Bureau of Land Management, National Operations Center Denver, CO.
- McDonald, R.C., R. Isbell, J.G. Speight, J. Walker and M. Hopkins. 1998. Australian soil and land survey: Field handbook. CSIRO publishing, Clayton, Australia.
- McKinley, D.C., A.J. Miller-Rushing, H.L. Ballard, R. Bonney, H. Brown, S.C. Cook-Patton, D.M. Evans, R.A. French, J.K. Parrish, T.B. Phillips, S.F. Ryan, L. A. Shanley, J.L. Shirk, K. F. Stepenuck, J. F. Weltzin, A. Wiggins, O.D. Boyle, R.D. Briggs, S.F. Chapin III, D.A. Hewitt, P.W. Preuss, M.A. Soukup. 2017. Citizen science can improve conservation science, natural resource management, and environmental protection. *Biol. Conserv.* 208:15–28. doi:10.1016/j.biocon.2016.05.015
- Miller, W., and D. Miller. 1987. A micro-pipette method for soil mechanical analysis. *Commun. Soil Sci. Plant Anal.* 18:1–15. doi:10.1080/00103628709367799
- Minasny, B., A.B. McBratney, D.J. Field, G. Tranter, N.J. McKenzie, and D.M. Brough. 2007. Relationships between field texture and particle-size distribution in Australia and their implications. *Soil Res.* 45:428–437. doi:10.1071/SR07051
- Moorberg, C.J. and D.A. Crouse. 2017. An open-source laboratory manual for introductory, undergraduate soil science courses. *Natural Sciences Education* 46: 170013. doi:10.4195/nse2017.06.0013
- Nauman, T.W., and M.C. Duniway. 2016. The automated reference toolset: A soil-geomorphic ecological potential matching algorithm. *Soil Sci. Soc. Am. J.* 80:1317–1328. doi:10.2136/sssaj2016.05.0151
- Nettleton, W., S. Brownfield, R. Burt, E. Benham, S. Baird, K. Hipple, et al. 1999. Reliability of Andisol field texture clay estimates. *Soil Horiz.* 40:36–49. doi:10.2136/sh1999.2.0036
- Newman, G., A. Wiggins, A. Crall, E. Graham, S. Newman, and K. Crowston. 2012. The future of citizen science: Emerging technologies and shifting paradigms. *Front. Ecol. Environ* 10:298–304. doi:10.1890/110294
- Northcote, K.H. 1971. Factual key for the recognition of Australian soils.
- Ogunkunle, A. 1993. Soil in land suitability evaluation: An example with oil palm in Nigeria. *Soil Use Manage.* 9:35–39. doi:10.1111/j.1475-2743.1993.tb00925.x
- Pipino, L.L., Y.W. Lee, and R.Y. Wang. 2002. Data quality assessment. *Commun. ACM* 45:211–218. doi:10.1145/505248.506010
- Post, D.F., A.R. Huete, and D.S. Pease. 1986. A comparison of soil scientist estimations and laboratory determinations of some Arizona soil properties. *J. Soil Water Conserv.* 41:421–424.
- Post, D.F., S.J. Parikh, R.A. Papp, and L. Ferreira. 2006. Evaluating the skill of students to determine soil morphology characteristics. *J. Nat. Resour. Life Sci. Educ.* 35:217–224.
- R Core Team. 2016. R: Language and environment for statistical computing. R Foundation for Statistical Computing, 2005; R Core Team, Vienna, Austria.
- Rawls, W.J., and Y.A. Pachepsky. 2002. Using field topographic descriptors to estimate soil water retention. *Soil Sci.* 167:423–435. doi:10.1097/00010694-200207000-00001
- Revelle, W. 2011. *psych*: Procedures for Psychological, Psychometric, and Personality Research. Version 1.7.5. Northwestern Univ., Evanston, IL. Available from <https://cran.r-project.org/package=psych>.
- Riesch, H., and C. Potter. 2014. Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions. *Public Underst. Sci.* 23:107–120. doi:10.1177/0963662513497324
- Rossiter, D.G., J. Liu, S. Carlisle, and A.-X. Zhu. 2015. Can citizen science assist digital soil mapping? *Geoderma* 259:71–80. doi:10.1016/j.geoderma.2015.05.006
- Rowell, D.L. 2014. *Soil science: Methods & applications*. Routledge.
- Sequeira, C.H., S.A. Wills, C.A. Seybold, and L.T. West. 2014. Predicting soil bulk density for incomplete databases. *Geoderma* 213:64–73. doi:10.1016/j.geoderma.2013.07.013
- Silva, J.H., J.L. Deenik, R.S. Yost, G.L. Bruland, and S.E. Crow. 2015. Improving clay content measurement in oxidic and volcanic ash soils of Hawaii by increasing dispersant concentration and ultrasonic energy levels. *Geoderma* 237:211–223. doi:10.1016/j.geoderma.2014.09.008
- Soil Science Division Staff. 2017. *Soil survey manual*. Gov. Print. Office, Washington, D.C.
- Soil Survey Staff. 2014a. Kellogg soil survey laboratory methods manual. Natural Resources Conservation Services. National Soil Survey Center, Lincoln, NE.
- Soil Survey Staff. 2014b. *Keys to soil taxonomy*. 12 ed. USDA-NRCS, Washington, DC.
- Soil Survey Staff. 1955. *Soil survey manual*. USDA handb. 18. USDA Soil Conservation Service. U.S. Gov. Print. Office, Washington, DC.
- Taylor, J., E. Kachergis, G. Toews, J. Karl, M. Bobo, M. Karl, S. Miller, and C.S. Spurrier. 2014. AIM-monitoring: A component of the BLM assessment, inventory, and monitoring strategy. Tech. Note 445. U.S. Dep. of Interior, Bureau of Land Management, National Operations Center, Denver, CO.
- Thien, S.J. 1979. A flow diagram for teaching texture-by-feel analysis. *J. Agron. Edu.* 8:54–55.
- Vos, C., A. Don, R. Priezt, A. Heidkamp, and A. Freibauer. 2016. Field-based soil-texture estimates could replace laboratory analysis. *Geoderma* 267:215–219. doi:10.1016/j.geoderma.2015.12.022
- Whitney, C.W., B.K. Lind, and P.W. Wahl. 1998. Quality assurance and quality control in longitudinal studies. *Epidemiol. Rev.* 20:71–80. doi:10.1093/oxfordjournals.epirev.a017973
- Young, F.J., J. Maatta and R.D. Hammer. 1991. Confidence intervals for soil properties within map units. *Spatial Variabilities of Soils and Landforms*: 213–229.
- Zobeck, T.M. 2004. Rapid soil particle size analyses using laser diffraction. *Appl. Eng. Agric.* 20:633–639. doi:10.13031/2013.17466