Environmental Science & Technology

Article

pubs.acs.org/est

Novel Proximal Sensing for Monitoring Soil Organic C Stocks and Condition

Raphael A. Viscarra Rossel,*[,†] Craig R. Lobsey,[†] Chris Sharman,[‡] Paul Flick,[§] and Gordon McLachlan[∥]

[†]CSIRO Land and Water, P.O. BOX 1700, Canberra, ACT 2601, Australia

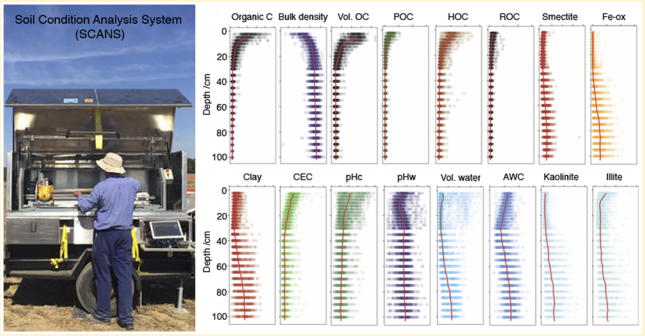
[‡]CSIRO Data61, Private bag 12, Hobart, TAS 7001, Australia

[§]CSIRO Data61, P.O. BOX 883, Kenmore, QLD 4069, Australia

[∥]CSIRO Agriculture, P.O. BOX 1700, Canberra, ACT 2601, Australia

Supporting Information

ABSTRACT: Soil information is needed for environmental monitoring to address current concerns over food, water and energy securities, land degradation, and climate change. We developed the Soil Condition ANalysis System (SCANS) to help address these needs. It integrates an automated soil core sensing system (CSS) with statistical analytics and modeling to characterize soil at fine depth resolutions and across landscapes. The CSS's sensors include a γ-ray attenuation densitometer to measure bulk density, digital cameras to image the measured soil, and a visible−near-infrared (vis−NIR) spectrometer to measure iron oxides and clay mineralogy. The spectra are also modeled to estimate total soil organic carbon (C), particulate, humus, and resistant organic C (POC, HOC, and ROC, respectively), clay content, cation exchange capacity (CEC), pH, volumetric water content, available water capacity (AWC), and their uncertainties. Measurements of bulk density and organic C are combined to estimate C stocks. Kalman smoothing is used to derive complete soil property profiles with propagated uncertainties. The SCANS provides rapid, precise, quantitative, and spatially explicit information about the properties of soil profiles with a level of detail that is difficult to obtain with other approaches. The information gained effectively deepens our understanding of soil and calls attention to the central role soil plays in our environment.



INTRODUCTION

To improve our current understanding of soil and its role in terrestrial ecosystems, we need new methods for assessing and monitoring soil properties, e.g., organic C content and composition, pH, and water content. The new methods should produce quantitative soil information to depth and across landscapes, to for example, help mitigate anthropogenic greenhouse gas emissions, climate change, and land degradation, improve soil condition, and improve food, water and energy securities, which are important for human well-being and economic development.[1]

Pedology and the use of laboratory-based soil analytical methods to measure the physical, biochemical, and mineralogical properties of soil have so far served us well, but they have not changed much in the past century.[2] Largely, laboratory methods have not kept up with the growing need for good quality, quantitative, inexpensive, spatial and temporal soil information to allow better soil and environmental use and management. Many of these conventional methods are constraining our further understanding of soil and its role in our environments because they are time-consuming and expensive and their measures do not truly represent soil at field condition; they require soil samples to be dried, crushed, ground, subsampled, prepared, and analyzed using often complex procedures. Furthermore, there is additional cost and effort needed to measure soil attributes at depth.

We need measurements of deeper soil to better understand the characteristics of soil profiles and their attributes, conditions, and functions. Measurements of soil at depth help us to, for example, quantify the soil organic C stocks, measure the available water capacity of agricultural soil, the concentration of nutrients in the root zone, assess constraints to root growth, quantify subsoil contamination and acidification, and determine the potential for off-site agricultural pollution such as leaching and runoff.

The past decades have witnessed growing interest in the development and use of proximal soil sensors in different applications to overcome the limitations of the conventional approach.[3] Proximal soil sensors can measure soil attributes

Received: February 16, 2017
Revised: April 14, 2017
Accepted: April 17, 2017
Published: April 17, 2017

DOI: 10.1021/acs.est.7b00889
Environ. Sci. Technol. 2017, 51, 5630−5641

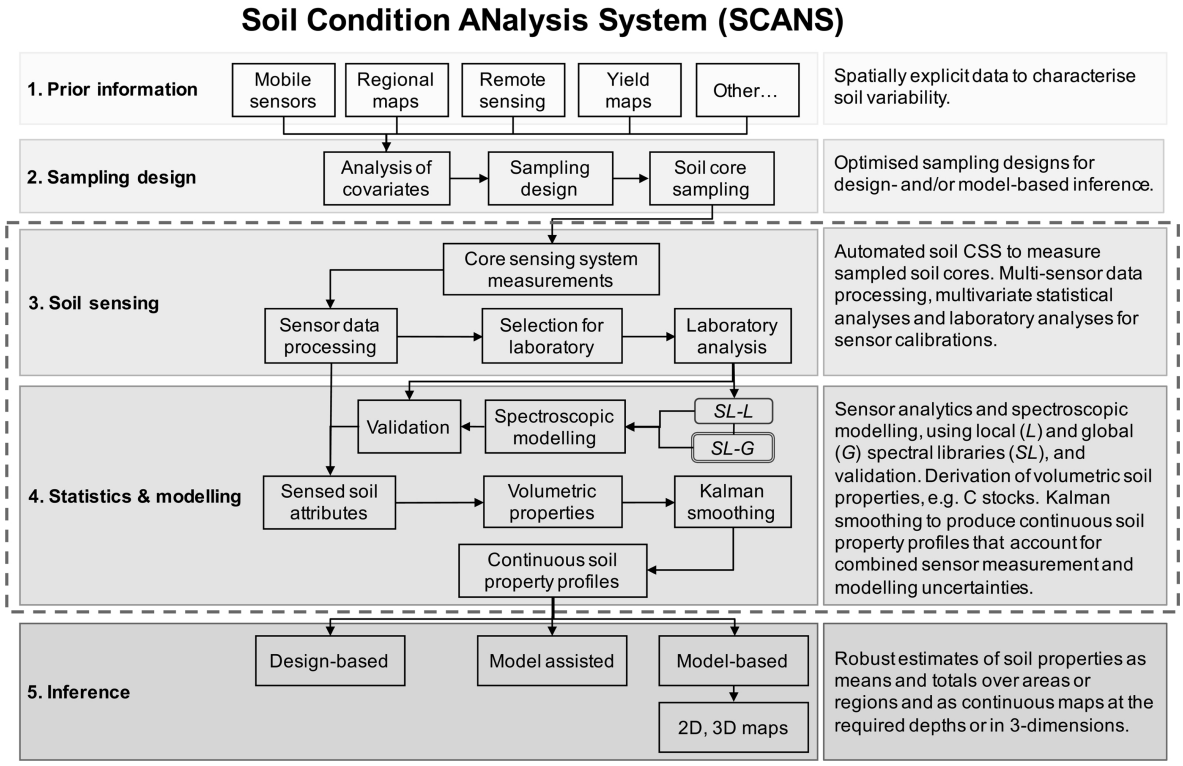


Figure 1. Soil Condition ANalysis System's workflow. The components of the workflow that are depicted within the dashed line are described here. Components that are outside of the broken line were described previously.[9]

rapidly, accurately and more cheaply than conventional laboratory methods. Sensors can also be cheaper to buy and use. They allow many more measurements to be made across the landscape, at depth and in time, so that the data can adequately characterize the spatial (lateral and vertical) and temporal variability of many soil properties. Furthermore, measurements can be made directly in the field so that the data better represent the soil under field conditions.[4]

Proximal soil sensing integrated with robust mathematical and statistical methods, and continued improvements in computing, is modernizing pedology and the methods for soil analyses.[5,6] Importantly, integrated sensing systems are providing new tools to deepen our understanding of soil and to help address our current environmental concerns. We propose that integrated, rapid soil core measurement systems combined with statistically sound analytics and modeling will facilitate the characterization of soil properties and their depth profiles. Such systems will allow us to undertake soil surveys in a far more efficient manner and will provide a natural complement to vehicle-mounted sensor systems for surveying soil.[7] For example, they may be used to help monitor soil organic C stocks after changes in land use or management and may form a good base for the development of auditable and verifiable soil C trading methodologies.[8,9] They may also help agronomically to determine subsoil constraints to production, allow site-specific soil management, and increase the accuracy of precision agriculture. Such systems may also facilitate the assessment and monitoring of soil contamination by heavy metals and other pollutants.

We have developed one such integrated system, the Soil Condition ANalysis System (SCANS) (Figure 1).

The SCANS uses a combination of proximal sensing technologies, smart engineering, and mathematics and statistics to characterize soil variation, laterally across landscapes and vertically down the profile (Figure 1). From Figure 1, the SCANS uses (1) prior information that may be gathered from different sources to characterize soil spatial variability. These covariates are used to inform (2) the sampling design and (5) the inference of the soil properties of interest, the latter, of course, only after (3) sensing and (4) analyzing the data and modeling. Thus, once the soil cores are sampled, (3) they are measured with the core sensing system (CSS) (Figure 1). The sensor data are processed, and a selected set of soil core subsamples are analyzed in the laboratory for (4) spectroscopic modeling and validation. The sensed soil attributes are then used to derive continuous soil property profiles with estimates of uncertainty. The soil property profiles can then be used to (5) infer the soil properties across the study area to the required depth, as means and totals and as continuous soil property maps (Figure 1). Viscarra Rossel et al.[9] described the implementation of steps 1, 2, and 5 in Figure 1. Our aims here are to describe the development of the SCANS automated soil CSS, its sensors, and the sensor data analyses and modeling used to derive depth profiles of soil organic C stocks and other soil biochemical, physical, and mineralogical properties (Figure 1, steps 3 and 4, enclosed by the dashed line).

MATERIALS AND METHODS

The CSS automatically measures biochemical, physical, and mineralogical properties of intact soil core samples that are (wet) under field condition. The visible−near-infrared (vis−NIR), active γ-ray attenuation and camera sensors measure the soil cores at user-defined depth intervals. Below we describe the sensing platform, its sensors, and the analyses and modeling of the sensor data to derive measures of soil properties and their depth profiles.

DOI: 10.1021/acs.est.7b00889
Environ. Sci. Technol. 2017, 51, 5630−5641

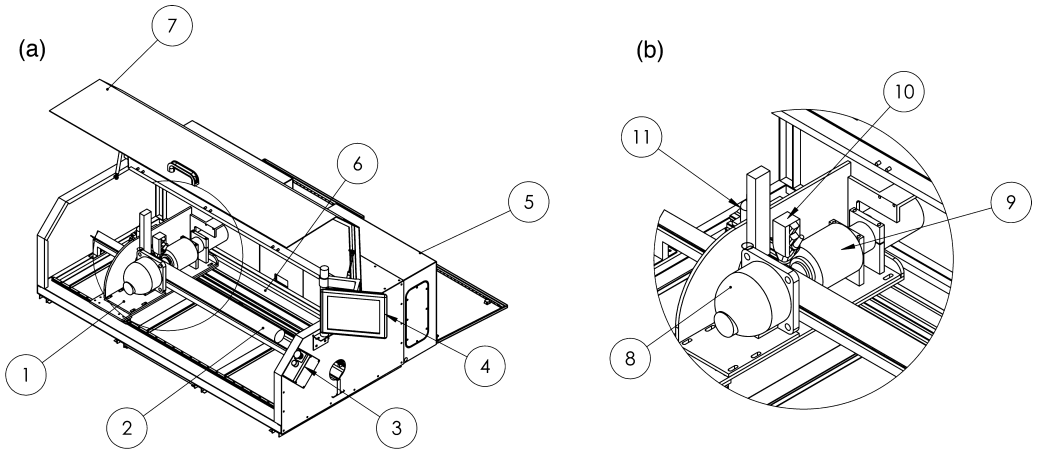


Figure 2. Schematic of the SCANS core sensing system: (a) (1) sensor head, (2) soil core, (3) emergency stop and reset buttons, (4) touch screen PC, (5) electronics boxes, (6) linear actuator, and (7) polycarbonate hood with safety sensors and (b) (8) γ-ray source, (9) γ-ray detector, (10) spectrometer contact probe attachment, and (11) cameras.

Figure 3. SCANS field-deployable core sensing system (CSS): (a) the SCANS sensing systems at a field site, (b) an operator loading a soil core on the CSS, and (c) the core sensing system.

Soil Core Sensing System. The multisensor system was developed to measure 1.2 m long soil core samples with diameters between 45 and 85 mm. Measurements of longer soil cores are possible with some small modifications. A schematic of the CSS is shown in Figure 2.

The platform's housing contains a support structure to rest the soil core, a linear actuator to move the sensor head along the core, a cable train to allow the movement of cables, and a linear rail to provide additional support to the sensor head (Figure 2), which holds four sensors: a vis−NIR spectrometer, an active γ-ray attenuation sensor, a visible camera, and a Lepton long wave infrared camera (see the Supporting Information). A photoelectric proximity sensor is used to detect the length and position of the core sample on the support platform. The platform includes a safety interlock system for the active γ-ray sensor that is triggered by opening the protective door or an emergency stop button. The interlock creates a safe system state by disabling motion and the sensors. A touch panel computer is used to control the system via software and a graphical user interface (Figure 2). Sensor calibration standards (e.g., for the vis−NIR sensor) are mounted on the core support platform for automated sensor calibration (Figures 2 and 3). In the Supporting Information, we provide details about the sensors used in the platform, its control system, and the measurement process.

Field Deployment. The CSS is installed in an enclosure and trailer for transport and operation in the field (Figure 3). For in-field use, the enclosure holds a generator to provide electrical power and an air compressor for the pneumatic system. The measurements can be controlled remotely from a desktop computer or tablet, and data can be retrieved using an integrated 3G cellular modem. The enclosure has storage compartments to carry ancillary equipment and soil cores that are being returned to the laboratory. The field-deployable CSS is shown in Figure 3.

Study Site and Soil Core Sampling. The site for this study was a 600 ha cattle grazing farm located in Northern New South Wales, Australia (S30.69, E151.48). The soil of the farm was classified predominantly as Kurosol in the Australian Soil Classification.[10] In the World Reference Base (WRB) classification, Kurosols might be Acrisols or Planosols.[11]

We used data from a mobile multisensor platform (with electromagnetic induction and passive γ-radiometric sensors, and a precise global navigation system) to produce spatially explicit covariates that we used to derive a stratified simple random sampling plan and 150 soil core sampling locations spread over the entire farm (see Figure 1, steps 1 and 2). The mobile multisensor platform used and the design of the sampling strategy are similar to those described by Viscarra Rossel et al.[9] We used the CSS to measure all 150 sampled soil cores (as in Figure 1, step 3).

Soil Core Sampling. We sampled the soil cores using a Geoprobe 7822DT core sampling rig and the DT325 sampling system. Cores were sampled directly into clear plastic liners

DOI: 10.1021/acs.est.7b00889
Environ. Sci. Technol. 2017, 51, 5630−5641

with end-caps, and immediately before measurement, we cut a longitudinal section of the liner with a DT325 liner cutter (Geoprobe, Salina, Kansas) to expose the core surface for the vis−NIR measurements and for capturing the images (Figure S2). In the Supporting Information, we also describe the different soil core sampling equipment that we have tested to extract intact soil cores for the sensing.

Sensor Measurements and Data Analytics. Measurements with the CSS can be made at fine depth resolutions, predetermined in software by the user. For our experiments here, the measurement resolution was set to 2.5 cm intervals from the surface to 30 cm and 5 cm intervals from 30 to 100 cm. Thus, there were a total of 26 measurements from each of the sensors (γ-ray densiometer, vis−NIR, and photographs, the latter were combined to produce a complete image of the soil core) per core, and the time taken to measure each core was approximately 15 min (i.e., the sensing at each location on the soil cores took approximately 35 s).

γ-Ray Densitometry. To measure the bulk density of the soil cores that are wet, under field condition, the CSS uses measurements of density from the γ-ray attenuation densitometer and estimates of water content from the vis−NIR spectrometer. The technique is described by Lobsey and Viscarra Rossel,[12] and we only briefly summarize it here. For a soil core that is under field condition, the γ-ray attenuation is a function of its mass and the mass attenuation coefficients of soil and water in the attenuation path. Using the Beer−Lambert law, this can be defined as

$$\frac{I}{I_0} = \exp[-x(\mu_s\rho_s + \mu_w\rho_w\theta)] \quad (1)$$

where I is the incident radiation at the detector, I_0 is the unattenuated radiation emitted from the source (determined using calibration standards), and x is the sample thickness in centimeters. Parameters μ_s and μ_w are in units of square centimeters per gram and represent the mass attenuation coefficients of the soil and water, respectively. Parameter ρ_w is the density of water, which is 1 g cm^{-3}, and θ is the volumetric water content of the soil in cm^{-3} cm^{-3}. The mass attenuation coefficient of soil (μ_s) depends on both the photon energy and its elemental constituents; however, at high photon energies (e.g., 0.662 MeV), the effects of varying soil compositions become negligible.[13] We used values of the soil (0.0770 cm^2 g^{-1}) and water (0.0832 cm^2 g^{-1}) mass attenuation coefficients derived by Lobsey and Viscarra Rossel.[12] At each γ-ray attenuation measurement on the soil core, we determine θ using vis−NIR spectroscopy (described below). Thus, using the sensors, we could calculate the dry bulk density of the soil (ρ_{b_γ}) using

$$\rho_{b_\gamma} = \frac{1}{x\mu_s}\ln\left(\frac{I_0}{I}\right) - \frac{\mu_w}{\mu_s}\rho_w\theta \quad (2)$$

Spectroscopic Modeling and Validations. Spectroscopic vis−NIR models (Figure 1, step 4) were used to simultaneously estimate, at the 26 measurement locations on each of the 150 soil cores, the total organic C content, the C composition characterized by the organic particulate, humus, and resistant organic C fractions (POC, HOC, and ROC, respectively),[14] the clay content, the cation exchange capacity (CEC), the pH measured in a 1:5 soil/water suspension (pH_w), the pH measured in a 1:5 soil/0.01 M $CaCl_2$ suspension (pH_c), the volumetric water content (θ), and the available water capacity (AWC).

We used the machine learning algorithm Cubist[15] for the spectroscopic modeling, but with different data sets for the calibrations, which depended on the availability of reference analytical data from the study site. When reference data from the study site were unavailable, estimates were made with general (or global) spectroscopic models derived using large Australian soil spectral libraries (Table 1). This was the case for pH_c, θ, CEC, and AWC.

The soil water library (Table 1) consists a total of 1429 soil vis−NIR reflectance spectra from 162 sites from across Australia. The vis−NIR spectra of soil samples from three different depths, 0−15, 13−30, and 30−60 cm, were recorded under air-dry conditions and at −1, −5, −10, −30, and −60 kPa using suction plates and −500 and −1500 kPa using pressure chambers.[18] For each sample in the data set, AWC was calculated as the difference in θ at −33 and −1500 kPa. This library was also used to correct for the effects of water on the vis−NIR spectra that were measured under field conditions (see below).

When we had a set of local, reference analytical data from our study site, as was the case for total organic C, POC, HOC, ROC, clay content, and pH_w, the spectroscopic models were developed using data from the ReSampling-Local (RS-LOCAL) algorithm that is described by Lobsey et al.[19] Briefly, RS-LOCAL uses a (small) number of representative spectra with reference analytical data from the local site to select an optimal subset of data from a large soil spectral library, in our case the Australian soil spectral libraries (see Table 1). The RS-LOCAL selection is made by iteratively removing data from the spectral library with the largest validation errors with respect to the local data. The result is a subset of data from the large spectral library, which is "customized" for the local predictions at the study site. Thus, the "customized" subset is combined with the local data for spectroscopic modeling. We refer the reader to Lobsey et al.[19] for a detailed description of the RS-LOCAL algorithm. Below we further describe our implementation.

Using the Kennard−Stone algorithm[20] and the 3900 spectra of the 150 soil cores (150 cores × 26 measurements at different depths in each core), we selected a representative set of 20 corresponding local soil core subsamples, which we could afford to analyze in the laboratory (Figure 1, step 3). The 20 samples were analyzed for total soil organic C by total combustion using a LECO carbon analyzer,[21] particulate, humus, and resistant C using the nuclear magnetic resonance (NMR) method.[14] We had 12 surface (0−30 cm) soil samples from a previous reconnaissance survey at our study site with data on clay content measured with the hydrometer method[22] and pH_w measured in a 1:5 soil/water suspension.[21] Using RS-LOCAL, a subset of the relevant Australian spectral library (i.e., the soil property or C fractions library in Table 1) was selected, and the resulting smaller "customized" library was combined with the small set of local samples from our study site. These combined data were used to develop spectroscopic models with Cubist to

Table 1. Soil Spectral Libraries Used in the Modeling

spectral library	soil properties	N	ref
vis−NIR soil properties	organic C, clay, pH_c, pH_w, CEC	18501	16
vis−NIR C fractions	POC, HOC, ROC	550	17
vis−NIR soil water	θ, AWC	1429	18

DOI: 10.1021/acs.est.7b00889
Environ. Sci. Technol. 2017, 51, 5630−5641

Table 2. Statistical Summary of the Soil Attribute Data Used in Spectroscopic Modeling

soil attribute	N	mean	standard deviation	minimum	median	maximum	skewness
organic C (%)	231	1.54	2.40	0.02	0.59	16.28	2.98
particulate C (%)	233	0.26	0.28	0.007	0.13	1.36	1.80
humus C (%)	233	0.82	0.59	0.13	0.67	3.01	1.29
resistant C (%)	233	0.38	0.31	0.01	0.32	1.64	1.55
clay (%)	329	25.3	19.53	0.20	20.30	96.40	0.99
pH_w	329	6.47	1.28	3.60	6.30	9.60	0.18
pH_c	4980	5.84	1.34	3.58	5.62	9.29	0.23
CEC ($cmol_c\ kg^{-1}$)	9397	17.86	14.47	0.50	13.90	106.50	1.53
θ ($cm^{-3}\ cm^{-3}$)	1429	0.26	0.21	0.00	0.22	0.94	0.63
AWC (%)	162	0.40	0.094	0.19	0.40	0.64	0.07

predict the soil properties. The implementation of Cubist for spectroscopic modeling was described in detail by Viscarra Rossel and Webster.[16]

Some of the soil attributes had strongly positively skewed distributions [skewness of >1 (Table 2)], and to stabilize their variances for spectroscopic modeling, described above, we transformed the data to approximate normal distributions by taking logarithms.

All of the spectroscopic models derived using data from either the Australian spectral library or RS-LOCAL, as described above, were validated using 50-repeat 10-fold cross validation,[23] and the validation statistics reported are the means of the 50 repeats. Additionally, the RS-LOCAL model for soil organic C was validated with an independent test data set that comprised 300 reference data with measurements of soil organic C made with a LECO carbon analyzer. These validation samples were selected using the Duplex algorithm[24] from the 3900 spectra of the 150 soil cores.

When the attributes were transformed to logarithms, the cross validation statistics were reported on that scale. However, the estimates of the independent validation of organic C and the estimates of the soil property profiles that were made with logarithmic models were back-transformed to their original scales by adjusting the variance of the estimates using[25]

$$\hat{y} = 10^{[\hat{\psi} + \frac{1}{N}\sum_{i=1}^{N}(\hat{\psi}_i - \psi_i)^2/2]} \quad (3)$$

where $\hat{y}$ values are the back-transformed data, $\hat{\psi}$ values are the estimates of the soil attributes, ψ values are the observed values, both on the logarithmic scale, and the numerator in the exponent represents the mean square error (MSE) of the model, or in the case of organic C, the MSE of the independent validation data set. We note that in the SCANS workflow (Figure 1), when the residuals, $(\hat{\psi}_i - \psi_i)$, in eq 3 are not normally distributed, the back-transformations are performed using the nonparametric "smearing estimate" proposed by Duan.[26]

To estimate uncertainty in the spectroscopic modeling, we used the method described by Viscarra Rossel.[27] The technique uses random sampling with replacement (i.e., the bootstrap) to produce multiple training data sets for spectroscopic modeling. These models are used to derive different realizations of the soil properties to form cumulative distribution functions for each estimate, from which means, variances, and confidence intervals can be calculated.

The volumetric soil organic C content (SOCv) was calculated by multiplying the organic C content (percent) and bulk density profiles. We also measured soil color using the visible portion of the spectra,[28] and the relative abundances of hematite, goethite, kaolinite, illite, and smectite from the heights and widths of wavelength-specific absorption features on continuum-removed spectra.[5,29]

Correcting the Vis–NIR Spectra for Water Present in the Soil. When measurements are taken in the field, soil vis–NIR spectra are affected by water, primarily at absorption near 1400 and 1900 nm. To allow in-field prediction of soil attributes, other than soil water, from the spectra of the soil cores that are under field condition, the SCANS allows the use of either external parameter orthogonalization (EPO)[30,31] to project spectra orthogonal to variations induced by water, or direct standardization (DS)[32] to transfer the wet field spectra so that it may be predicted with the laboratory-derived spectral libraries. The spectra used to develop the transformation matrices for corrections of water using these methods were derived using a soil water spectral library[18] (Table 1).

In the study presented here, before using the vis–NIR spectra for spectroscopic modeling, we corrected them using EPO. We used DS to remove the effects of water before measuring the Fe and clay mineralogy of the cores. In this case, we used DS because unlike EPO, it does not alter the corrected spectra.[32]

Repeatability of the Core Sensing System Measurements. To test the repeatability of the CSS, we measured a single core from our study site seven times over 2 days. On the first day, the operator took four measurements: the first at 9 a.m. (17 °C), one at 11 a.m. (26 °C), one at 2 p.m. (33 °C), and one at 5 p.m. (36 °C). On the second day, the operator took three measurements: at 10 a.m. (26 °C), at 1 p.m. (35 °C), and finally at 4 p.m. (38 °C).

As described above, the sensor data were used to derive measures of bulk density, organic C, SOCv, POC, HOC, ROC, clay content, CEC, pH_w, pH_c, θ, and AWC, at each of the 26 measurement intervals on the soil core. For each soil property, we measured the mean and standard error of the repeated measurements and graphed our results.

Deriving Continuous Soil Property Depth Profiles. To improve the estimates of the sensed soil properties along the 1 m soil core (i.e., with depth), we used Kalman filtering and smoothing, implemented in the KFAS package[33] in the R software.[34] An advantage of using Kalman smoothing over other techniques for deriving soil property profiles (e.g., spline interpolation) is that it allows the uncertainty in the sensor measurements to be incorporated in the process and thus allows its propagation to the filtered estimates of the soil attribute profiles.

Consider a soil property, x (e.g., soil organic C), along the 1 m soil core. We start our sensor measurements at the surface and then take measurements at defined depth intervals, k. These measurements are imprecise, but we want to maintain a

DOI: 10.1021/acs.est.7b00889
Environ. Sci. Technol. 2017, 51, 5630–5641

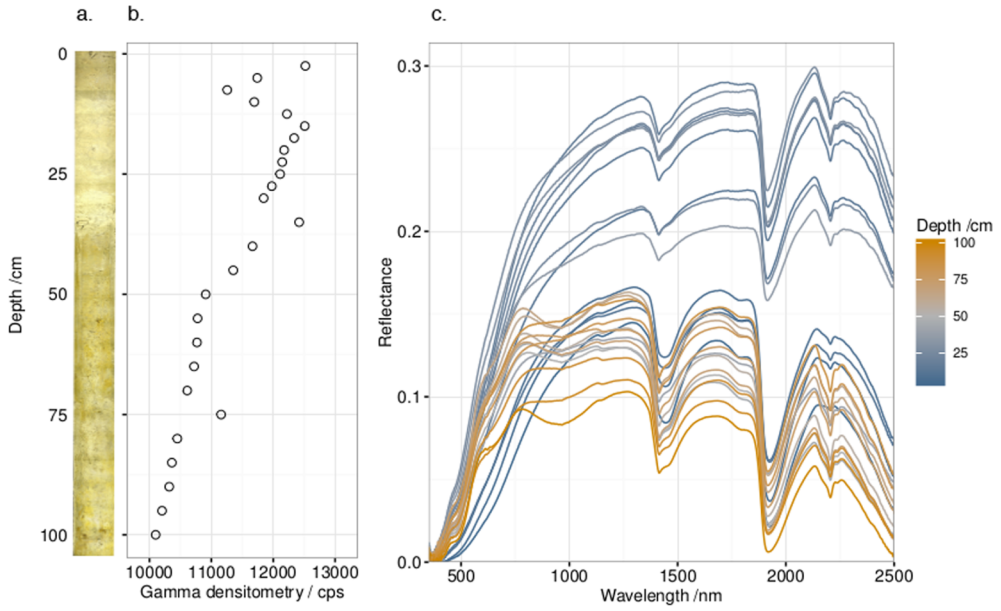


Figure 4. Raw data from the sensors: (a) the image of a soil core showing a bleached horizon around 10−30 cm, (b) decreasing γ-ray counts with depth, indicating the increasing density of the soil with depth, and (c) the vis−NIR spectra of the soil core, showing a fairly abrupt change in the spectra around 30 cm, where absorptions due to iron oxides, clay minerals, and water are more pronounced. This soil is likely to have more clay below 30 cm and a duplex texture profile.

model that defines the magnitude of x along the soil core and its rate of change with depth. Thus, we define the linear state space $\mathbf{x}_k$ by

$$\mathbf{x}_k = \begin{bmatrix} x \\ \dot{x} \end{bmatrix} \tag{4}$$

where $\dot{x}$ is the rate of change of soil property with depth. Note that we have assumed that k is discrete (i.e., fixed measurement intervals), although this can vary depending on the measurement plan defined by the user (see above). Then, the state transition model between the $k-1$ and k depths is defined by

$$\mathbf{x}_k = \begin{bmatrix} 1 & \Delta k \\ 0 & 1 \end{bmatrix} \mathbf{x}_{k-1} + \mathbf{w}_k \tag{5}$$

where $\mathbf{w}_k \sim N(0,\mathbf{Q}_k)$ is the process noise, which we assume to be normally distributed with variance $\mathbf{Q}_k$. It can be fixed or derived using maximum likelihood estimation (MLE) (see ref 33). At each depth, a noisy measurement, $\mathbf{z}_k$, of the true soil property value is made with the particular sensor so that the observation model can be defined by

$$\mathbf{z}_k = [1 \quad 0]\mathbf{x}_k + \mathbf{v}_k \tag{6}$$

where $\mathbf{v}_k \sim N(0,\mathbf{R}_k)$ is the observation noise that we assumed to be normally distributed with variance $\mathbf{R}_k$. $\mathbf{R}_k$ for the spectroscopic estimates of soil properties was taken from the respective spectroscopic model bootstrap variances (see above). For the γ-ray attenuation measurements of bulk density, $\mathbf{R}_k$ was the sum of two components. The first represents the standard error of the γ-ray attenuation measurements of apparent density, which was 0.029 g cm^{-3}, from Lobsey and Viscarra Rossel[12], and the second is the bootstrap variance of the spectroscopic estimates of θ used in the correction of the apparent densities to derive bulk densities (see above and ref 12). We assumed that such errors were independent and not autocorrelated with depth. Thus, the Kalman filter recursively predicts the soil property at the particular depth using only the previously estimated state and the sensor measurement at the current depth and its uncertainty matrix. The Kalman smoother then estimates the state of the system at depths k during a second backward recursion and using the filtered state estimates (see refs 33 and 35).

RESULTS

Sensor Measurements. An example of the data recorded by the SCANS CSS is shown in Figure 4. It shows a true color image of a soil core, which is useful for performing visual inspection and for record keeping, the γ-ray attenuation counts,

Table 3. Spectroscopic Modeling Used To Estimate Soil Properties and Their Performance Statistics[a]

soil attribute	method	n	m	R^2	RMSE
Repeated 10-fold Cross-Validation					
$\log_{10}$ organic C (%)	RS-LOCAL	211	20	0.83	0.39
$\log_{10}$ particulate C (%)	RS-LOCAL	213	20	0.74	0.15
$\log_{10}$ humus C (%)	RS-LOCAL	213	20	0.71	0.35
$\log_{10}$ resistant C (%)	RS-LOCAL	213	20	0.71	0.17
clay (%)	RS-LOCAL	317	12	0.70	10.17
pH_w	RS-LOCAL	317	12	0.71	0.70
pH_c	general	4980		0.76	0.66
$\log_{10}$ CEC ($cmol_c$ kg^{-1})	general	9397		0.70	0.23
θ (cm^{-3} cm^{-3})	general	1429		0.96	0.04
AWC (%)	general	162		0.75	0.048
Independent Validation					
organic C (%)	RS-LOCAL	211	20	0.81	0.41

[a]In the table, n is the number of spectral library samples and m is the number of local samples used with RS-LOCAL. For all of the attributes listed, the coefficient of determination R^2 and the root-mean-square error (RMSE) were computed on predictions made using a repeated 10-fold cross-validation. Additionally for soil organic C, the statistics were also computed using an independent data set of 300 samples.

DOI: 10.1021/acs.est.7b00889
Environ. Sci. Technol. 2017, 51, 5630−5641

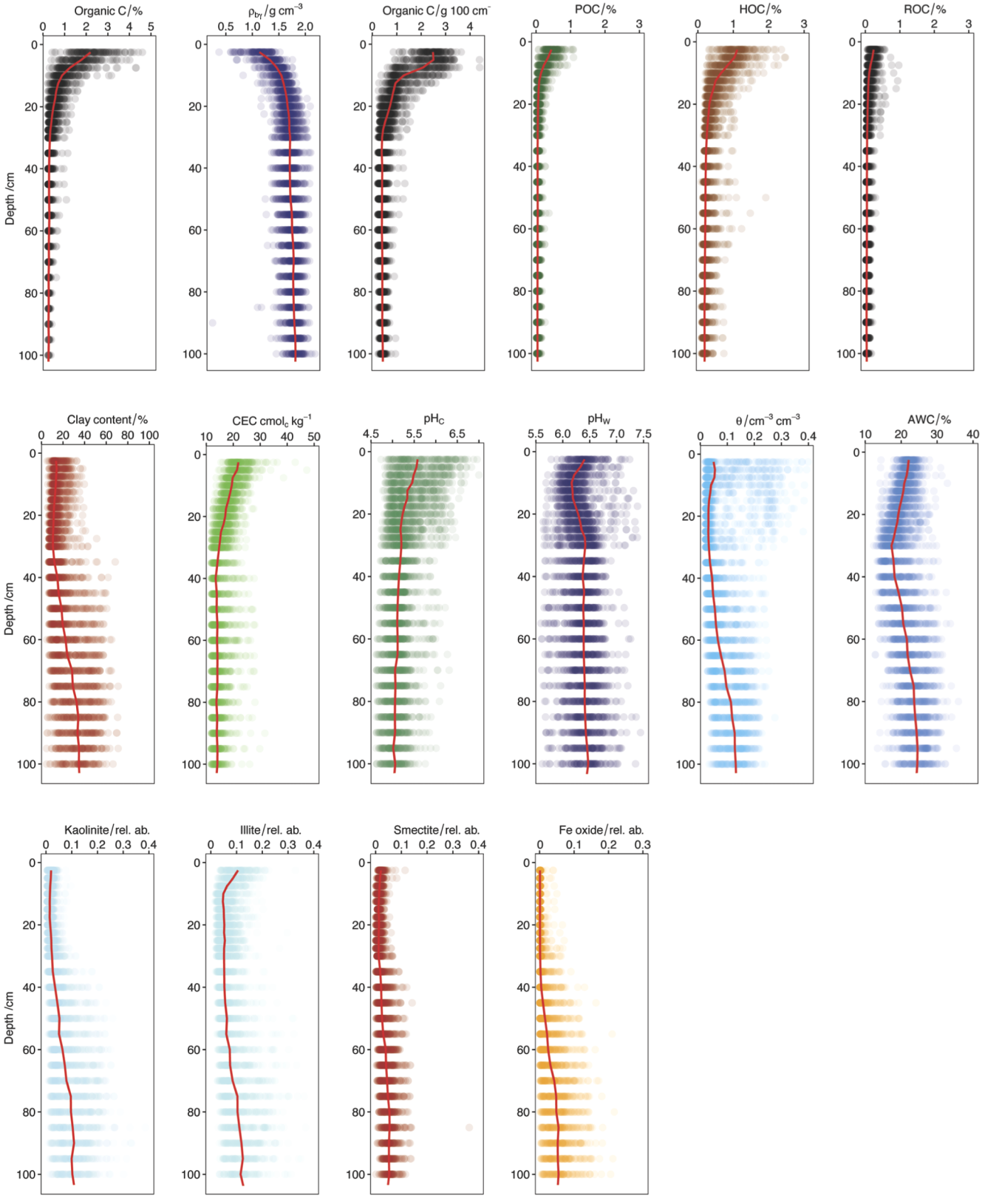


Figure 5. Sensed soil properties of all of the 150 soil cores at each of the 26 depth increments. The bulk density (ρ_{b_γ}) was measured with the γ-ray densitometer. Kaolinite, illite, smectite, and iron oxide were measured directly from their characteristic vis−NIR absorptions. Spectroscopic models were used to estimate organic C, particulate, humus, resistant C (POC, HOC, and ROC, respectively), clay content, cation exchange capacity (CEC), pH measured in 0.01 M $CaCl_2$ and water (pH_c and pH_w, respectively), volumetric water content (θ), and available water capacity (AWC).

and the vis−NIR spectra measured at each depth increment of the soil core.

A summary of the spectroscopic modeling and the performance statistics of the models are given in Table 3.

DOI: 10.1021/acs.est.7b00889
Environ. Sci. Technol. 2017, 51, 5630−5641

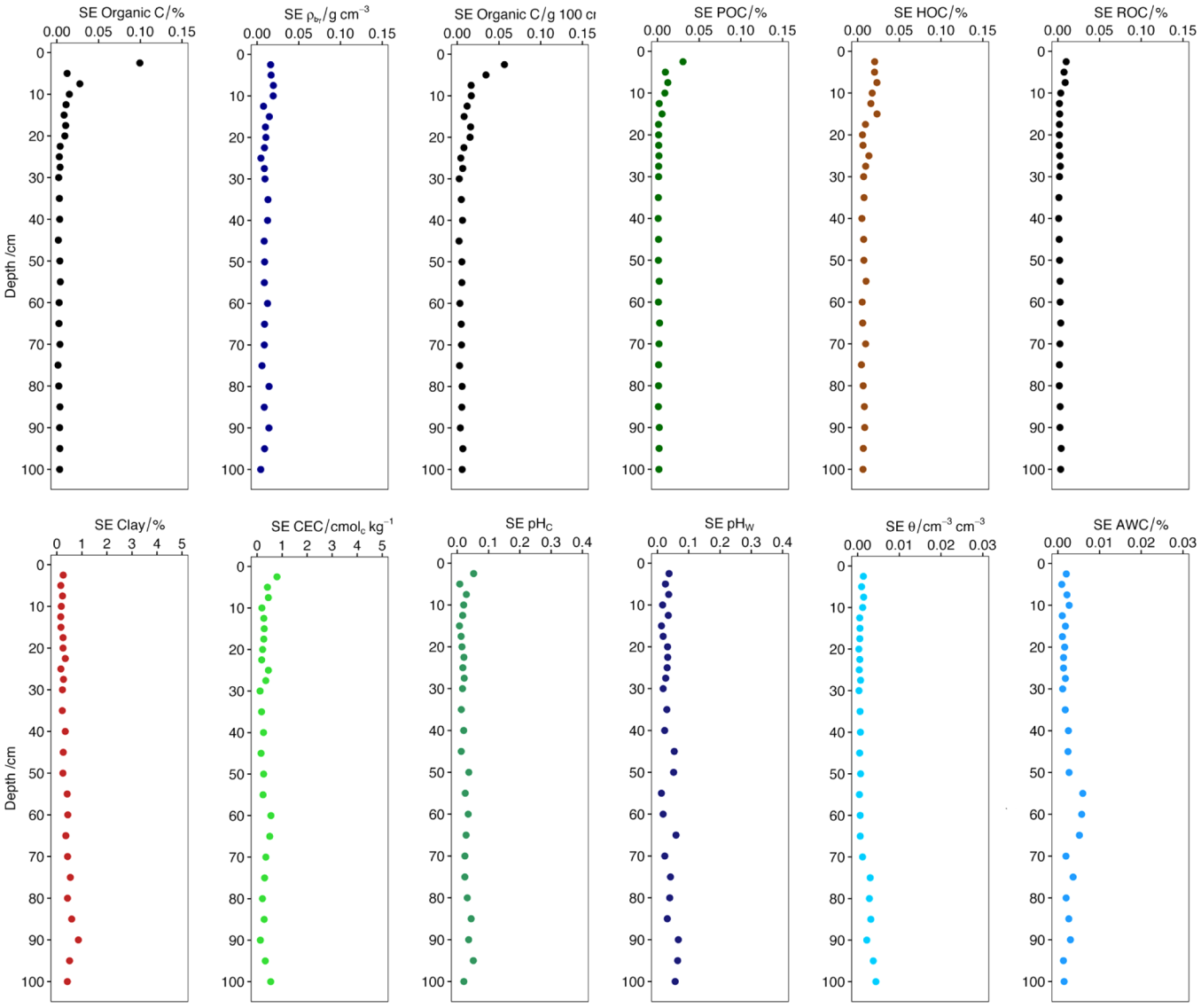


Figure 6. Repeatability of the core sensing system measurements showing the standard errors of the seven repetitions made on a single soil core, at different times and temperatures over 2 days. We provide graphs of the repeated measurements and confidence limits in the Supporting Information.

The models for θ and organic C content were excellent with R^2 values of >0.8; those for pH_c and AWC were very good with R^2 values of >0.75, and those for POC, HOC, ROC, pH_w, clay content, and CEC were good with R^2 values of ≥0.70 (Table 3). The performance statistics in Table 3 are within the ranges of those reported in the literature for studies on a local scale.[17,36−38]

The sensed soil properties of all of the 150 soil cores at each of the 26 depths are shown in Figure 5. They represent the variability of the soil properties in the profiles of our study site (closed discs in Figure 5). The median soil property values at each of the measured depths are shown by the solid lines in Figure 5.

Repeatability of the Measurements. The standard errors of the repeated measurements on the same soil core, taken at different times and temperatures over 2 days, are shown in Figure 6.

Generally, the repeated CSS measurements of different soil attribute profiles produced small standard errors and were repeatable and precise (Figure 6). There was no particular pattern or trend in the measurements taken over the 2 days and at the different temperatures between 17 and 38 °C. The Supporting Information includes graphs of the repeated measurements for each soil attribute.

Soil Property Depth Profiles. Figure 7 shows for a single soil core the spectroscopic predictions of organic C, clay content, and CEC, and the implementation of Kalman smoothing.

The uncertainties of the profile estimates of organic C were small (Figure 7) because our spectroscopic modeling used a data set selected with RS-LOCAL and the 20 well-selected samples from our study site that were representative of the different soil depths. The estimates of clay content were more accurate at the surface above 30 cm [smaller error bars in the clay content profile (Figure 7)] because for these predictions we used an RS-LOCAL data set with 12 surface soil samples form our study site. Below 30 cm, the estimates of clay content were more uncertain (Figure 7) because the data set used for the modeling did not contain local samples from below 30 cm. We did not have any local samples to help with the modeling of CEC, and so we used a general model derived with the Australian soil

DOI: 10.1021/acs.est.7b00889
Environ. Sci. Technol. 2017, 51, 5630−5641

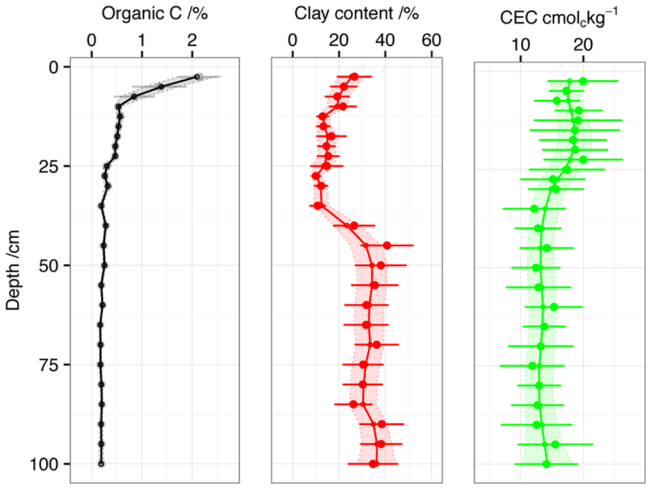


Figure 7. Example for a single core of the implementation of Kalman smoothing used to derive accurate soil attribute depth profiles. The spectroscopic model predictions of soil organic C, clay content, and cation exchange capacity (CEC) are represented by the discs; the error bars are the measurement errors, depicted as 95% confidence limits, derived from the bootstrap of the spectroscopic models, and the solid lines and shaded ribbons are the Kalman filtered estimates of the respective soil properties and their uncertainties. In this case, too, the uncertainties are represented by 95% confidence limits.

property spectral library (Table 1). Thus, the uncertainties of the CEC profile estimates were all relatively large (Figure 7).

Compared to the unfiltered discrete spectroscopic predictions with 95% confidence limits (discs with error bars in Figure 7), the Kalman estimates provide continuous profile estimates with narrower uncertainties (shaded ribbons in Figure 6). For example, Figure 8 shows the Kalman filtered continuous soil property profiles of two soil cores from our study site.

The soil organic C content of these profiles and that of the C fractions, POC, HOC, and ROC, decrease exponentially with depth. In both profiles, the humus fraction was the largest. The POC and ROC fractions were present at higher concentrations near the surface than at depth (Figure 8a,b). There was an inverse relationship between organic C content and bulk density, ρ_{b_γ} (Figure 8a,b). The SOCv profiles show that a larger portion of the stock occurs within the top 30 cm (Figure 8a,b).

The soil depicted in Figure 8a has a duplex texture profile, with clay content and the abundance of clay and iron minerals increasing below 30 cm. The volumetric water content, θ, and the AWC of the profiles had shapes similar to that of the clay content profile (Figure 8a). The soil has a bleached horizon between around 10 and 30 cm, as shown by the brighter RGB color values (Figure 8a) and the brighter appearance of the core in this layer (Figure 4). Kaolinite is the dominant clay mineral, particularly at depth. Both illite and smectite were less abundant, but their presence suggests that there is interstratification of the clay minerals in this soil. Goethite was the dominant iron oxide (Figure 8a). The CEC of the soil is larger near the surface above 25 cm than at depth. This might be due to organic C contributing primarily to the CEC near the surface and the clay minerals, particularly kaolinite, which has a CEC smaller than that of illite or smectite, at depth. The pH values measured in water and $CaCl_2$ are similar near the surface but increasingly different at depth, with pH_c becoming slightly more acidic with depth (Figure 8a).

The soil profile shown in Figure 8b has a more gradational texture profile. The θ and AWC profiles have shapes similar to that of the clay content profile (Figure 8b), but the uncertainties of the AWC estimates are larger, particularly at depth. Illite is somewhat more abundant near the surface, but otherwise, kaolinite is more abundant at depth (Figure 7b). Hematite is the dominant iron oxide in this profile (Figure 8b). The abundance of hematite increases gradually in the subsoil below 30 cm. The CEC of the soil is largest near the surface and decreases gradually to ~30 cm. As before, the larger CEC near the surface is likely to be due to the larger amounts of soil organic C, but also the greater abundance of illite (Figure 8b). pH_w was roughly 1 unit larger than pH_c throughout the profile (Figure 8b).

DISCUSSION

Different proximal sensing systems have been developed to improve the efficiency of soil analyses;[3] however, few are automated, and fewer deal with the problem of acquiring soil measurements at depth. The two main approaches involve either inserting sensors into the soil profile and making measurements *in situ* or sampling undisturbed soil core samples and measuring the soil *ex situ*. A few systems are available, at least one commercially, that enable *in situ* measurements using penetrometers, electrical conductivity sensors, and vis−NIR spectrometers with fiber optics.[39−43]

Some researchers have used sensors to manually measure soil profiles and/or core samples using vis−NIR reflectance and X-ray fluorescence (XRF) spectrometers.[44,45] For example, vis−NIR spectroscopy was used to measure the clay content, color, and mineralogy of soil profiles in the field.[5] It has also been used to measure the soil organic C content of soil core samples, in cross section[46] and longitudinally.[47] The latter study first cut the cores in half vertically and measured the vis−NIR reflectance from the plane surface of the cores at 5 cm intervals.

There are no integrated sensing systems for measuring the properties of soil profiles, although there are commercial systems for measuring sediment and rock core samples in geology, mining and exploration, and oceanography, for example, the GEOTEK,[48] HyLogger (FLSmidth Pty. Ltd., Welshpool, Australia), and Itrax (Cox Analytical Systems, Mölndal, Sweden) systems.

The SCANS (with its CSS) is unique because it provides an integrated soil sensing and data analytics solution for quantifying the properties of soil profiles, their depth functions, and their variation in space and time. Therefore, it can be useful for different environmental and agronomic applications. We used a complementary set of sensors in the CSS that cover different portions of the electromagnetic spectrum so that we might gather useful information about a range of soil properties. The CSS is modular so that as technologies develop, other sensors (e.g., portable XRF and mid-infrared spectrometers) may be relatively easily incorporated.

Although results from sensing may not be as accurate per individual measurement as those from conventional laboratory analysis, sensing is rapid and allows the collection of larger amounts of spatially explicit data using cheaper, less laborious methods. As an ensemble, the information gained will be more precise and informative. Measurements with the CSS are automated and rapid, and can be taken at fine depth resolutions predetermined by the user, depending on the application. Measuring a 1 m core at 2 cm intervals will result in 50 measurements along the core and will take approximately

DOI: 10.1021/acs.est.7b00889
Environ. Sci. Technol. 2017, 51, 5630−5641

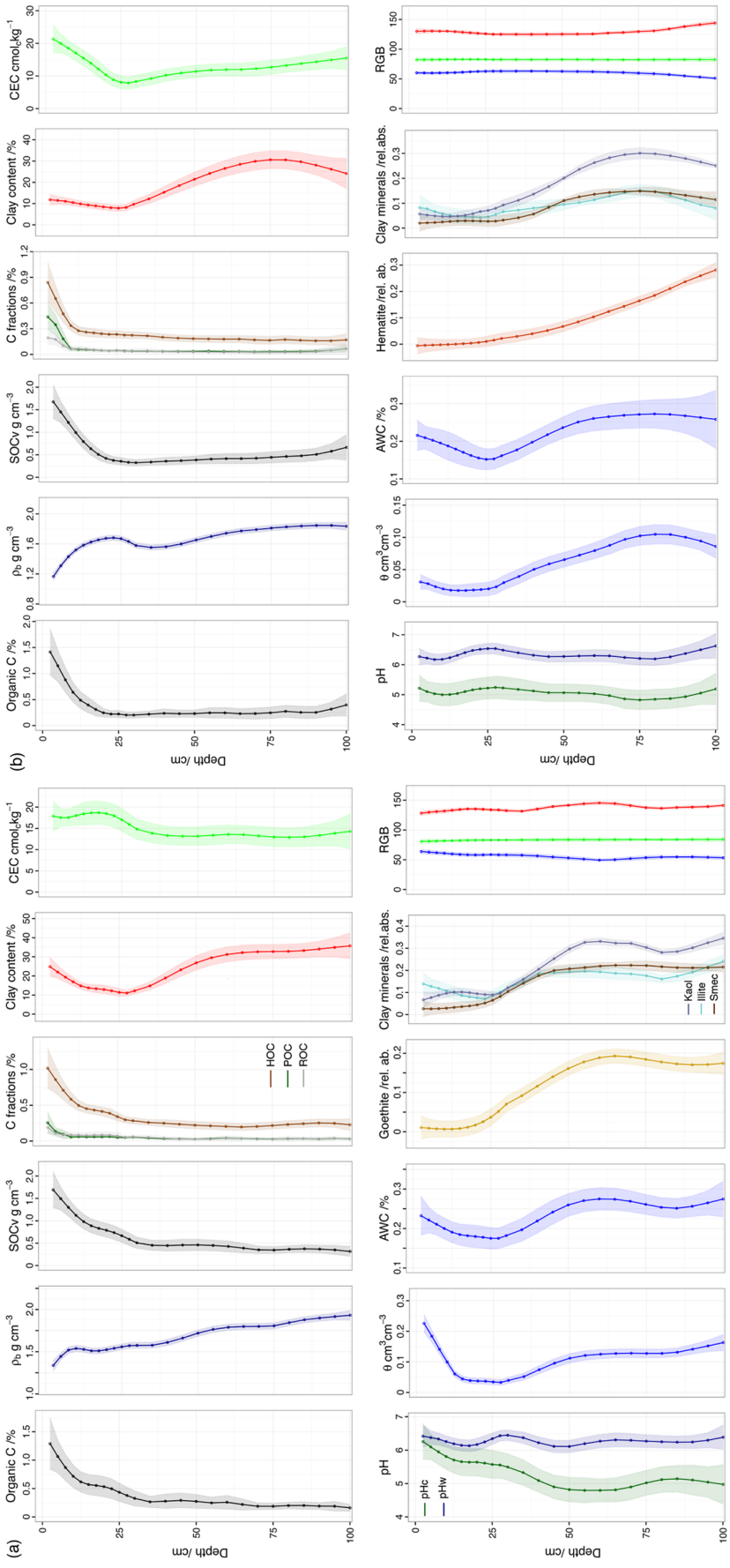


Figure 8. Measurements made on two soil core samples (a and b) from different locations on our study site. The soil attribute profiles show the locations of the Kalman filtered sensor measurements (discs) and their uncertainties (shaded ribbons), represented by 95% confidence intervals.

DOI: 10.1021/acs.est.7b00889
Environ. Sci. Technol. 2017, 51, 5630−5641

30 min. The repeatability and consistency of the system's measurements were improved by automating the CSS's calibration and measurement routines. Automation also improved the timeliness and cost efficiency of the system. An operator needs to be present only to prepare and change cores and to keep check of the measurements being taken.

As with any soil sampling, extraction of soil cores for measurement with the CSS should be performed only when suitable conditions prevail. Soil coring should be performed using standardized practice and suitable equipment (e.g., using appropriate coring tips to account for different soil types and conditions and to prevent compaction). We have not yet used the CSS with very sandy soil or with soil that has soft or weak consistency. However, using the plastic liners and appropriate sampling equipment, as described in the Supporting Information, will help to ensure that the soil cores remain intact and under field conditions for sensing. More recently, we have used the system in soil with gravel layers, and in this case, the plastic liners helped to hold the soil and gravel in place. The sensors easily differentiated between soil and gravel, so that it was easy to account for them in our analyses.

In some situations, taking the measurements *in situ* by inserting the sensors into the soil profile[41] might be advantageous. The difficulties with the *in situ* approach, however, are that penetration into the soil can be difficult in dry or hard-setting soil, it is difficult to prepare and control the measurements and the sensing areas, and direct measurements of bulk density at depth are difficult if not impossible.

The CSS has sensors that use direct and indirect inference.[3] Measures of density are made directly using γ-ray attenuation. Measurements of color, iron oxide, and clay mineralogy are also taken directly from the vis−NIR spectra. The other soil attributes, however, are measured indirectly using empirical calibrations of the reference soil attributes to vis−NIR spectra recorded with the spectrometer. The accuracies of the spectroscopic estimates depend on the models and how they are derived. When "local" (or site-specific) data are available (here we had local data on organic C, POC, HOC, ROC, clay, and pH_w), they should be used in the spectroscopic modeling (e.g. with RS-LOCAL) to improve the accuracy of the predictions. If "local" data are unavailable, then general (or "global") models using existing soil spectral libraries can be made as long as the library contains spectra from soil samples that are similar in composition to those that are present in the study area. As we have shown here (and others elsewhere[49,50]), compared to general spectroscopic models, local models produce more accurate predictions of soil attributes.

We developed the SCANS because of the urgent need for good quality spatial and temporal soil profile information for a wide range of applications.

The SCANS CSS could be used to effectively monitor soil organic C for accounting purposes[9] and to increase the frequency of adoption of best agronomic practices that will improve the condition of soil and reduce greenhouse gas emissions by farmers. Potentially, this could generate significant carbon sequestration to reach the "4 pour 1000" proposal made by French authorities ahead of the 21st Conference of Parties to the United Nations Framework Convention on Climate Change (COP21).

The flexibility of SCANS allows farmers and landowners to select, depending on their situations, either a direct measurement or a modeling approach for their organic C sequestration projects.[51] Models can be used to establish soil C storage projects by predicting the potential magnitude and uncertainty of soil organic C change over time due to specific management actions. The SCANS CSS can provide the data needed to parametrize such mechanistic soil organic C models (e.g., RothC[52,53]), which are otherwise limited by the paucity of data available to run them.

The SCANS would be useful in agronomic applications too, for example, helping to better understand nutrient mineralization, determine subsoil constraints to production, and to devise strategies for enhancing the water holding capacity and infiltration of soil. Our system simultaneously quantifies the variability of a range of biochemical, physical, and mineralogical soil properties that will help farmers to assess and monitor the condition of soil in the root zone. This could help to promote sustainable management and more targeted decision making, for example, in site-specific soil and crop management and precision agriculture.[54]

An important environmental application of the SCANS CSS could be to assess and monitor soil contamination by heavy metals and other pollutants. The use of proximal soil vis−NIR and XRF sensing for such applications has been reported in the literature.[45,55]

The SCANS provides continuous soil attribute information with a level of detail that is difficult to obtain with other approaches. Its ability to simultaneously sense many important soil properties to depth and across the landscape allows rapid, inexpensive, and spatially explicit assessments of soil. The soil information gained can help to baseline and assess whether the condition of soil, its functions, and its productivity are improving or deteriorating over time. In this way, the information from the SCANS could be used to promote good soil and environmental management, refine the sustainability of farming practices to boost food production, and sequester carbon.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.7b00889.

Descriptions of the sensors used in the SCANS core sensing system, its control system, the measurement process, and the repeatability of the measurements (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: raphael.viscarra-rossel@csiro.au. Phone: +61 2 6246 5945.

ORCID

Raphael A. Viscarra Rossel: 0000-0003-1540-4748

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the Australian Government's Filling the Research Gap Round 2 for funding this research through Project 1194194-91, "An innovative solution for accurate and affordable estimates of soil carbon". We also thank Carbon Link Pty. Ltd., Cam Banks at "Lakeview" and Craig Carter at "Tallawang", for providing us with access to the field sites and for soil sampling. We thank Jeff Baldock for the fractionation of the soil samples

DOI: 10.1021/acs.est.7b00889
Environ. Sci. Technol. 2017, 51, 5630−5641

and Iain Coombe, David Biggins, and Seija Tuomi for their contributions.

REFERENCES

(1) Viscarra Rossel, R. A.; Bouma, J. *Agricultural Systems* **2016**, *148*, 71−74.
(2) Black, C., Ed. *Methods of soil analysis. Parts 1 and 2*; Agronomy Monographs; American Society of Agronomy and Soil Science Society of America: Madison, WI, 1965.
(3) Viscarra Rossel, R. A.; Adamchuk, V. I.; Sudduth, K. A.; McKenzie, N. J.; Lobsey, C. *Adv. Agron.* **2011**, *113*, 243−291.
(4) Doetterl, S.; Stevens, A.; Van Oost, K.; van Wesemael, B. *Soil Science Society of America Journal* **2013**, *77*, 1430−1435.
(5) Viscarra Rossel, R.; Cattle, S.; Ortega, A.; Fouad, Y. *Geoderma* **2009**, *150*, 253−266.
(6) Hartemink, A. E.; Minasny, B. *Geoderma* **2014**, *230*−*231*, 305−317.
(7) Adamchuk, V. I.; Hummel, J. W.; Morgan, M. T.; Upadhyaya, S. K. *Computers and Electronics in Agriculture* **2004**, *44*, 71−91.
(8) de Gruijter, J.; McBratney, A.; Minasny, B.; Wheeler, I.; Malone, B.; Stockmann, U. *Geoderma* **2016**, *265*, 120−130.
(9) Viscarra Rossel, R. A.; Brus, D.; Lobsey, C.; Shi, Z.; McLachlan, G. *Geoderma* **2016**, *265*, 152−163.
(10) Isbell, R. *The Australian Soil Classification*, Revised edition; CSIRO Publishing: Collingwood, Australia, 2002.
(11) World reference base for soil resources. World Soil Resources Report 103; Food and Agriculture Organization of the United Nations: Rome, 2006.
(12) Lobsey, C. R.; Viscarra Rossel, R. A. *European Journal of Soil Science* **2016**, *67*, 504−513.
(13) Luo, X.; Wells, L. *Trans. ASAE* **1992**, *35*, 17−26.
(14) Baldock, J. A.; Sanderman, J.; Macdonald, L. M.; Puccini, A.; Hawke, B.; Szarvas, S.; McGowan, J. *Soil Research* **2013**, *51*, 561−576.
(15) Quinlan, J. Learning with continuous classes. Proceedings AI'92, 5th Australian Conference on Artificial Intelligence, 1992; pp 343−348.
(16) Viscarra Rossel, R. A.; Webster, R. *Eur. J. Soil Sci.* **2012**, *63*, 848−860.
(17) Viscarra Rossel, R. A.; Hicks, W. S. *European Journal of Soil Science* **2015**, *66*, 438−450.
(18) Viscarra Rossel, R. A. Potential of visible-near infrared spectroscopy to measure soil water. Report to the Grains Research and Development Corporation (GRDC) project 'Doing it better, doing it smarter. Managing soil water in Australian agriculture'; 2012.
(19) Lobsey, C.; Viscarra Rossel, R.; Roudier, P.; Hedley, C. *Eur. J. Soil Sci.* **2017**, ZZZ.
(20) Kennard, R. W.; Stone, L. A. *Technometrics* **1969**, *11*, 137−148.
(21) Rayment, G.; Lyons, D. *Soil chemical methods: Australasia*; CSIRO Publishing: Collingwood, Australia, 2011.
(22) Gee, G.; Bauder, J. In *Methods of Soil Analysis. Part 1*, 2nd ed.; Klute, A., Ed.; Agronomy Monograph 9; American Society of Agronomy and Soil Science Society of America: Madison, WI, 1986; pp 383−411.
(23) Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*; Springer Series in Statistics; Springer: Dordrecht, The Netherlands, 2009.
(24) Snee, R. D. *Technometrics* **1977**, *19*, 415−428.
(25) Newman, M. C. *Environ. Toxicol. Chem.* **1993**, *12*, 1129−1133.
(26) Duan, N. *J. Am. Stat. Assoc.* **1983**, *78*, 605−610.
(27) Viscarra Rossel, R. *J. Near Infrared Spectrosc.* **2007**, *15*, 39−47.
(28) Viscarra Rossel, R.; Minasny, B.; Roudier, P.; McBratney, A. *Geoderma* **2006**, *133*, 320−337.
(29) Clark, R. N.; Roush, T. L. *Journal of Geophysical Research: Solid Earth* **1984**, *89*, 6329−6340.
(30) Roger, J.-M.; Chauchard, F.; Bellon-Maurel, V. *Chemom. Intell. Lab. Syst.* **2003**, *66*, 191−204.
(31) Minasny, B.; McBratney, A. B.; Bellon-Maurel, V.; Roger, J.-M.; Gobrecht, A.; Ferrand, L.; Joalland, S. *Geoderma* **2011**, *167*−*168*, 118−124.
(32) Ji, W.; Viscarra Rossel, R. A.; Shi, Z. *European Journal of Soil Science* **2015**, *66*, 555−565.
(33) Helske, J. *KFAS: Kalman Filter and Smoother for Exponential Family State Space Models*, R package version 1.2.1; 2016.
(34) R Core Team. *R: A language and environment for statistical computing*; R Foundation for Statistical Computing: Vienna, 2016.
(35) Grewal, M.; Andrews, A. *Kalman Filtering: Theory and practice*; Information and System Sciences; Prentice Hall: Upper Saddle River, NJ, 1993.
(36) Bellon-Maurel, V.; McBratney, A. *Soil Biol. Biochem.* **2011**, *43*, 1398−1410.
(37) Soriano-Disla, J. M.; Janik, L. J.; Viscarra Rossel, R. A.; Macdonald, L. M.; McLaughlin, M. J. *Appl. Spectrosc. Rev.* **2014**, *49*, 139−186.
(38) Viscarra Rossel, R. A.; et al. *Earth-Sci. Rev.* **2016**, *155*, 198−230.
(39) Ben-Dor, E.; Heller, D.; Chudnovsky, A. *Soil Sci. Soc. Am. J.* **2008**, *72*, 1113−1123.
(40) Hummel, J. W.; Ahmad, I. S.; Newman, S. C.; Sudduth, K. A.; Drummond, S. T. *Trans. ASAE* **2004**, *47*, 607−618.
(41) Kweon, G.; Lund, E.; Maxton, C.; Drummond, P.; Jensen, K. *Journal of Biosystems Engineering* **2009**, *34*, 382−389.
(42) Sudduth, K. A.; Myers, D. B.; Kitchen, N. R.; Drummond, S. T. *Geoderma* **2013**, *199*, 12−21.
(43) Wetterlind, J.; Piikki, K.; Stenberg, B.; Söderström, M. *European Journal of Soil Science* **2015**, *66*, 631−638.
(44) Zhu, Y.; Weindorf, D. C.; Zhang, W. *Geoderma* **2011**, *167*−*168*, 167−177.
(45) Horta, A.; Malone, B.; Stockmann, U.; Minasny, B.; Bishop, T.; McBratney, A.; Pallasser, R.; Pozza, L. *Geoderma* **2015**, *241*−*242*, 180−209.
(46) Kusumo, B. H.; Hedley, M. J.; Hedley, C. B.; Tuohy, M. P. *Plant Soil* **2011**, *338*, 233−245.
(47) Li, S.; Shi, Z.; Chen, S.; Ji, W.; Zhou, L.; Yu, W.; Webster, R. *Environ. Sci. Technol.* **2015**, *49*, 4980−4987.
(48) *Multi-Sensor Core Logger*; Geotek, 2011.
(49) Guerrero, C.; Stenberg, B.; Wetterlind, J.; Viscarra Rossel, R. A.; Maestre, F. T.; Mouazen, A. M.; Zornoza, R.; Ruiz-Sinoga, J. D.; Kuang, B. *European Journal of Soil Science* **2014**, *65*, 248−263.
(50) Ramirez-Lopez, L.; Behrens, T.; Schmidt, K.; Rossel, R. V.; Dematte, J.; Scholten, T. *Geoderma* **2013**, *199*, 43−53.
(51) Wu, Y. Z.; Chen, J.; Ji, J. F.; Tian, Q. J.; Wu, X. M. *Environ. Sci. Technol.* **2005**, *39*, 873−878.
(52) Gebbers, R.; Adamchuk, V. I. *Science* **2010**, *327*, 828−831.
(53) Petrokofsky, G.; Kanamaru, H.; Achard, F.; Goetz, S. J.; Joosten, H.; Holmgren, P.; Lehtonen, A.; Menton, M. C.; Pullin, A. S.; Wattenbach, M. *Environmental Evidence* **2012**, *1*, 6.
(54) Jenkinson, D.; Hart, P.; Rayner, J.; Parry, L. *INTECOL Bulletin* **1987**, *15*, 1−8.
(55) Jenkinson, D. S. *Philos. Trans. R. Soc. London* **1990**, *329*, 361−368.

DOI: 10.1021/acs.est.7b00889
Environ. Sci. Technol. 2017, 51, 5630−5641