Geoderma 363 (2020) 114177

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Mapping soil organic carbon and clay using remote sensing to predict soil workability for enhanced climate change adaptation

S.S. Paul[a,*,1], N.C. Coops[b], M.S. Johnson[c], M. Krzic[a,d], A. Chandna[a], S.M. Smukler[a]

[a] Soil Science Program, Faculty of Land and Food Systems, University of British Columbia, 2357 Main Mall, Vancouver, BC V6T 1Z4, Canada
[b] Department of Forest Resources Management, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada
[c] Institute for Resources, Environment and Sustainability, University of British Columbia, 2202 Main Mall, Vancouver, BC V6T 1Z4, Canada
[d] Department of Forest and Conservation Sciences, University of British Columbia, 2357 Main Mall, Vancouver, BC V6T 1Z4, Canada

ARTICLE INFO

Handling Editor: Kristin Piikki

Keywords:
Clay content
Landsat indices
Random forest
Soil organic carbon
Soil plasticity limits

ABSTRACT

Climate change is presenting sizeable challenges for agricultural production around the world. In some regions, shifting precipitation patterns in the spring and fall are negatively impacting farm operation by reducing the number of "workable days" or the days fields can be worked with heavy equipment without damaging soil structure. This can be particularly problematic for farms on clay soils and/or poor drainage. Approximating a water content threshold at which a soil is not workable due to soil structure destruction can be helpful for planning effective farm operations. In this study, we applied advanced remote sensing and machine learning tools to produce digital maps of soil organic carbon (SOC) and clay (CL) content and used them in existing pedotransfer functions (PTFs) to predict a workability threshold (WT) across a study area in Delta, British Columbia, Canada. We combined field data, soil and vegetation indices derived from multiple Landsat satellite images, topographic indices, and soil survey information to digitally map SOC and CL of the agricultural lands in Delta using random forest (RF) and generalized boosted regression model (GBM). When validated against an independent field dataset, the RF model outperformed GBM for all accuracy measures (coefficient of determination – R^2, concordance correlation coefficient – CCC, and normalized root mean square error – nRMSE). We then spatially applied several PTFs using our digital maps to estimate the plasticity limits of the soil and produce WT map. The WT map was then tested against independent field samples of the soil water content at −10 kPa and we achieved R^2 of 0.59, CCC of 0.70, and nRMSE of 0.15. Our analysis showed that 40% of the fields in the study area had WT < 30%, a threshold that is already being impacted by reduced workable days. This WT map could be used to improve spatial prioritizations of investments for climate change adaptation at farm to regional scales.

1. Introduction

Agricultural production worldwide has become highly mechanized to reduce labor, increase efficiency and to meet the demand of the growing global population. Mechanized farming operations often must be done under a narrow range of weather conditions to avoid adverse impacts because soil conditions at the time of mechanized operation determine the level of degradation of soil structure. This is becoming more challenging as weather patterns are rapidly shifting and becoming more unpredictable in many parts of the world (Chipanshi et al., 2018). Excessive precipitation, for example, can cause poor soil trafficability, restricting the use of farm equipment during critical times of the growing season (Kolberg et al., 2019; Servadio et al., 2016). Soil trafficability is acondition at which a soil provides sufficient tire-traction for the farm vehicles without causing the soil structural deformation (Earl, 1997). Use of farm machinery when trafficable conditions are not optimal destroys soil structure and leads to soil compaction (Müller et al., 2011). This is of particular concern in humid regions where precipitation can cause soil saturation during certain parts of the year, resulting in considerable delays in field preparation or harvesting which may shorten growing season and potentially reduce crop yields. Therefore, the shifting precipitation patterns, projected by various climate models (Fischer and Knutti, 2016), will likely reduce the number of days when agricultural soils are workable with heavy equipment without causing soil degradation (Tomasek et al., 2017). Site-specific optimum water content information for soil workability could be of

* Corresponding author.
E-mail address: siddpaul@mail.ubc.ca (S.S. Paul).
[1] https://orcid.org/0000-0001-5243-8416.

https://doi.org/10.1016/j.geoderma.2020.114177
Received 14 August 2019; Received in revised form 10 November 2019; Accepted 7 January 2020

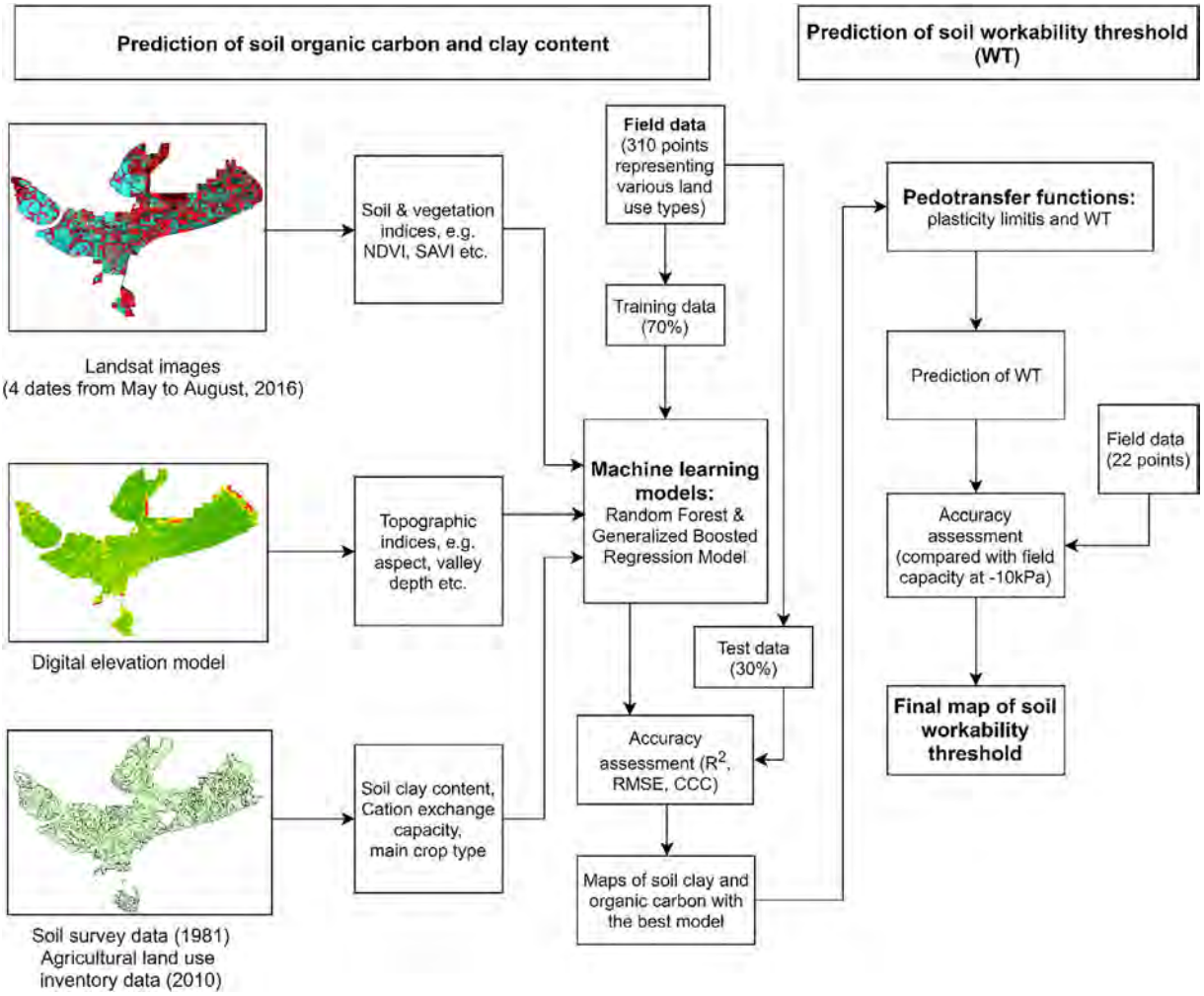

Fig. 1. Schematic diagram showing the steps of producing the maps of soil organic carbon (SOC), clay (CL), and the workability threshold. R^2, RMSE, and CCC refer to the coefficient of determination, root mean square error, and concordance correlation coefficient, respectively.

value for scheduling farm management operations as well as evaluating the impacts of changing climatic conditions and developing associated adaptation strategies.

In the literature, a soil workability threshold (WT) is defined as the optimum soil water content at which mechanical tillage operations lead to maximum number of soil aggregates (Dexter and Bird, 2001; Müller et al., 2011). WT is a combination of soil trafficability and the capacity of soil to be operated without causing substantial damage to its structure (Earl, 1997). WT depends on a number of soil properties, including texture, soil organic carbon (SOC), bulk density (Obour et al., 2017). The soil bulk density can increase as a result of overburden pressure imparted to soil by machinery, while the WT is mainly controlled by the balance of precipitation, drainage, and evapotranspiration. However, both bulk density and WT are strongly related to SOC and soil texture (Gupta and Larson, 1979). Given these relationships, many previous studies proposed methodologies to estimate WT using these two key soil properties. For example, Dexter and Bird (2001) applied SOC and texture data and the water retention characteristics of the soil to determine WT, while some authors, like Kretschmer (1996) and Mueller et al. (1990) proposed the use of consistency limits of cohesive soils for this purpose. Additionally, Bueno et al. (2006), Mapfumo and Chanasyk (1998), Rounsevell (1993), and Rutledge and Russell (1971) found that WT is highly correlated to 95–99% of the soil water content at field capacity for different soil types.

Determination of WT using soil consistency is mainly based on the Atterberg's plasticity limits – lower plastic limit (LPL) and upper plastic limit (UPL) (Campbell, 1991). Soil exhibits liquid behavior and can freely flow at the water content above UPL, but it shows friability and breaks apart under pressure at the water content below LPL (Keller and Dexter, 2012). Soil needs to be at this friable stage during mechanical operation for optimum tillage; hence, determining this threshold is critical (Keller et al., 2007; Mueller et al., 2003). Since WT is certainly associated with the soil water content at LPL, estimation of WT using the plasticity limits has been widely used for different soil types (Mueller et al., 2003; Smedema, 1993). The applications of this technique may not be effective for non-cohesive sandy soils; however, it is of value for cohesive, clay-rich soils that often have poor drainage (Dexter and Bird, 2001). Furthermore, there are well-established pedotransfer functions (PTFs) that can determine LPL and UPL from the soil texture and SOC data. For example, Kværnø et al. (2007) validated PTFs derived LPL and UPL values against field measurements achieving R^2 from 0.94 to 0.97. While the PTFs are extensively utilized, the most direct way of measuring LPL and UPL is from the remoulded soil at the laboratory (Obour et al., 2017). However, these PTFs can provide a considerable advantage in terms of time and cost effectiveness when landscape scale spatial variability of WT is of interest.

Given that soil properties are highly heterogeneous, spatially explicit and landscape scale information on WT would be helpful for designing effective climate adaptive management strategies. Yet, such information is not widely available to agricultural producers except in some selected regions. Mueller et al. (2003) demonstrated that available soil survey data and PTFs can be combined to assess the landscape

scale spatial variability of WT. Kværnø et al. (2007) presented a Norwegian case study where they examined the nature and extent of variability in soil texture, SOC, and WT within various soil map units using the approach described by Mueller et al. (2003). But the mapping of WT at the landscape scale, using advanced digital soil mapping (DSM) of SOC and clay (CL) that integrates remote sensing (RS) and machine learning (ML) techniques (e.g. random forest and generalized boosted regression model), is still limited and needs further development. Advanced DSM is currently being used widely for predictive mapping of different soil properties with high accuracies (Heung et al., 2016; McBratney et al., 2003). Hence, landscape scale mapping of WT using these state-of-the-art tools coupled with the PTFs of plasticity limits can provide particular benefits for devising regional soil management strategies to enhance adaption to changing climatic conditions. It is, however, unclear how effectively PTFs may be used for developing DSMs of WT and what modeling approach is best suited for these types of data.

To address these research gaps, the specific objectives of this study were to (1) produce maps of SOC and CL using advanced RS technique and two ML models including random forest (RF) and generalized boosted regression model (GBM), (2) compare the outcomes of the ML models for mapping SOC and CL, and (3) generate landscape scale map of WT based on the PTFs described in Mueller et al. (2003). We conducted the study in the highly intensive agricultural landscape of Delta, British Columbia, Canada, where a combination of the soils with moderate to fine texture, poor natural drainage, and increasing spring and fall precipitation has amplified concerns for soil workability.

2. Methods

Our approach in this study included a combination of field sampling and geospatial analysis using DSM and RS techniques, and Fig. 1 shows the steps involved in our analysis.

2.1. Study area

The study area represents the agricultural landscape within the City of Delta (49.08 N, 123.06 W), British Columbia, Canada and contains an area of 120 km^2 (Fig. 2). Our analysis was restricted to land within the British Columbia's Agricultural Land Reserve. The study area is in the Fraser River delta and close to the ocean with an average elevation of 10 m above mean sea level. This area comprises a mild, humid climate with a mean annual temperature of 11.1 °C and a mean annual precipitation of 1189 mm based on 30-year climate record (Environment Canada, 2019). The area is characterized by highly fertile, silty clay loam to silt loam soil with known issues of poor drainage. Delta is one of the most productive agricultural regions of British Columbia and produces a major share of the province's vegetable and blueberry crops. For the vegetable crops grown in the region, trafficability, is a major concern as a number of mechanized operations are required for preparation of the crop in the spring including discing, tillage, forming beds, nutrient applications and planting. In that fall, heavy equipment is again used for operations like harvest, discing, tillage and sowing of cover crops. Trafficability is not as much of a problem for blueberry production, as mechanized operations of mowing grasses alleyways and harvesting are done during typically dry summers. During the winter, however, farmers do use heavy equipment to assit with pruning operations.

2.2. Soil sampling and laboratory analysis

Across the study area, we collected a total of 310 soil samples at the 0–15 cm depth (Table 1) that were representative of different land use types (i.e., various annual and perennial crops, grassland, and hedgerow). At each sampling plot, we collected 4 samples (Fig. 2) from an area roughly covering the area of a Landsat image pixel (i.e., 900 m^2). The soils from these 4 samples were composited to get a representative sample for that plot, while the center of the plot was recorded with a GNSS Pro 6H Differential Global Positioning System (DGPS) (Trimble Inc., Sunnyvale, California, USA) with post-processing accuracy varying from 10 to 50 cm.

We sent 25% of the samples (n = 75) to the Technical Service Laboratory of British Columbia Ministry of Environment for determination of SOC using the combustion elemental analysis with a Vario EL Cube Elemental Analyzer (Elementar, Langenselbold, Germany) and for soil clay (CL) using the hydrometer method (Kroetsch and Wang, 2007). We also analyzed all the samples (n = 310) in TENSOR 37 spectrometer (Bruker Instruments, Ettlingen, Germany) using mid-infrared spectroscopy where the data from the standard laboratory analysis were utilized for calibration and validation in a Partial least square regression (PLSR) model.

2.3. Environmental variables for predicting SOC and CL

Topographic indices: We used the provincial 25 m Terrain Resource Information Management (TRIM) digital elevation model (Bc TRIM, 2012) to derive a suite of topographic indices. We resampled the digital elevation model to 30 m. SAGA 2.1.2 software was then utilized to produce analytical hill shading (AHS), aspect, catchment area (CA), channel network base level (CNBL), closed depressions (CD), convergence index (CI), cross sectional curvature (CSC), longitudinal curvature (LC), slope-length factor (LS), multiresolution index of the ridge top flatness (MRRTF), multiresolution index of valley bottom flatness (MRVBF), negative topographic openness (NOpen), positive topographic openness (POpen), relative slope position (RSP), slope, terrain ruggedness index (TRI), total wetness index (TWI), valley depth (VD), and total curvature (TC) from the DEM. The details of the computation of these indices are described here: http://www.saga-gis.org/saga_tool_doc/2.1.3/a2z.html. Behrens et al. (2010), Lacoste et al. (2014), and Schillaci et al. (2017) utilized some or all these topographic indices for producing DSM at different scales.

Existing soil survey and agricultural land use inventory: We used the existing detailed Canadian soil survey (1981) information and the data from British Columbia Ministry of Agriculture's agricultural land use inventory (2010) to derive environmental predictors. We extracted sand%, silt%, clay%, and cation exchange capacity (CEC) from the polygon-based soil survey map using a 30 m grid and produced raster layers for each of the soil properties using inverse distance weighting (IDW) interpolation. In case of multi-component map units, we used the dominant (i.e., covering > 50% of the unit area) category and assigned values accordingly; however, there were only a few cases with multi-component map units. The data from agricultural land use inventory (ALUI) represents detail land use information of every crop field within the agricultural land reserve in the study area (BC Ministry of Agriculture, 2016). This polygon dataset was directly rasterized to produce the ALUI covariate at 30 m spatial resolution.

Landsat image derived indices: We downloaded the Landsat 8 Level-2 surface reflectance images (Path 47, Row 26) for 2016 from the United States Geological Survey Earth Explorer data warehouse. Four Landsat scenes captured on the dates May 31, July 02, July 18, and August 19 of 2016 were used in this study to capture the seasonal variability of the agricultural landscape. Acquiring images with minimum or no cloud cover was a significant challenge and limited the choice during the image selection process. We derived a suite of soil and vegetation indices (Table 2) and several image textural variables, including homogeneity (Homo), contrast (Cont), and dissimilarity (Diss) of the images. These indices were developed for each of the four images of 2016. To produce the textural variables, we used the grey level co-occurrence matrix (GLCM) (Clausi, 2002) and derived the textural variables for the Near Infrared (B5), Short-wave Infrared-1 (B6), and Short-wave Infrared-2 (B7) bands of the Landsat images. We applied the 'glcm' package in R to generate these textural variables (Zvoleff, 2016).

Fig. 2. Map showing the fields sampled across the study area. At each field, 4 plots (P1, P2, P3, and P4) were sampled and 4 sub-samples (a, b, c, and d) were composited at each plot to get a representative sample from an area of 900 m^2 (an area covered by a Landsat satellite image pixel).

2.4. Random forest and generalized boosted regression model for predicting SOC and CL

We then used two ML models – Random Forest (RF) (Breiman, 2001) and Generalized Boosted Regression Model (GBM) (Friedman, 2001) to predict SOC and CL. These decision tree-based ensemble models are comprised of a number of nodes and leaves where the nodes perform an 'if-then' statement based on the inferred relationships between the dependent variables and a set of predictor variables. The leaves represent the 'end-nodes' where a decision is made for the prediction (Bui and Moran, 2001; Heung et al., 2014). In RF and GBM, outputs from an ensemble of decision trees are combined to improve the prediction accuracy and thus, they are being utilized for predicting complex soil landscape where the prediction is usually dependent on a large number of variables (Heung et al., 2016; Schillaci et al., 2017). RF and GBM models have their own procedures for measuring the variable importance (VI). RF calculates the percent increase in mean square error (%IncMSE) of prediction by removing the variables one by one

Table 1
Total number of soil samples collected at the 0–15 cm depth for different land use types.

Land use type	Number of samples
Annual crop	193
Perennial crop (blueberry)	60
Grassland	37
Hedgerow	20
Total	310

from the model and accordingly, determines the importance of each of the variables (Breiman, 2001). On the other hand, GBM measures the 'relative importance' score for each variable based on the empirical improvement of the model attained by splitting on a variable at the nodes and averaging over all boosted trees (Friedman, 2001). In this study, we used the 'randomForest' (Brieman et al., 2015) and 'gbm' (Ridgeway, 2015) packages in R to implement these models. We used 70% of the field data (n = 310) for training the models, while the rest of the data were applied for accuracy assessment. We utilized coefficient of determination (R^2), concordance correlation coefficient (CCC), root mean square error (RMSE), and normalized root mean square error (nRMSE) to assess the accuracy of the predicted outputs. nRMSE is the RMSE value normalized by dividing by the difference between the maximum and minimum values of the observed data (Shen et al., 2016). Eqs. (1)–(4) below describe these accuracy measures.

$$R^2 = \frac{(\sum_{i=1}^{n}(Y_i - \bar{X})(X_i - \bar{X}))^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \sum_{i=1}^{n}(X_i - \bar{X})^2} \quad (1)$$

$$CCC = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2 + (\bar{X} - \bar{Y})^2} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - X_i)}{n}} \quad (3)$$

$$nRMSE = \frac{1}{Z}\sqrt{\frac{\sum_{i=1}^{n}(Y_i - X_i)}{n}} \quad (4)$$

In Eqs. (1)–(4), X and Y represent measured and predicted values, respectively; n is the number of samples; X_i and Y_i are the paired i^{th} values from the measured and predicted data, respectively; $\bar{X}$ and $\bar{Y}$ are the mean of the predicted and observed data, respectively; ρ is the Pearson correlation coefficient between the measured and predicted values; σ_x and σ_y are the corresponding variances of the measured and predicted values; Z is the difference between the maximum and minimum values of the observed data. Between the two models, the most accurate prediction of SOC and CL were used in the subsequent step for estimating WT.

2.5. Prediction and validation of WT

In this study, we used a series of PTFs for predicting WT. At first, we determined the UPL and LPL using Eqs. (5) and (6) where the SOC and CL maps produced in step 2.4 were used as the inputs. Eqs. (5) and (6) were modified after Olson (1975) considering that SOC comprises 50% of the total organic matter in the soil (Pribyl, 2010). These maps of UPL and LPL were then applied to Eq. (7) to predict WT which is the maximum soil water content that provides optimum workability (Kværnø et al., 2007; Mueller et al., 2003). In all equations, the soil water content is represented as gravimetric % ().

$$UPL = 11.9 + 0.92 \times CL\% + 0.08 \times SOC\% \ (after \text{ Olson, } 1975) \quad (5)$$

$$LPL = 7.15 + 0.199 \times CL\% + 1.957 \times SOC\% \ (after \text{ Olson, } 1975) \quad (6)$$

$$WT = LPL - 0.15 \times (UPL - LPL) \ (\text{Kretschmer, } 1996; \text{ Mueller et al., } 2003) \quad (7)$$

To validate the prediction of WT, we collected an additional set of soil samples from 22 locations across the study area using a Conditioned Latin Hypercube Sampling technique (Minasny and McBratney, 2006) which selected stratified random sampling locations based on the spatial variability of the topography and soil types of the study area. The samples were collected as undisturbed cores at 0–7.5 cm depth and analyzed in the lab to determine the soil water content at field capacity at −10 kPa. We herein used −10 kPa matric potential for field capacity since our PTF of WT is assumed to be equal to the water content at this matric potential (Kværnø et al., 2007). We used Richard's Pressure Plate Apparatus (Richards and Fireman, 1943) for this purpose. The samples were completely hydrated, weighed, placed inside the pressure plate chamber, and allowed to equilibrate at −10 kPa pressure. We then oven-dried the samples at 105 °C for 48 h and weighed again. We determined the soil bulk density from the mass of oven-dried soils and the volume of the sampling core. Particle density was also calculated using the mass of oven-dried soils and the total volume of soil particles. We then calculated the porosity of soil using Eq. (8). Finally, the soil water content at −10 kPa (or at field capacity) was determined by applying Eq. (9).

$$Porosity\ of\ soil = \left(1 - \frac{Bulk\ density}{Particle\ density}\right) \times 100 \quad (8)$$

$$Soil\ water\ content\ at\ -10\,\text{kPa} = Water\ lost\ at\ -10\,\text{kPa} \times \frac{Porosity\ of\ Soil}{Total\ water\ lost} \quad (9)$$

Thereafter, we extracted the WT values of these 22 locations from the predicted map and tested for the accuracy measures described in the previous section (i.e., R^2, CCC, RMSE, and nRMSE).

3. Results and discussion

3.1. Summary of soil properties from the field plot data

The SOC and CL in the study area varied substantially as would be expected in a region where land use is very heterogeneous and texture is influenced by the dynamics of the large adjacent river (Table 3). The range of SOC values in the sampled field plots may be attributed to the variable nutrient and soil management practices in these fields, the

Table 2
List of soil and vegetation indices. In the formula, R, B, NIR, SWIR1, SWIR2 refer to the red, blue, near infra-red, short wave infrared-1, short wave infrared-2 bands of Landsat 8 satellite imagery, respectively, while L refers to canopy background adjustment factor.

Soil and vegetation indices	Formula	References
Normalized Difference Vegetation Index (NDVI)	(NIR − R)/(NIR + R)	Rouse et al. (1974)
Soil Adjusted Vegetation Index (SAVI)	(1 + L) (NIR − R)/(NIR + R + L), L = 0.5	Huete (1988)
Normalized Difference Moisture Index (NDMI)	(NIR − SWIR1)/(NIR + SWIR1)	Hunt and Rock (1989)
Soil Brightness Index (SBI)	$\sqrt{((R)^2 + (NIR)^2)}$	Elvidge and Lyon (1985)
Normalized Difference Tillage Index (NDTI)	(SWIR1–SWIR2)/(SWIR1 + SWIR2)	Van Deventer et al. (1997)
Clay Minerals Ratio (CMR)	SWIR1/SWIR2	Carranza and Hale (2002)
Bare Soil Index (BSI)	((SWIR1 + R) − (NIR + B))/((SWIR1 + R) + (NIR + B)) * 100 + 100	Rikimaru et al. (2002)

Table 3
Summary statistics for soil organic carbon (SOC) and clay (CL).

Soil Property	Minimum	Maximum	Mean	Standard Deviation	Range
SOC (%)	0.58	4.76	2.82	0.78	4.18
CL (%)	6	29	18.81	7.82	27

types of crop (i.e., annual vs. perennial crops) and non-production perennial vegetation (i.e., hedgerow, grass margin) scattered in and around the field. On the other hand, the range of the CL content was much larger than that of SOC, with plots closer to the river exhibiting higher CL content.

3.2. Selection of environmental variables for RF and GBM models

Of the 80 environmental variables derived from multiple sources, including topography (i.e., digital elevation model), soil survey and ALUI, and Landsat imagery, our analysis identified between 29 and 41 variables of high importance depending on the modeling approach. We ran the models for both SOC and CL with including all the environmental variables and then, identified the top predictors based on a threshold value of 5 for %IncMSE in RF and a threshold value of 1 of relative importance score in GBM. This process reduced the number of variables by 35–50% depending on the soil property and the type of model (Figs. 3 and 4). Overall, the topographic variables and soil survey and ALUI variables were found to be stronger predictors than the Landsat variables based on the VI scores for both models. It is also important to note that the Landsat indices, derived from the images of pre-growing and post-harvest seasons, when soils were left without cover, explained more variance predictors than the indices developed from the growing season images. The CEC was found to be the most dominant predictor for SOC in both models, while multiresolution index of the ridge top flatness (MRRTF) and land use types from ALUI

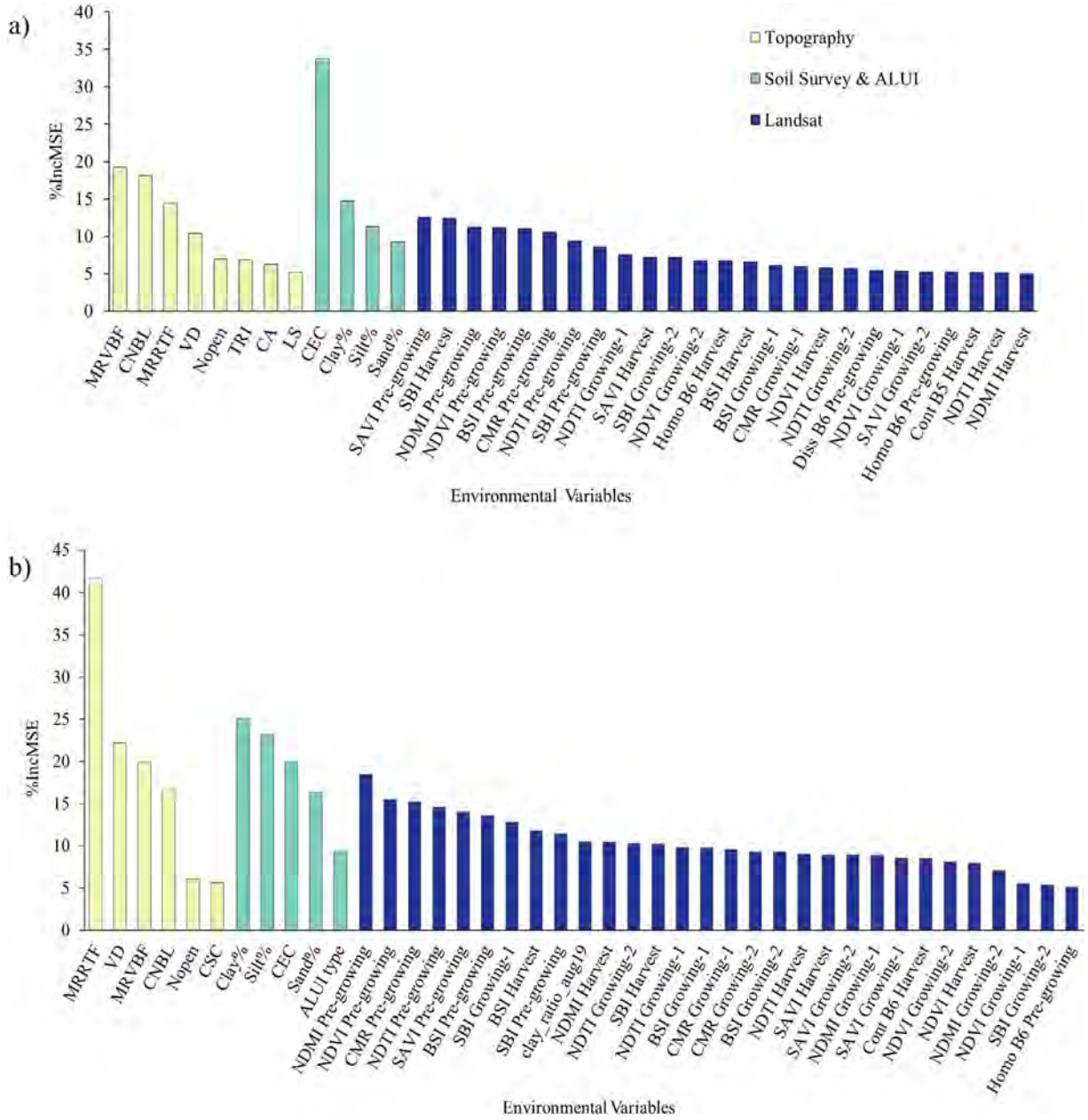

Fig. 3. Most important environmental variables in random forest model – a) for predicting soil organic carbon, b) for predicting clay. See section 2.3 for the full name of the variables.

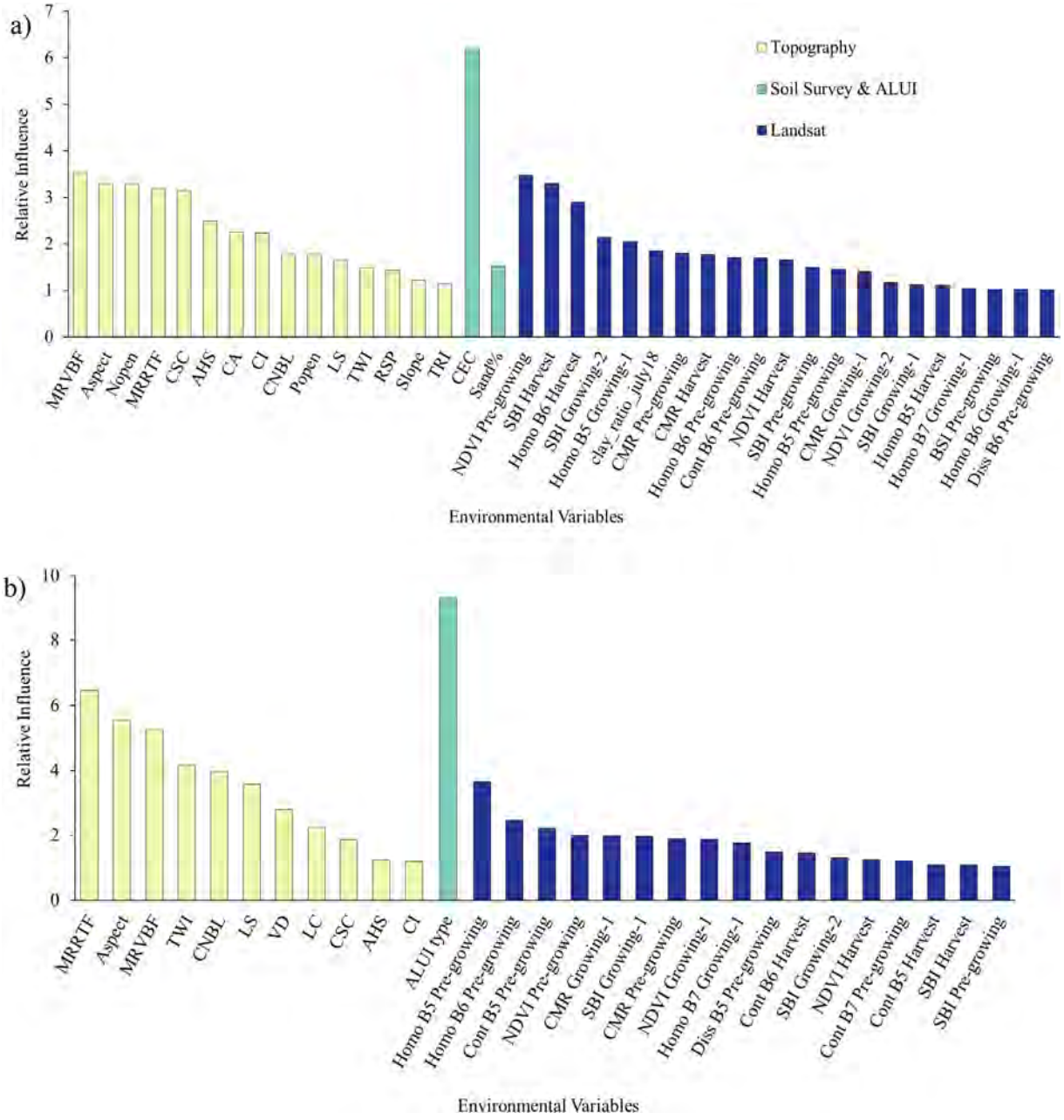

Fig. 4. Most important environmental variables in generalized boosted regression model – a) for predicting soil organic carbon, b) for predicting clay. See section 2.3 for the full name of the variables.

Table 4
Accuracy of prediction of soil organic carbon (SOC) and clay (CL) using random forest (RF) and generalized boosted regression model (GBM). R^2, CCC, and nRMSE represent coefficient of determination, concordance correlation coefficient, and normalized root mean square error (nRMSE) respectively.

Model	Accuracy metrics	SOC	CL
RF	R^2	0.55	0.62
	CCC	0.70	0.72
	nRMSE	0.12	0.15
GBM	R^2	0.39	0.41
	CCC	0.55	0.63
	nRMSE	0.14	0.20

were the most important variables for predicting CL using RF and GBM models, respectively. We also performed Pearson correlation analysis on the selected variables presented in Figs. 3 and 4 and removed any variable from the final model if it had a correlation coefficient value of ≥80% with another variable.

3.3. Assessment of the performance of RF and GBM models

In our study, both ML models performed reasonably well, but there were important differences in their accuracy metrics depending on the soil property. Both RF and GBM models predicted CL more accurately than SOC, although the differences were minor (Table 4). We found that RF outperformed GBM for both SOC and CL for all accuracy metrics (i.e., R^2, CCC, and nRMSE). However, the differences between the models were more obvious when R^2 values were compared for both SOC and CL. The nRMSE and CCC of predictions were somewhat close to each other, especially for SOC. The R^2 and CCC of SOC prediction using RF were 38% and 12% higher, respectively than those using GBM, while the nRMSE was 14% less with the RF model. We found similar

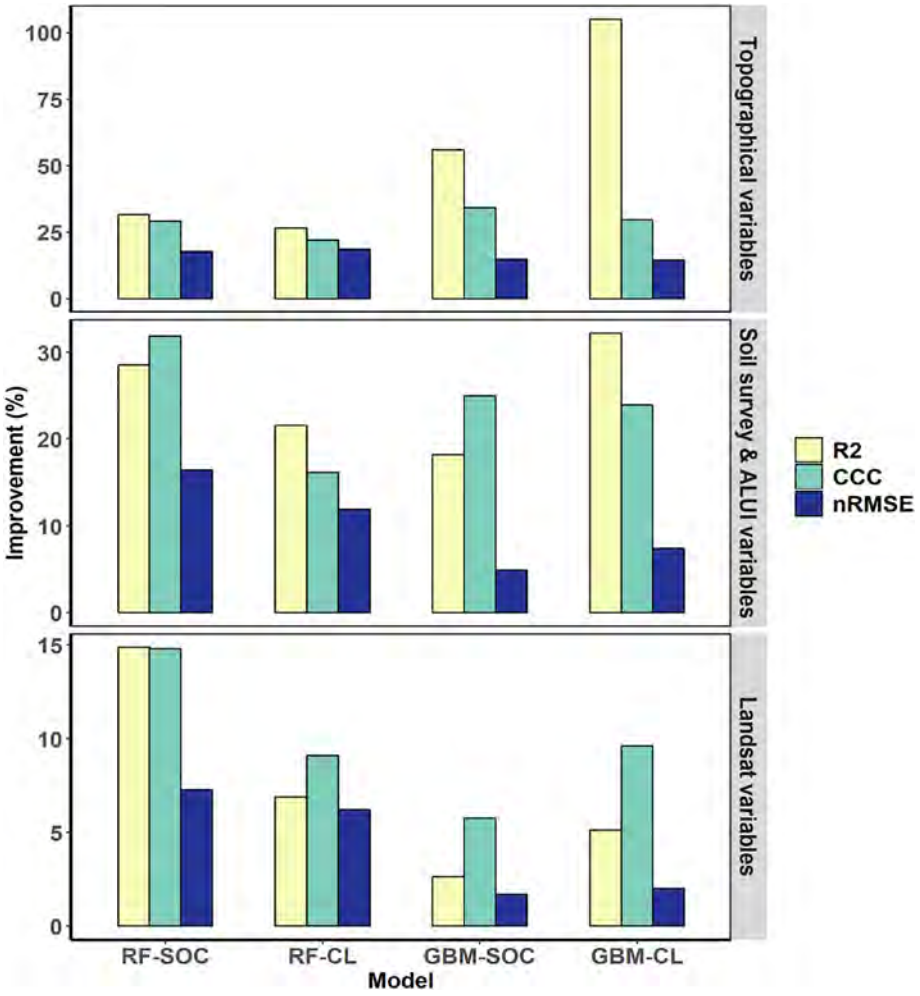

Fig. 5. Improvement (%) in model accuracy in terms of R^2 (coefficient of determination), CCC (concordance correlation coefficient), and nRMSE (root mean square error) for the (a) topographic, (b) soil survey & ALUI, and (c) Landsat variables using both random forest (RF) and generalized boosted regression (GBM) models for predicting soil organic carbon (SOC) and clay (CL).

results for the prediction of CL where the differences in R^2, CCC, and nRMSE were 51%, 26%, and 25%, respectively between the outcomes of RF and GBM models. We also tested the model accuracy when including all the 80 variables for all cases to test if discarded variables had any effect on the model performance. We found that the model accuracies were not impacted, except the R^2 and CCC of RF-CL prediction were decreased by 3% and 2%.

That the RF model performed better than GF in our study is consistent with Wang et al. (2018) who mapped SOC in a semi-arid rangeland of Australia, predicting SOC with R^2 values ranging between 0.42 and 0.48 and CCC values ranging from 0.56 to 0.62 for various sets of environmental variables, compared with values for R^2 (0.55) and CCC (0.7) for RF in the present study. In contrast, the study by Yang et al. (2016) reported that GBM performed better than RF for modeling SOC in an alpine ecosystem, especially in the areas with greater vegetation cover.

Landscape scale prediction of CL using these techniques is relatively sparse in literature, but our findings support the results of Chagas et al. (2016) who predicted CL with RF and obtained an R^2 of 0.56, which is slightly lower than what we were able to achieve in this study. Our results did differ from those of Sindayihebura et al. (2017) who demonstrated that GBM better predicted CL across Burundi's central plateau in Africa. Such differences in performance of these ML models in various studies have been mainly attributed to the dissimilarities in landscapes and environmental conditions, the scale of prediction, and the type and quality of environmental variables used (Were et al., 2015). Therefore, selecting a single ML model as the best method for predicting landscape scale soil properties is difficult, and largely site and case dependent (Ließ et al., 2016). In our study, the RF model likely outperformed GBM because of the RF model's simple parameterization and reduced susceptibility to overfitting as well as greater descriptive power that enables the model to decipher the complex and hierarchical relationships between the environmental variables and the target soil properties (Wang et al., 2018).

We also examined the influence of each variable category on the prediction accuracy of the models where the results reported in Table 4 were used as the benchmark for comparison (Fig. 5). The topographic variables were the most important category of predictors in our study, regardless of model or soil property. However, they showed significantly stronger predictive capability for some cases of GBM as compared to RF. For example, when topographic variables were included in the GBM model, the R^2 of SOC prediction was improved by 56% whereas the improvement was 32% for RF. Similarly, R^2 of CL prediction was improved by 105% for GBM and by 27% for RF when the topographical variables were included. The second most important predictor category was the soil survey and ALUI variables followed by the Landsat variables. We also observed that Landsat variables were more influential for RF predictions than those of GBM. For instance, the R^2, CCC, and nRMSE of the RF prediction of SOC was improved by 15%, 7%, and 15%, respectively when Landsat variables were added to the model, but the improvements were 3%, 2%, 6%, respectively when modeled with GBM. We obtained similar results for the prediction of

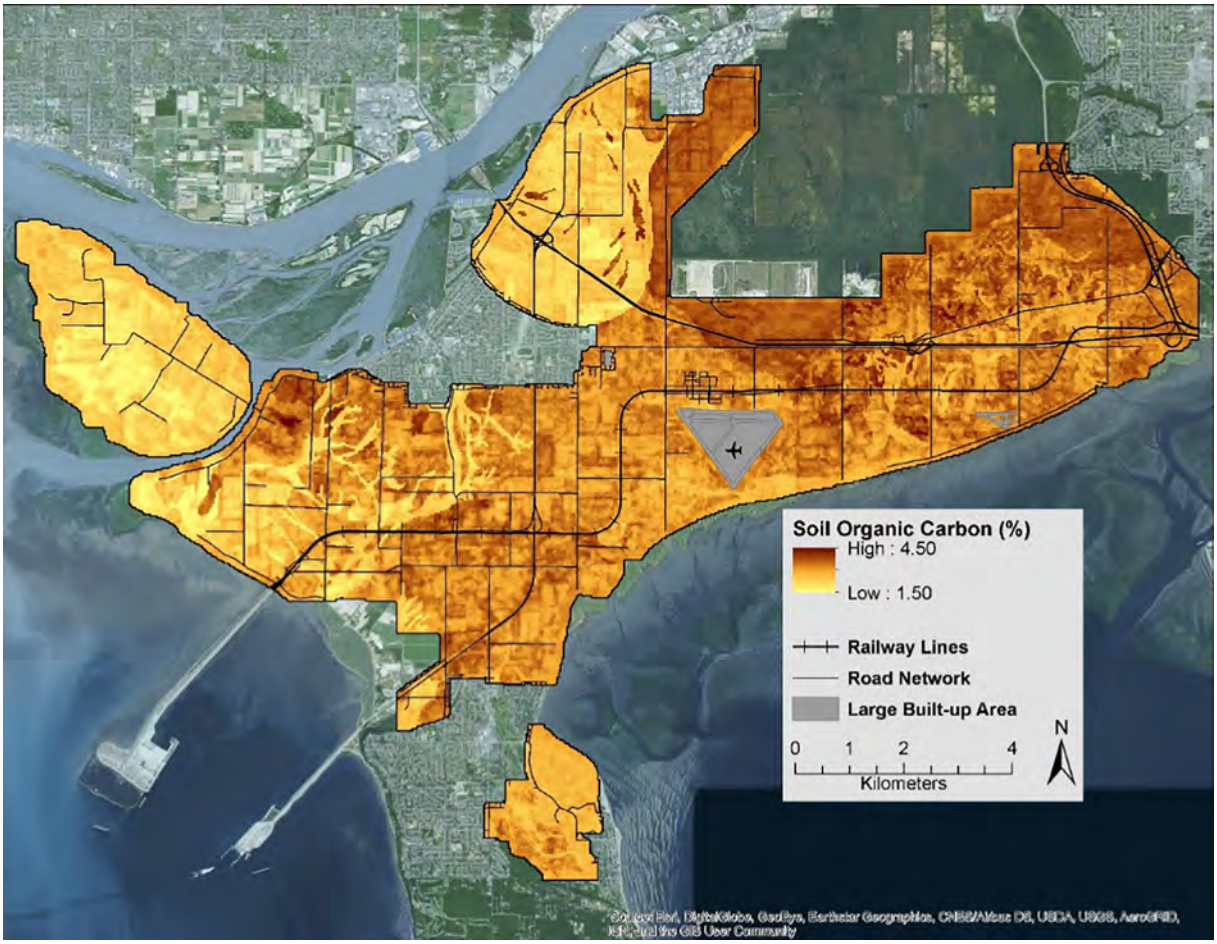

Fig. 6. Soil organic carbon (SOC) map of the study area predicted using the random forest model.

CL.

As mentioned above, the performance of the ML models is highly dependent on site characteristics, and the predictive capability of the environmental variables differs from case to case. For instance, the strong performance of soil and vegetation information derived from Landsat variables was reported by Chagas et al. (2016) and Grinand et al. (2017) for predicting CL and SOC, respectively using the RF model whereas Grimm et al. (2008) and Schillaci et al. (2017) found topographic indices along with the soil survey and land use data as the most important predictors of SOC in their studies using both RF and GBM models. The findings of the latter two examples agree with our results. The topography of a site is related to the erosion and deposition of soil materials (Cavazzi et al., 2013); hence, it was expected that topographic variables would have a strong correlation with the target soil properties. In addition, the soil survey data used in our study captured the inherent characteristics of the soil landscape and a much greater distribution of field-based data than in our study. Although these data were collected almost 40 years ago, soil texture is unlikely to change and changes in SOC are likely to have been correlated with land use. The agricultural land use data informed the model about the agricultural practices of the study area. This would likely be even more useful if we had data of land use changes over time. A study by Schillaci et al. (2017) found that incorporating multiple Landsat images spanning the whole season was more effective for predicting soil properties at landscape scale, but we did not observe a strong predictive influence for the Landsat variables, even after including a large number of indices derived from four Landsat images representing the whole growing season. Kheir et al. (2010) reported that removal of the Landsat variables increased the overall accuracy of their predictions. However, in our case, Landsat variables improved the accuracy of the models by 6 to 15% (Fig. 5); although the improvement was not large, Landsat variable including contributed to model performance.

Even though the variable selection process substantially reduced the number of the variables and satisfactory predictions were obtained using them, we attempted to assess the model performance with even fewer variables so that we can identify the key predictors and minimize the analysis effort. To accomplish this, we ran the models including only two top predictors from each category. For instance, the RF-SOC model only included MRVBF, CNBL, CEC, Clay%, SAVI pre-growing, and SBI harvest as the predictor variables. We then compared the outcomes of these reduced models with the results of the full model as shown in Table 4. Interestingly, we achieved relatively similar accuracies for both RF and GBM models, with some variation depending on the soil property and accuracy metric. The accuracy remained the same for all the metrics of RF-SOC model; however, the accuracy of RF-CL model decreased by 20%, 7%, and 34% in terms of R^2, CCC, and nRMSE, respectively. On the other hand, for GBM model, the results did not change substantially for either soil property, where R^2 of the predictions decreased by 5% and 9% for SOC and CL, respectively. These findings somewhat agree with the results obtained by Wang et al. (2018) where prediction accuracies for SOC remained unchanged or slightly improved when a more parsimonious model was used, although their variable selection approach was more complex compared with ours. Based on these results, we conclude that identifying only a few key environmental covariates based on the variable importance scores for a specific geographical area can improve the accuracy of digital soil maps. This may substantially reduce the analysis effort for producing the additional covariates.

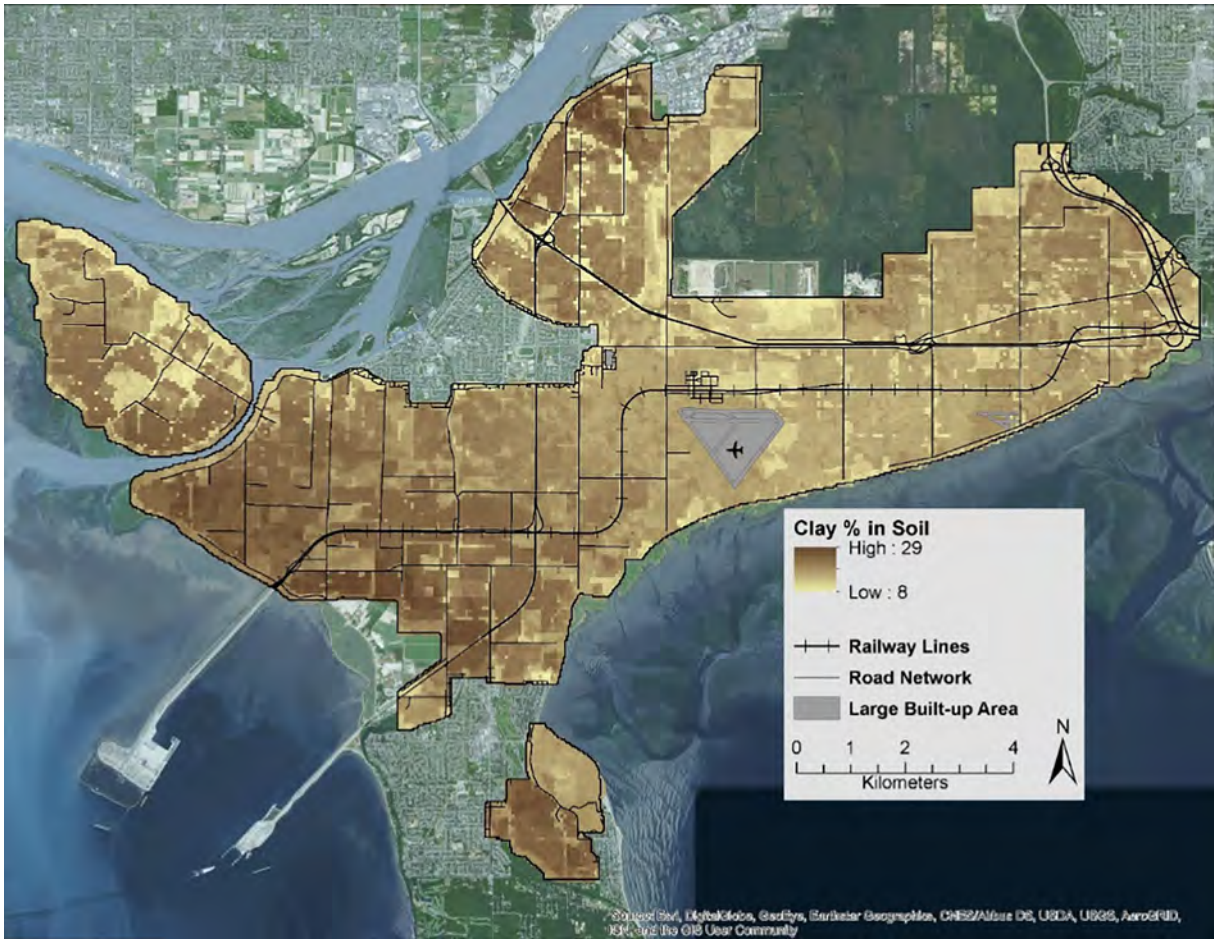

Fig. 7. Clay (CL) map of the study area predicted using the random forest model.

3.4. Spatial distribution of SOC and CL

Given that the RF model performed better than GBM, it was used to predict the SOC and CL for the entire study area at a 30 m spatial resolution. The SOC content in our study area varied from 1.5 to 4.5% (Fig. 6). Although there were several patches of crop fields with a higher concentration of SOC distributed across the area; generally, the fields in the north-eastern region had the highest concentration of SOC. This region of the study area borders Burns Bog, which is a unique ombrotrophic raised peat bog in the Fraser River Delta (Hebda et al., 2000). Moreover, this region was dominated by perennial highbush blueberry production which is known for sequestering carbon in the soil (Nemeth et al., 2017). Together, the historic bog ecosystem followed by a perennial cropping system resulted in high SOC of this small region of the study area. Fields dominated by intensive annual crop production exhibited lower SOC concentration. In addition, the fields with lower SOC, especially on Westham Island have known issues of soil salinity (Lussier et al., 2019). High salinity reduces crop production resulting in inputs of organic matter and in turn lower SOC concentrations (Rietz and Haynes, 2003).

The study area is characterized by CL content that ranged between 8 and 29% (Fig. 7), where the western half of Delta was dominated by fields with higher CL. However, the highest CL values were observed in the fields adjacent to the river which deposited a large amount of fine sediments on those fields over the course of soil formation. Although Clay% from the soil survey data was one of the strongest predictors of SOC in the RF model, the CL and SOC were inversely correlated in our maps with a Pearson correlation coefficient of −0.19. Numerous studies have reported the opposite trend, showing the chemical adsorption of carbon onto the surface of clay minerals led to greater SOC in clay-rich soils than in coarse textured ones (Johannes et al., 2017; Singh et al., 2018). In our study area, the fields with higher CL are also characterized by poor drainage, which in combination with intensive tillage, often results in the destruction of soil structure, compaction, and increased soil erosion (Müller et al., 2011), causing loss of SOC (Lilly et al., 2018). Thus, the observed reverse relationship between CL and SOC suggests that intensive tillage is most likely resulting in SOC losses across the study region. Our maps will provide a base-line for long-term monitoring across the region and enable tracking of future changes in SOC.

3.5. Distribution of WT and accuracy of prediction

The prediction of WT for agricultural land in Delta resulted in higher than standard accuracy when validated with independent soil samples analyzed for field capacity at −10 kPa. The validation of the predicted WT map (Fig. 8) resulted in R^2 of 0.59, CCC of 0.70, and nRMSE of 0.15. The WT ranged from 20 to 42% across the study area, where fields with high SOC and/or low CL exhibited high WT, and vice versa. Although we have found no other studies that have used advanced DSM to predict WT, Kværnø et al. (2007) produced a map of WT using simple kriging of the estimated values from soil map units; however, they concluded that their WT map did not capture the differences between soil types, especially at the boundaries between two or more soil units. Our map of WT is a substantial improvement in this regard as it effectively captured the variation in soil properties in a more continuous manner across the study area.

The WT map highlights that a substantial portion of the crop fields

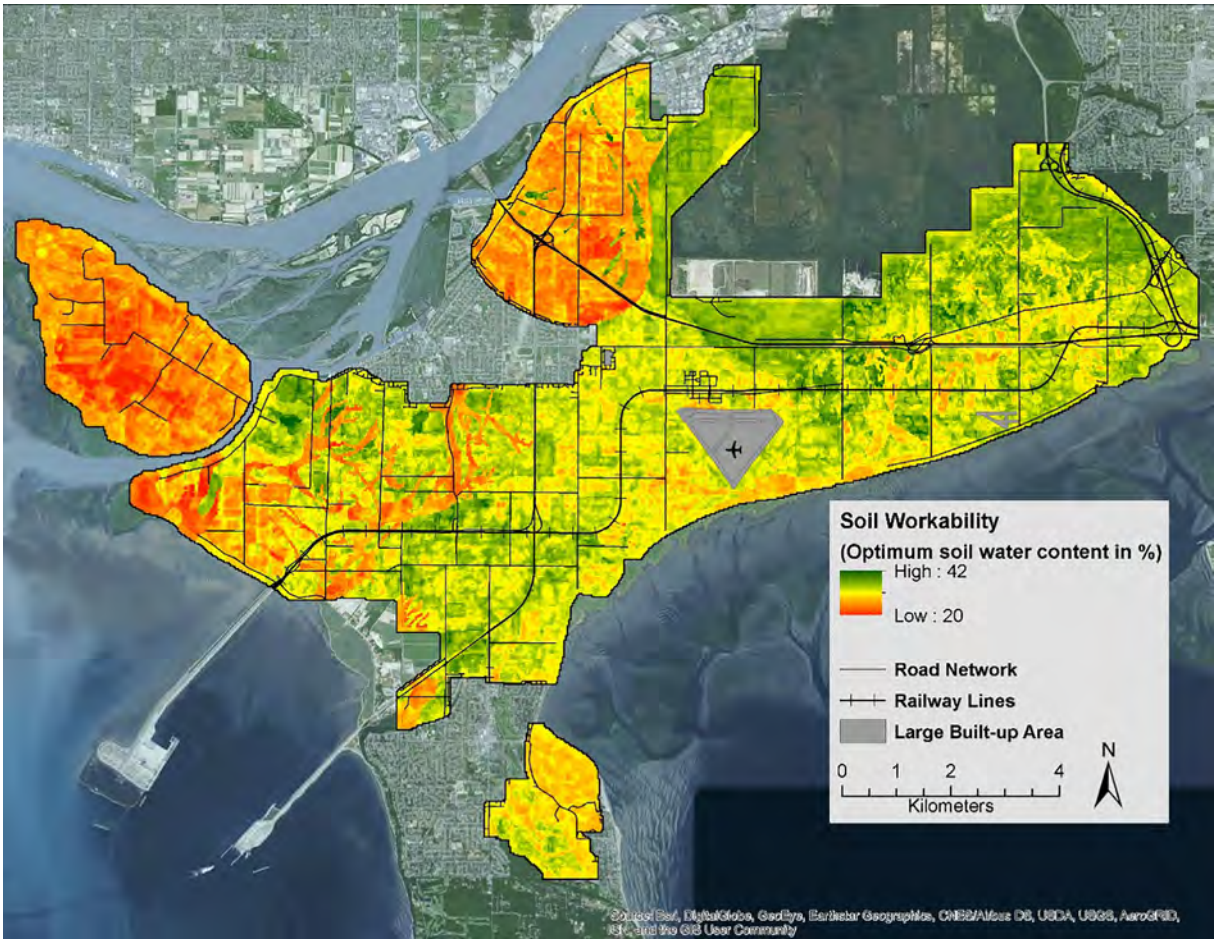

Fig. 8. Soil workability (WT) map of the study area where the values represent the optimum gravimetric soil water content (%) above which soil may be degraded with any mechanical operation.

in Delta will likely face serious challenges due to poor workable conditions especially during the wet part of the year (i.e., spring and fall). Our map showed that 40% of the crop fields in Delta had a WT < 30% (category-1), while 59% had a WT ranging from 30 to 40% (category-2), and 1% had WT > 40% (category-3). Based on a study by Neufeld et al. (2017), which tracked soil water content (at 0–15 cm depth) across 26 fields in Delta with both heavy tillage and no-tillage practices, category-2 and -3 fields would have been workable by April 11th in the typical spring conditions of 2016, while fields in category-1 would be workable only a week later. Alternatively, in the unusually wet spring of 2017, fields in category-2 and -3 would not have been workable until May 15th and category-1 fields at least two weeks later. This pattern could also be observed in the fall of 2016 where category 1 fields were workable only until September 15th while category 2–3 fields could have been workable for another month. Wet conditions in the spring and the fall could result in large differences in the number of workable days between category-1 fields and those in category-2 and -3. This situation is expected to become more challenging in the coming years as climate models predicted a 7% increase in precipitation for the region by 2050 occurring mainly in the spring and fall (BC Agriculture & Food Climate Action Initiative, 2015). Adapting to this shifting precipitation pattern will likely require substantial investment at the farm level to enhance SOC or install drainage infrastructure field with low WT and at the regional level to improve water conveyance.

4. Conclusions

Producing spatially explicit information of soil workability is critical for effective management and climate change adaptation in agricultural lands. We combined advanced remote sensing and machine learning tools with existing PTFs to produce a digital map of WT for an intensive agricultural landscape in Delta, British Columbia, Canada. We predicted SOC and CL across the landscape using RF and GBM models and found that RF was the best approach for both soil properties. Combining the digital maps of SOC and CL using a number of PTFs to produce a map of WT did not result in much reduction in accuracy. The WT map identified that 40% of the crop fields in the study area had a WT < 30%, a threshold that will likely result in substantially fewer workable days than the other fields in the region. Our analysis demonstrates an effective approach to spatially predict WT across a heterogeneous agricultural landscape. These results can be used to formulate efficient farming strategies in the study area for more effective climate change adaptation. However, results may vary depending on the soil type, climate, and agricultural management practices; thus, future analysis should focus on validating the methodology in other geographical contexts.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to acknowledge the funding support from the Farm

Adaptation Innovator Program of the British Columbia Climate Action Initiative. We also thank the producers for access to their fields for sampling purposes and the undergraduate assistants for their sincere and diligent assistance during field sampling and lab work.

References

BC Agriculture & Food Climate Action Initiative, 2015. BC Agriculture & Climate Change Regional Adaptation Strategies series – Fraser Valley. Victoria, BC.

BC Ministry of Agriculture, 2016. Fraser Valley Regional District agricultural land use inventory, Summer 2011–2013. Abbosford, BC.

BC TRIM, 2012. British Columbia Terrain Resource Information Mapping (TRIM) Digitial Map Products.

Behrens, T., Zhu, A.X., Schmidt, K., Scholten, T., 2010. Multi-scale digital terrain analysis and feature selection for digital soil mapping. Geoderma 155, 175–185. https://doi.org/10.1016/j.geoderma.2009.07.010.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Brieman, L., Cutler, A., Liaw, A., Wiener, M., 2015. randomForest package. Repos. CRAN.

Bueno, J., Amiama, C., Hernanz, J.L., Pereira, J.M., 2006. Penetration resistance, soil water content, and workability of grasslands soils under two tillage systems. Trans. ASABE 49, 875–882.

Bui, E., Moran, C., 2001. Disaggregation of polygons of surficial geology and soil maps using spatial modelling and legacy data. Geoderma 103, 79–94.

Campbell, D.J., 1991. Liquid and plastic limits. In: Smith, K., Mullins, C. (Eds.), Soil Analysis–Physical Methods. Dekker Inc., New York, pp. 367–398.

Carranza, E.J.M., Hale, M., 2002. Mineral imaging with Landsat Thematic Mapper data for hydrothermal alteration mapping in heavily vegetated terrane. Int. J. Remote Sens. 23, 4827–4852.

Cavazzi, S., Corstanje, R., Mayr, T., Hannam, J., Fealy, R., 2013. Are fine resolution digital elevation models always the best choice in digital soil mapping? Geoderma 195, 111–121.

Chagas, C., de Carvalho Junior, W., Bhering, S.B., Calderano Filho, B., 2016. Spatial prediction of soil surface texture in a semiarid region using random forest and multiple linear regressions. Catena 139, 232–240.

Chipanshi, A., Fitzmaurice, J., De Jong, R., Bogdan, D., Lewis, M., Kroetsch, D., Lee, D., 2018. Assessment of soil trafficability across the agricultural region of the Canadian Prairies with the gridded climate data set. Soil Till. Res. 184, 128–141.

Clausi, D.A., 2002. An analysis of co-occurrence texture statistics as a function of grey level quantization. Can. J. Remote Sens. 28, 45–62.

Dexter, A.R., Bird, N.R.A., 2001. Methods for predicting the optimum and the range of soil water contents for tillage based on the water retention curve. Soil Till. Res. 57, 203–212.

Earl, R., 1997. Prediction of trafficability and workability from soil moisture deficit. Soil Till. Res. 40, 155–168. https://doi.org/10.1016/S0167-1987(96)01072-0.

Elvidge, C.D., Lyon, R.J.P., 1985. Influence of rock-soil spectral variation on the assessment of green biomass. Remote Sens. Environ. 17, 265–279.

Environment Canada, 2019. Canadian Climate Normals 1981–2010 [WWW Document]. URL http://climate.weather.gc.ca/climate_normals/results_1981_2010_e.html?searchType=stnName&txtStationName=delta&searchMethod=contains&txtCentralLatMin=0&txtCentralLatSec=0&txtCentralLongMin=0&txtCentralLongSec=0&stnID=766&dispBack=0 (accessed 3.23.19).

Fischer, E.M., Knutti, R., 2016. Observed heavy precipitation increase confirms theory and early models. Nat. Clim. Chang. 6, 986.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232.

Grimm, R., Behrens, T., Märker, M., Elsenbeer, H., 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island—Digital soil mapping using Random Forests analysis. Geoderma 146, 102–113.

Grinand, C., Maire, G. Le, Vieilledent, G., Razakamanarivo, H., Razafimbelo, T., Bernoux, M., 2017. Estimating temporal changes in soil carbon stocks at ecoregional scale in Madagascar using remote-sensing. Int. J. Appl. Earth Obs. Geoinf. 54, 1–14. https://doi.org/10.1016/j.jag.2016.09.002.

Gupta, S.C., Larson, W.E., 1979. Estimating soil water retention characteristics from particle size distribution, organic matter percent, and bulk density. Water Resour. Res. 15, 1633–1635. https://doi.org/10.1029/WR015i006p01633.

Hebda, R.J., Gustavson, K., Golinski, K., Calder, A.M., 2000. Burns Bog Ecosystem Review: Synthesis Report for Burns Bog, Fraser River Delta, South-western British Columbia, Canada. Environmental Assessment Office Victoria, BC.

Heung, B., Bulmer, C.E., Schmidt, M.G., 2014. Geoderma Predictive soil parent material mapping at a regional-scale : a Random Forest approach. Geoderma 214, 141–154. https://doi.org/10.1016/j.geoderma.2013.09.016.

Heung, B., Ho, H.C., Zhang, J., Knudby, A., Bulmer, C.E., Schmidt, M.G., 2016. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. Geoderma 265, 62–77.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309.

Hunt Jr, E.R., Rock, B.N., 1989. Detection of changes in leaf water content using near-and middle-infrared reflectances. Remote Sens. Environ. 30, 43–54.

Johannes, A., Matter, A., Schulin, R., Weisskopf, P., Baveye, P.C., Boivin, P., 2017. Optimal organic carbon values for soil structure quality of arable soils. Does clay content matter? Geoderma 302, 14–21.

Keller, T., Dexter, A.R., 2012. Plastic limits of agricultural soils as functions of soil texture and organic matter content. Soil Res. 50, 7–17.

Keller, T., Arvidsson, J., Dexter, A.R., 2007. Soil structures produced by tillage as affected by soil water content and the physical quality of soil. Soil Till. Res. 92, 45–52.

Kheir, R.B., Greve, M.H., Bøcher, P.K., Greve, M.B., Larsen, R., McCloy, K., 2010. Predictive mapping of soil organic carbon in wet cultivated lands using classification-tree based models: the case study of Denmark. J. Environ. Manage. 91, 1150–1160.

Kolberg, D., Persson, T., Mangerud, K., Riley, H., 2019. Impact of projected climate change on workability, attainable yield, profitability and farm mechanization in Norwegian spring cereals. Soil Till. Res. 185, 122–138.

Kretschmer, H., 1996. Koernung und Konsistenz, in: Blume, H.-P., Felix-Henningsen, P., Fischer, W.R., Frede, H.G., Horn, R., Stahr, K. (Eds.), Handbuch Der Bodenkunde. Ecomed.

Kroetsch, D., Wang, C., 2007. Particle size distribution. In: Carter, M.R., Gregorich, E.G. (Eds.), Soil Sampling and Methods of Analysis. CRC Press, Boca Raton, FL, pp. 713–725.

Kværnø, S.H., Haugen, L.E., Børresen, T., 2007. Variability in topsoil texture and carbon content within soil map units and its implications in predicting soil water content for optimum workability. Soil Till. Res. 95, 332–347. https://doi.org/10.1016/j.still.2007.02.001.

Lacoste, M., Minasny, B., McBratney, A., Michot, D., Viaud, V., Walter, C., 2014. High resolution 3D mapping of soil organic carbon in a heterogeneous agricultural landscape. Geoderma 213, 296–311. https://doi.org/10.1016/j.geoderma.2013.07.002.

Ließ, M., Schmidt, J., Glaser, B., 2016. Improving the spatial prediction of soil organic carbon stocks in a complex tropical mountain landscape by methodological specifications in machine learning approaches. PLoS One 11, e0153673.

Lilly, A., Baggaley, N.J., Loades, K.W., McKenzie, B.M., Troldborg, M., 2018. Soil erosion and compaction in Scottish soils: adapting to a changing climate. Dundee.

Lussier, J.M., Krzic, M., Smukler, S.M., Bomke, A.A., Bondar, D., 2019. Short-term effects of grassland set-asides on soil properties in the Fraser River delta of British Columbia. Can. J. Soil Sci. 99, 136–145.

Mapfumo, E., Chanasyk, D.S., 1998. Guidelines for safe trafficking and cultivation, and resistance–density–moisture relations of three disturbed soils from Alberta. Soil Till. Res. 46, 193–202.

McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma. https://doi.org/10.1016/S0016-7061(03)00223-4.

Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Comput. Geosci. 32, 1378–1388. https://doi.org/10.1016/j.cageo.2005.12.009.

Mueller, L., Tille, P., Kretschmer, H., 1990. Trafficability and workability of alluvial clay soils in response to drainage status. Soil Till. Res. 16, 273–287.

Mueller, L., Schindler, U., Fausey, N.R., Lal, R., 2003. Comparison of methods for estimating maximum soil water content for optimum workability. Soil Till. Res. 72, 9–20. https://doi.org/10.1016/S0167-1987(03)00046-1.

Müller, L., Lipiec, J., Kornecki, T.S., Gebhardt, S., 2011. Trafficability and workability of soils. In: Encyclopedia of Agrophysics. Springer, pp. 912–924.

Nemeth, D., Lambrinos, J.G., Strik, B.C., 2017. The effects of long-term management on patterns of carbon storage in a northern highbush blueberry production system. Sci. Total Environ. 579, 1084–1093. https://doi.org/10.1016/J.SCITOTENV.2016.11.077.

Neufeld, K., Paul, S., Smukler, S., 2017. Improving On-Farm Drainage Management to Reduce the Impacts of Climate Change in Delta, BC. Vancouver, Canada.

Obour, P.B., Lamandé, M., Edwards, G., Sørensen, C.G., Munkholm, L.J., 2017. Predicting soil workability and fragmentation in tillage: a review. Soil Use Manag. 33, 288–298.

Olson, G.O., 1975. Engineering characteristics of Ohio soil series.

Pribyl, D.W., 2010. A critical review of the conventional SOC to SOM conversion factor. Geoderma 156, 75–83. https://doi.org/10.1016/J.GEODERMA.2010.02.003.

Richards, L.A., Fireman, M., 1943. Pressure-plate apparatus for measuring moisture sorption and transmission by soils. Soil Sci. 56, 395–404.

Ridgeway, G., 2015. The gbm package. Generalized boosted regression models (Documentation on the R Package 'gbm', version 1.6–3).

Rietz, D.N., Haynes, R.J., 2003. Effects of irrigation-induced salinity and sodicity on soil microbial activity. Soil Biol. Biochem. 35, 845–854.

Rikimaru, A., Roy, P.S., Miyatake, S., 2002. Tropical forest cover density mapping. Trop. Ecol. 43, 39–47.

Rounsevell, M.D.A., 1993. A review of soil workability models and their limitations in temperate regions. Soil Manage. 9, 15–20.

Rouse, J., Jr., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS, in: NASA. Goddard Space Flight Center 3d ERTS-1 Symposium. NASA, United States, College Station, TX. pp. 309–317.

Rutledge, P.L., Russell, D.G., 1971. Work Day Probabilities for Tillage Operations in Alberta. Edmonton.

Schillaci, C., Acutis, M., Lombardo, L., Lipani, A., Fantappiè, M., Märker, M., Saia, S., 2017. Spatio-temporal topsoil organic carbon mapping of a semi-arid Mediterranean region: the role of land use, soil texture, topographic indices and the influence of remote sensing data to modelling. Sci. Total Environ. 601–602, 821–832. https://doi.org/10.1016/j.scitotenv.2017.05.239.

Servadio, P., Bergonzoli, S., Beni, C., 2016. Soil tillage systems and wheat yield under climate change scenarios. Agronomy 6, 43.

Shen, W., Li, M., Huang, C., Wei, A., 2016. Quantifying live aboveground biomass and forest disturbance of mountainous natural and plantation forests in Northern Guangdong, China, Based on Multi-Temporal Landsat, PALSAR and field plot data Wenjuan. Remote Sens. 8, 1–24. https://doi.org/10.3390/rs8070595.

Sindayihebura, A., Ottoy, S., Dondeyne, S., Van Meirvenne, M., Van Orshoven, J., 2017. Comparing digital soil mapping techniques for organic carbon and clay content: case study in Burundi's central plateaus. Catena 156, 161–175.

Singh, M., Sarkar, B., Sarkar, S., Churchman, J., Bolan, N., Mandal, S., Menon, M., Purakayastha, T.J., Beerling, D.J., 2018. Stabilization of soil organic carbon as influenced by clay mineralogy. In: Advances in Agronomy. Elsevier, pp. 33–84.

Smedema, L.K., 1993. Drainage performance and soil management. Soil Technol. 6, 183–189.
Tomasek, B.J., Williams II, M.M., Davis, A.S., 2017. Changes in field workability and drought risk from projected climate change drive spatially variable risks in Illinois cropping systems. PLoS One 12, e0172301.
Van Deventer, A.P., Ward, A.D., Gowda, P.H., Lyon, J.G., 1997. Using Thematic Mapper data to identify contrasting soil plains and tillage practices. Photogramm. Eng. Remote Sens. 63, 87–93.
Wang, B., Waters, C., Orgill, S., Cowie, A., Clark, A., Li Liu, D., Simpson, M., McGowen, I., Sides, T., 2018. Estimating soil organic carbon stocks using different modelling techniques in the semi-arid rangelands of eastern Australia. Ecol. Indic. 88, 425–438. https://doi.org/10.1016/J.ECOLIND.2018.01.049.
Were, K., Bui, D.T., Dick, Ø.B., Singh, B.R., 2015. A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. Ecol. Indic. 52, 394–403.
Yang, R., Zhang, G., Liu, F., Lu, Y., Yang, Fan, Yang, Fei, Yang, M., Zhao, Y., Li, D., 2016. Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an Ecol. Indic. 60, 870–878. https://doi.org/10.1016/j.ecolind.2015.08.036.
Zvoleff, A., 2016. Package 'glcm'.