BioOne COMPLETE

Comparing direct and indirect approaches to predicting soil texture class

Author: Saurette, Daniel D.

Comparing direct and indirect approaches to predicting soil texture class

Daniel D. Saurette [iD]

[a]Ontario Ministry of Agriculture, Food and Rural Affairs, 1 Stone Road West, 3rd Floor SE, Guelph, ON N1G 4Y2, Canada; [b]School of Environmental Sciences, University of Guelph, 50 Stone Road East, Guelph, ON N1G 2W1, Canada

Corresponding author: **Daniel D. Saurette** (email: Daniel.Saurette@ontario.ca)

Abstract

Soil texture, or the relative proportions of sand, silt, and clay, is a key soil attribute that influences many important physical, chemical, and biological properties of soils. Digital soil mapping is increasingly used to predict soil texture; however, few comparisons have been made between direct prediction of a texture class, and the indirect prediction of texture class by first predicting sand, silt, and clay content, and subsequently converting the predictions to a texture class. We predicted soil texture class for the 5–15 and 30–60 cm depth intervals of the Ottawa soil survey project using direct and indirect approaches which yielded a similar overall accuracy (28–36%) and kappa (0.19–0.27). The predicted soil maps had a similar spatial distribution of soil texture classes. We then used the Euclidean distance between the texture classes to adjust the model performance metrics, revealing the indirect approach provided the better soil texture class prediction. When comparing the predictions, the 5–15 and 30–60 cm maps were in perfect agreement for 53% and 42% of the study area, respectively, and in both cases texture class predictions were within one texture class for over 87% of the map area. For many studies, including legacy soil surveys, texture class information is available, and particle size distribution data are generally lacking. This study confirms that direct prediction of soil texture class performs almost equally with indirect prediction.

Key words: digital soil mapping, texture, particle size, additive log ratio, weighted metrics

Résumé

La texture du sol, c'est-à-dire les proportions relatives de sable, de limon et d'argile, est une caractéristique capitale qui influe sur de nombreuses propriétés physiques, chimiques et biologiques du sol. On recourt de plus en plus à la cartographie numérique des sols pour en prévoir la texture. Cependant, on a rarement comparé la méthode de prévision directe à la méthode indirecte qui prévoit, en un premier temps, les proportions de sable, de limon et d'argile, puis convertit le résultat en classe de texture. Les auteurs ont prévu la classe de texture des sols d'Ottawa échantillonnés à une profondeur de 5–15 cm et de 30–60 cm par les deux approches et ont obtenu des résultats similaires pour l'exactitude générale (28–36 %) et l'indice kappa (0,19–0,27). Les cartes du sol illustraient une distribution spatiale semblable pour les classes de texture. Ensuite, les auteurs ont utilisé la distance euclidienne entre les classes de texture pour ajuster les mesures de la performance du modèle et ont constaté que l'approche indirecte donne les meilleures prévisions. Quand on compare celles-ci, les cartes pour les intervalles de profondeur de 5–15 cm et de 30–60 cm s'accordent à la perfection avec 53 % et 42 % des sols examinés, respectivement. Dans les deux cas, la classe de texture prévue se situe à moins d'une classe de distance pour au-delà de 87 % de la surface couverte par la carte. Si la classe de texture est souvent connue, même pour les levés les plus anciens, les données sur la granulométrie manquent habituellement. Cette étude confirme que la méthode de prévision directe de la texture du sol est presque aussi efficace que la méthode de prévision indirecte. [Traduit par la Rédaction]

Mots-clés : cartographie numérique du sol, texture, granulométrie, rapport logarithmique additif, pondération des mesures

Introduction

Soil texture is a unit of classification which expresses the proportions of the individual particle size fractions (PSFs; sand, silt, and clay) and is the most well-known composition in soil science (Odeh et al. 2003; Krzic et al. 2021). The relative proportions of the PSFs can be used to assign a texture class based on any one of a multitude of soil texture classifica-tion schemes (e.g., United States Department of Agriculture (USDA; Soil Science Division Staff 2017), Canadian System of Soil Classification (SCWG 1998), "Aisne" triangle (Jamagne 1967), etc.). Soil texture is of critical importance because it affects countless physical, chemical, and biological properties of soil (Liu et al. 2020). Given the importance of soil texture, there is a need for high-resolution maps of PSFs or texture

class, and digital soil mapping (DSM) has provided a mechanism by which these products can be generated with relative efficiency (Malone and Searle 2021a).

Soil texture is typically determined either by field estimation of texture class, or by laboratory analysis by determining the proportions of sand, silt, and clay in a bulk sample, usually by hydrometer, pipette, or laser diffraction. As such, in a DSM context, texture class can be predicted directly using a categorical modelling approach (Laborczi et al. 2016; Maynard and Levi 2017; Gomez et al. 2019) or inferred by predicting sand, silt, and clay content using a regression modelling approach (Poggio and Gimona 2017; Pahlavan-Rad and Akbarimoghaddam 2018; Amirian-Chakan et al. 2019; Liu et al. 2020). In some instances, a combination approach has been used where sand, silt, and clay content were estimated for observations, where only texture class is known based on a priori information such as a subset of samples (Dharumarajan and Hegde 2022; Richer-de-Forges et al. 2022) or on texture class centroids (Levi 2017; Malone and Searle 2021a). The combination approach is particularly relevant since soil texture class data are more abundant than particle size data due to their ease of acquisition (Levi 2017).

Compositional data, such as the composition of sand, silt, and clay which determines soil texture class, are non-negative and with a constant sum (Odeh et al. 2003; Greenacre 2021). Odeh et al. (2003) noted that modelling individual components of composition was not meaningful, and thus introduced the use of the additive log ratio (ALR) transformation of PSF data prior to kriging and compared the approach to kriging of the untransformed PSF data and compositional kriging. They concluded that kriging of the untransformed data resulted in many areas not summing to unity (100%) and that the ALR transformation technique outperformed both compositional kriging and kriging of the untransformed data. Three log-ratio transformations have since been extensively used for soil texture modelling applications: the ALR, the isometric log ratio, and the centred log ratio (Odeh et al. 2003; Wang et al. 2020; Zhang et al. 2020; Poggio et al. 2021). These data transformations are flexible in that they can be used to preprocess compositional data before modelling. As such, with the formalization of DSM (McBratney et al. 2003), these techniques are now commonly coupled with machine learning. Geostatistical approaches to mapping soil properties have limitations such as the assumption of normally distributed residuals, challenges associated with nonlinear relationships between dependent and independent variables, reliance on the variogram that might not adequately capture spatial heterogeneity, and computational requirements for large data sets (Wadoux et al. 2020). Conversely, machine learning approaches are a data-driven approach and do not make assumptions about the distribution of the data and can identify complex nonlinear relationships between covariates and the dependent variable to partition the data more effectively and improve predictions (Thompson et al. 2012; Wadoux et al. 2020).

Laborczi et al. (2016) noted that texture class maps could be derived using direct and indirect prediction and that these could be evaluated using validation statistics. Despite many studies that leverage DSM workflows to predict either PSFs or texture classes, we only found one study that predicted texture class both directly using a categorical modelling approach and indirectly by first modelling PSFs and then converting these to texture class (Zhang et al. 2020). Furthermore, of those studies predicting texture classes directly, the vast majority do not account for the magnitude of misallocations when reporting validation statistics. This is problematic since there appears to be a disconnect between the model performance metrics used to evaluate regression models and those used to evaluate classification models. The fundamental difference is that in a classification model, the results are binary — either the prediction is correct, or it is incorrect — whereas, in regression, performance metrics are typically calculated as a departure from a reference line (e.g., regression line), and hence inherently consider the magnitude of the error. To address this concern, the taxonomic distance can be used to adjust the accuracy assessment of soil class predictions to account for the magnitude of a misclassification (Minasny and McBratney 2007; Rossiter et al. 2017). Laborczi et al. (2016, 2019) adopted this approach to soil texture class prediction by using the two-dimensional distance between the centroids of the texture classes of the USDA texture triangle to weight the accuracy metrics; however, this should be viewed as a three-dimensional problem, with sand, silt, and clay of the composition representing coordinates, allowing the distance to be calculated between the centroids of the texture classes using the Euclidean distance.

To address some of these gaps, this study does the following:

1. highlights a workflow in which texture classes are predicted directly and indirectly using the random forest model;
2. reviews and contrasts model performance for the two approaches;
3. provides techniques based on the Euclidean distance and texture class separation for a quantitative comparison of texture class maps derived from the two approaches;
4. and reinforces the importance of adjusting categorical goodness of fit metrics for the magnitude of misallocations when evaluating model performance and reporting to aid in interpretability.

Materials and methods

Study area

The study area is the City of Ottawa, Canada, which is approximately 280 000 ha and is flanked to the north by the Ottawa River (Fig. 1). The study area is complex both in terms of bedrock geology and surficial (quaternary) geology. Bedrock is mainly of the Paleozoic era, dominated by dolomite and limestone with smaller areas of shales and sandstones, with the exception of a large Precambrian unit in the northwest of the study area (Schut and Wilson 1987). The southwest portion of the study area is characterized by the Smith Falls limestone plain and dominated by thin till veneers over bedrock (Schut and Wilson 1987). The areas east of the Smith Falls plain were mostly all subjected to inundation from the Cham-

Fig. 1. Ottawa study area with sample locations (triangles) over the digital elevation model. The inset map shows the Ottawa study area (red polygon) in relation to Ontario, Canada. Topographic base map courtesy of the Ontario GeoHub, Ontario Ministry of Natural Resources and Forestry. Map projection: NAD83, Lambert Conformal Conic. [Colour online]
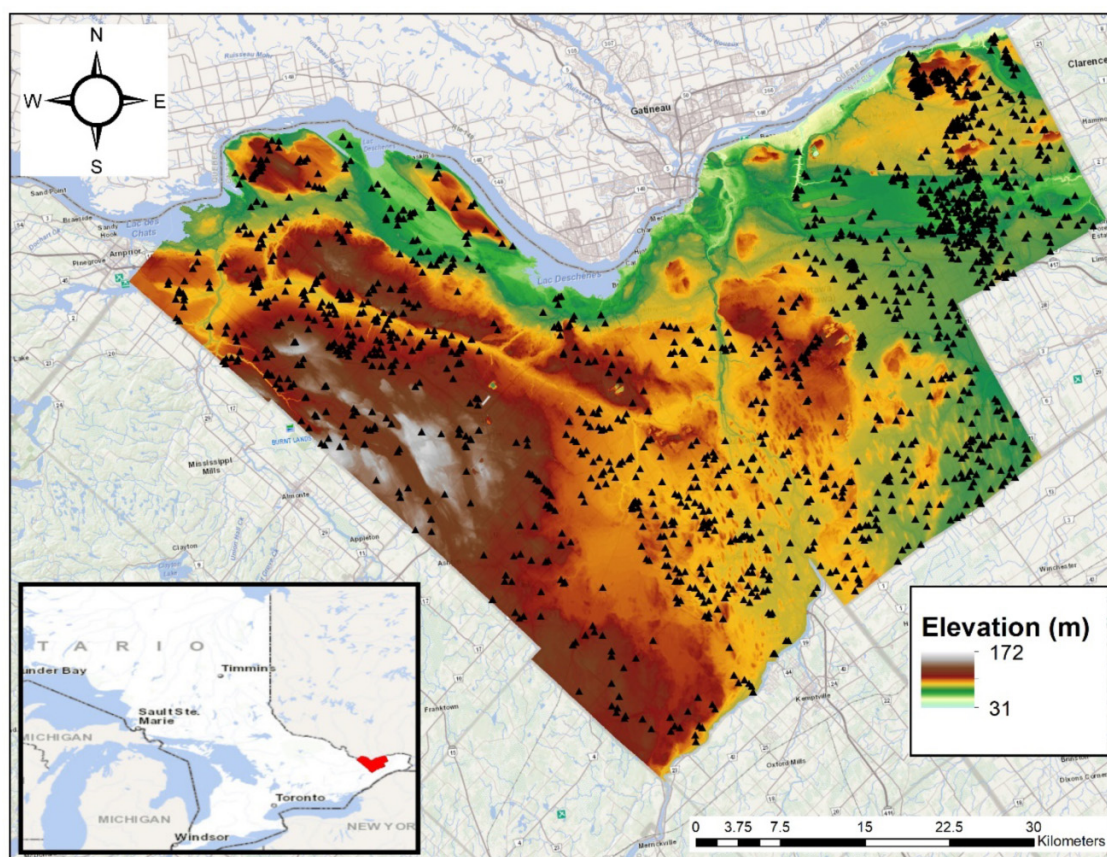
plain Sea after deglaciation, resulting in a complex distribution of clay plains, beach ridges, marine sand deposits, and estuarine deposits, with large organic deposits then accumulating in low-lying areas. For a detailed description of the soils and quaternary and bedrock geology in the study area, see Schut and Wilson (1987), Bélanger et al. (1995), and MacDonald and Harrison (1979).

Recent soil survey data collection efforts (2016–2019) completed by the Ontario Ministry of Agriculture, Food and Rural Affairs to update the soil maps for the study area resulted in the collection of 1622 soil profiles described and sampled based on pedogenic horizons. Sample sites were selected by a combination of conditioned Latin hypercube sampling (Minasny and McBratney 2006), expert knowledge-based sampling, and opportunistic sampling where sample sites were not accessible and alternative sites had to be identified in the field.

Soil analysis and data processing

Soil samples were analysed for particle size distribution (sand, silt, and clay content) using the pipette method with hydrogen peroxide pre-treatment for the removal of organic matter (Sheldrick and Wang 1993). Soil profile data were harmonized to the GlobalSoilMap.net standard depth intervals of 0–5, 5–15, 15–30, 30–60, and 60–100 cm (Arrouays et al.

2014). The equal area quadratic spline approach was used to harmonize each soil profile from the field-described horizon thicknesses to the standard depth intervals using the "easpline" function from the "ithir" package (Bishop et al. 1999; Malone 2018). Two depths were then selected, 5–15 and 30–60 cm, for modelling in this study. Shallow soils were encountered in the study area, resulting in only 1525 sites with particle size data at the 30–60 cm depth interval. Particle size distribution data are compositional (sand + silt + clay = 100%), and since the spline procedure is applied independently to each particle size, unity of the composition is not guaranteed. For this reason, spline-estimated values of sand, silt, and clay for each horizon were then normalized to a sum of 100 to enforce the compositional nature of the data using eq. 1:

$$(1) \quad nPSF = \frac{PSF}{sand + silt + clay} \times 100$$

where nPSF is the normalized PSF (sand, silt, or clay), PSF is the spline-estimated particle size fraction (sand, silt, or clay), and sand, silt, and clay are the spline-estimated values. The error introduced by the spline procedure was minimal. For the 5–15 cm depth interval, the sum of the spline-estimated PSFs was outside the range of 99%–101% in only two cases. Similarly, for the 30–60 cm depth, only 10 observations fell out-

side the range of 99%–102%. Once complete, the texture class, as per the texture triangle of the Canadian System of Soil Classification (SCWG 1998), was assigned to each record using the "oss.texture" function in the "onsoilsurvey" package in the R programming language (R Core Team 2020; Saurette 2021).

Environmental covariates

A total of 64 continuous covariates and three categorical covariates were considered for inclusion in the predictive models (Supplementary Table S1). The continuous covariates can be grouped into categories based on their provenance. The vast majority of the covariates (55) were terrain derivatives generated from a LiDAR-derived, 10 m resolution digital terrain model using either SAGA-GIS implemented in the "rsaga" package (Brenning et al. 2018; Conrad et al. 2015) or WhiteboxTools implemented in the "whitebox" package (Lindsay 2018; Wu 2019). Of the remaining continuous covariates, six were gamma-radiometric data (Natural Resources Canada 2019) and three were computed from a time series of cloud-free normalized difference vegetation index (NDVI) images generated from Sentinel-2 imagery captured from June to September in each of the 2017–2020 growing seasons using Google Earth Engine (median, maximum, and standard deviation). The three categorical covariates were quaternary geology (six classes), bedrock geology (four classes), and soil order (five classes), the latter generated from the legacy soil map (Ontario Geological Survey 2010, 2011; Ontario Ministry of Agriculture, Food and Rural Affairs 2019; Schut and Wilson 1987). All covariates were resampled to the same resolution as the digital terrain model (10 m) using bilinear resampling in the case of continuous covariates (gamma-radiometric and NDVI) and nearest neighbour resampling in the case of categorical covariates (quaternary geology, bedrock geology, soil order). Both resampling techniques were implemented using the "resample" function from the "raster" package (Hijmans 2021). Details for each of the covariates are outlined in the Supplementary Materials (Supplementary Table S1).

To reduce the number of environmental covariates and mitigate potential overfitting of the predictive models due to the collinearity of the predictors, the variance inflation factor (VIF) technique was applied to the set of 64 continuous covariates (Neter et al. 1983). The VIF (eq. 2) calculates how much an environmental covariate's (or regressor's) variability can be explained by the remaining covariates (regressors) in the regression model (Craney and Surles 2002):

$$(2) \quad \text{VIF}_j = \frac{1}{1 - r_j^2}$$

where r_j^2 is the coefficient of determination from fitting a linear regression between the jth independent variable and all other independent variables. The process is repeated until only covariates above a selected threshold remain. Commonly used thresholds for VIF analysis are five and 10 (O'brien 2007; Pourghasemi et al. 2017); we selected the lower cutoff to be more conservative and retain fewer covariates. Using a threshold VIF value of five, 23 of the 64 continuous covariates were retained as predictors for modelling. The VIF procedure was implemented using the "oss.seqVIF" in the

"onsoilsurvey" package (Saurette 2021). Categorical covariates were encoded using one hot encoding (Kuhn 2019; Kuhn and Johnson 2013) and were all retained for modelling.

Predictive modelling approaches

Soil texture class maps were generated using two approaches: direct prediction of texture class using categorical (classification) prediction, and indirect prediction where sand, silt, and clay content were predicted and then converted to texture class. A training data matrix was developed for each of the two depth increments (5–15 and 30–60 cm) by extracting the covariate values to the sampling locations. Each matrix contained the sample site identifier, the sand, silt, and clay content, the texture class, and the environmental covariate values. For each depth, the data were split into training (70%) and testing (30%) partitions using random sampling to ensure direct and indirect prediction approaches used identical input data. Detailed descriptions of the direct and indirect modelling approaches are outlined below.

Direct prediction

In the direct prediction approach, the random forest algorithm (Breiman 2001) was trained using the "caret" package (Kuhn 2018). Repeated (five repeats) 10-fold cross-validation was used to train the model using only the training data partition (70%) and select the optimal model hyperparameter (*mtry*). The final model from the repeated cross-validation was evaluated using the independent test data partition. The optimal *mtry* was then used to train the final random forest model using the entire data set, which was subsequently used to generate the predicted maps of the soil texture class.

Indirect prediction

In the indirect prediction approach, sand, silt, and clay content were modelled using the random forest algorithm. Soil PSFs are compositional (sum to 100) and therefore are modelled as a composition, typically using a log-ratio transformation. In this study, the ALR was used to retain the compositional nature of the three PSFs, therefore reducing the three-part composition to two parts. We selected the silt component of the composition as the denominator for the ALR calculations; the two remaining components were named alrClay (eq. 3) and alrSand (eq. 4)

$$(3) \quad \text{alrClay} = \ln\frac{\text{clay}}{\text{silt}}$$

$$(4) \quad \text{alrSand} = \ln\frac{\text{sand}}{\text{silt}}$$

The random forest model was therefore trained on the ALR-transformed clay (alrClay) and ALR-transformed sand (alrSand). Repeated (five repeats) 10-fold cross-validation was used to train each of the two models (alrSand and alrClay), at each depth, using the same 70% training partition as the direct prediction method, to select the optimal *mtry* hyper-

parameter values for each model. The optimal models were evaluated using the test data set on back-transformed values of sand, silt, and clay. The final models were trained using the optimal hyperparameters and the entire data set to predict maps of alrSand and alrClay. These maps were then back-transformed using the inverse ALR to generate maps of sand, silt, and clay content. Finally, these maps were then converted to texture class maps on a pixel-by-pixel basis with the "oss.texture.r" function in the "onsoilsurvey" package (Saurette 2021).

Model evaluation

The direct prediction models were evaluated using the overall accuracy and Cohen's kappa coefficient calculated from the confusion matrix generated from the observed and predicted values for the test data set. The rows of the confusion matrix represent the predictions, whereas the columns represent the observations. The diagonal of the confusion matrix represents correct classifications, while the off-diagonal represents incorrect classifications or misallocations. The overall accuracy is calculated as the total correct predictions divided by the total number of observations (Congalton 2001). Cohen's kappa coefficient takes into consideration chance agreement between observations and predictions (Cohen 1960; Rossiter et al. 2017).

The indirect prediction models were evaluated in two different ways. First, the sand, silt, and clay predictions (continuous properties) were evaluated using Lin's concordance correlation coefficient (CCC; Lin 1989) and the root mean square error (RMSE) by back-transforming the residuals from the alrClay and alrSand training models to sand, silt, and clay content. Secondly, the sand, silt, and clay predictions were then used to determine the texture classes, which were then evaluated using the classification metrics described above for the direct prediction (i.e., overall accuracy and Cohen's kappa coefficient).

The overall accuracy and Cohen's kappa coefficient, however, do not consider the magnitude of a misclassification. For example, if an observation is classified as a clay and the model predicts a clay loam, this is less egregious than if the model predicted the sand. Rossiter et al. (2017) proposed the use of taxonomic distance to adjust the accuracy of soil class predictions and outlined four approaches to computing a weight matrix: expert opinion, numerical distance, hierarchical distance, and error loss function. This approach was adapted to adjust the accuracy statistics for the categorical predictions in this study using the Euclidean distance between the centroids of the texture classes. First, the coordinates of the centroids for each texture class were identified on the texture triangle, with sand, silt, and clay content representing the coordinates in three-dimensional space. The Euclidean distance between the centroids of each pair of texture classes (e.g., distance from clay to sand) was generated and stored as a distance matrix (Table 1). To convert the distances to weights, the matrix was normalized to a range of 0–1. Finally, since the normalized Euclidean distances indicate increasing dissimilarity as the distance value increases, the weights were recalculated as 1 − normalized distance. Therefore, in the weight matrix, correct classifications carry a weight of 1 and all misallocations carry a weight <1 proportional to the Euclidean distance between the classes (Table 2).

Comparing direct and indirect texture class predictions

Two approaches were used to compare the predicted soil texture maps generated from the direct and indirect methods: a quantitative assessment using the Euclidean distance, and a semiquantitative approach using the texture class separation, both implemented on a pixel-by-pixel basis. In the first approach, for each of the 5–15 and 30–60 cm depth intervals, the predicted soil texture class maps generated from the direct and indirect prediction methods are stacked (overlaid). Then, on a pixel-by-pixel basis, the Euclidean distance between the texture classes of the two maps is assigned based on the distance matrix previously calculated (Table 1). For example, for a given pixel, if the direct prediction approach was classified as a clay, and the indirect prediction approach was classified as a loam, the Euclidean distance assigned to this pixel would be 40. The output of this process is a new raster map where the pixel values represent the Euclidean distance between the direct and indirect prediction approaches. As such, a distance of zero would indicate no difference between the texture class predicted by both approaches, whereas a difference greater than zero indicates a difference in the prediction, with the magnitude of the difference increasing as distance (dissimilarity) increases.

The second approach is to generate a "texture class separation" map. Similar to the Euclidean distance approach described above, for each of the 5–15 and 30–60 cm depth intervals, the predicted soil texture class maps generated from the direct and indirect prediction methods are stacked (overlaid). Then, on a pixel-by-pixel basis, the texture class separation between the texture classes of the two maps is assigned. The "texture class separation" between two texture classes is simply the number of classes between the two textures in question following the most direct route through the texture triangle. For example, using the Canadian soil texture triangle, a heavy clay and a clay are adjacent on the texture triangle (share a boundary or vertex), and therefore are one class apart, whereas a heavy clay and a clay loam are two classes apart, with the clay appearing in between them. A "texture class separation" was assigned to all pairs of texture classes, and these values were used to assign, on a pixel-by-pixel basis, the "texture class separation" between the maps from the two approaches. This is a more intuitive representation of the difference between the maps: for example, a pixel with a value of 0 indicates no difference between the two approaches (e.g., both approaches predicted sand), and a pixel with a value of 3 indicates the difference between the two approaches was three texture classes (e.g., one approach predicted sand, the other loam).

Finally, we can use the standard categorical goodness of fit metrics to evaluate the agreement between the two approaches. Therefore, at each depth, the texture class maps from the direct and indirect approaches were overlaid, and

Table 1. Distance matrix providing the Euclidean distance between the texture class centroids for the 13 texture classes of the Canadian System of Soil Classification (SCWG 1998).

	HC	SiC	C	SC	SiCL	CL	SCL	Si	SiL	L	SL	LS	S
HC	0												
SiC	43	0											
C	28	32	0										
SC	50	60	31	0									
SiCL	59	17	42	65	0								
CL	49	32	22	34	32	0							
SCL	66	66	42	18	66	35	0						
Si	101	58	82	99	42	65	94	0					
SiL	80	41	57	72	25	39	66	27	0				
L	67	45	40	43	38	19	34	59	32	0			
SL	82	72	55	38	67	41	21	85	59	29	0		
LS	96	92	72	47	89	60	30	105	80	51	22	0	
S	105	104	82	55	101	72	41	118	93	64	34	13	0

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

Table 2. Weight matrix applied for correction of model evaluation metrics of texture class calculated from the Euclidean distance between centroid coordinates (sand, silt, and clay) of the texture classes from the Canadian System of Soil Classification (SCWG 1998).

	HC	SiC	C	SC	SiCL	CL	SCL	Si	SiL	L	SL	LS	S
HC	1.00												
SiC	0.63	1.00											
C	0.76	0.73	1.00										
SC	0.57	0.49	0.74	1.00									
SiCL	0.50	0.86	0.65	0.44	1.00								
CL	0.59	0.73	0.82	0.71	0.73	1.00							
SCL	0.44	0.44	0.65	0.85	0.44	0.70	1.00						
Si	0.14	0.51	0.31	0.16	0.64	0.44	0.21	1.00					
SiL	0.32	0.65	0.51	0.39	0.79	0.67	0.44	0.77	1.00				
L	0.43	0.62	0.66	0.64	0.68	0.84	0.71	0.50	0.73	1.00			
SL	0.30	0.39	0.53	0.67	0.43	0.66	0.82	0.28	0.50	0.75	1.00		
LS	0.19	0.22	0.39	0.60	0.25	0.49	0.74	0.11	0.32	0.57	0.82	1.00	
S	0.11	0.12	0.30	0.53	0.14	0.39	0.66	0.00	0.21	0.46	0.71	0.89	1.00

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

the weighted overall accuracy and kappa were calculated using the weight matrix, as described above.

Results

Particle size analysis

For both depths, sand content was the dominant particle size, with a mean of 40.4% and 39.3% for the 5–15 and 30–60 cm depths, respectively (Table 3). The silt was intermediate, and clay was the least important particle size averaging 25.8% and 29.2% across the 5–15 and 30–60 cm depths. This is not unexpected as the study area is dominated by loamy moranial deposits and sandy deposits, whereas the clay plains occupy much smaller areas, despite the clay content of those soils ranging as high as 84.9%.

Training and testing data

The results of the random sampling to create the training (70%) and testing data sets (30%) for the 5–15 and 30–60 cm depths show close agreement between the frequencies within each data set across the 11 texture classes represented in the data (Fig. 2). Sandy loam, loam, and clay loam texture classes dominated the 5–15 cm depth interval, whereas, at the deeper sampling interval, heavy clay and clay textures both increased significantly compared to the surface sampling. This can be attributed to many of the glaciomarine deposits, which were very high in clay content, having lighter-textured deposits at the surface. Commonly these soils were either medium to moderately fine at the surface and grade to clay or heavy clay in the subsoil. It should be noted that only 11 of the 13 classes from the Canadian soil texture triangle were represented in the data set; there were no observations,

Table 3. Descriptive statistics for sand, silt, and clay content for the full data set at each depth after harmonization to standardized depth intervals.

Depth	Particle size	Minimum	25%	Median	Mean	75%	Maximum
5–15 cm	Sand	0.0	16.8	36.0	40.4	63.6	96.0
	Silt	0.4	21.2	34.7	33.8	45.2	79.5
	Clay	0.0	12.7	22.1	25.8	36.3	83.1
30–60 cm	Sand	0.0	9.8	33.3	39.3	65.2	99.8
	Silt	0.0	19.9	32.3	31.5	42.8	79.5
	Clay	0.0	9.8	22.9	29.2	46.5	84.9

Fig. 2. Bar plot showing the distribution of texture classes for the training and testing data sets for the 0–15 cm (*a*) and 30–60 cm (*b*) depth intervals. HC, heavy clay; SiC, silty clay; C, clay; SiCL, silty clay loam; CL, clay loam; SCL, sandy clay loam; SiL, silt loam; L, loam; SL, sandy loam; LS, loamy sand; S, sand. [Colour online]
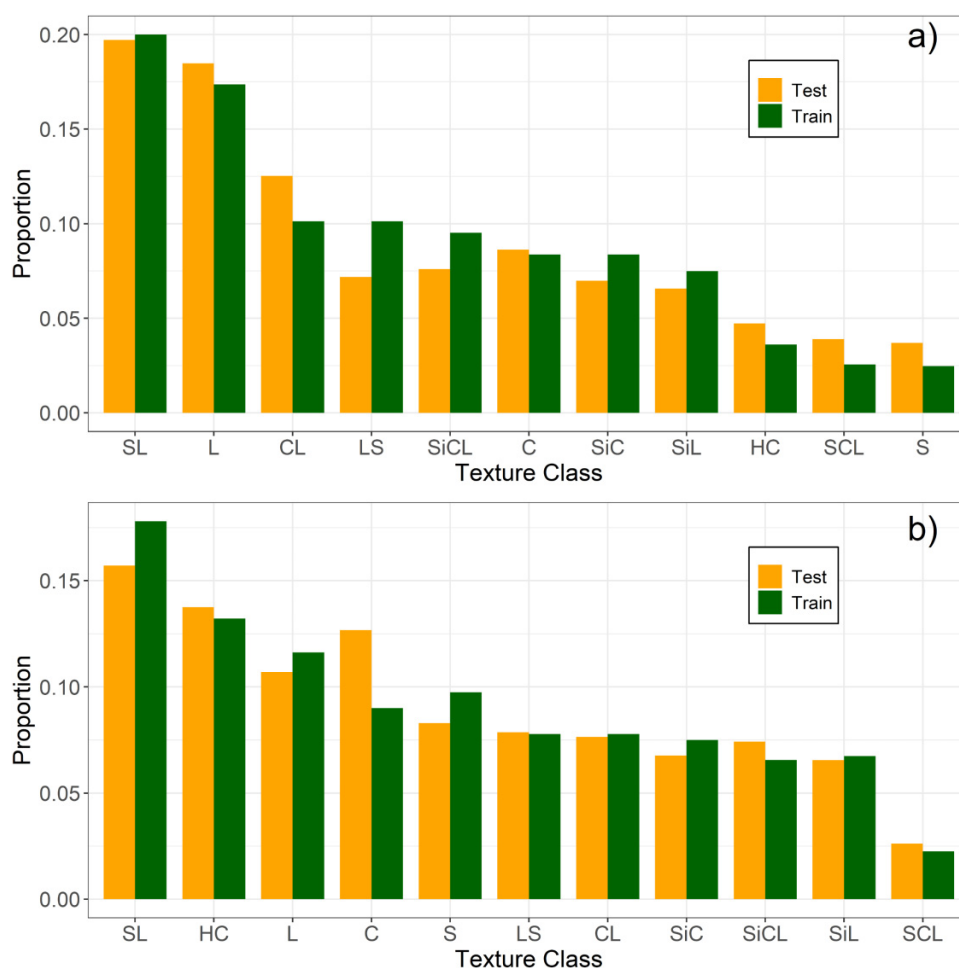
and therefore no predictions for silt and sandy clay texture classes.

For the continuous variables, sand, silt, and clay distributions between the training and test data sets were well aligned (Fig. 3). At the 5–15 cm depth, clay was positively skewed for both the training and test data sets, whereas both the sand and silt were multimodal, and the silt test data set had more evident multimodality than the training data. Regardless, the distributions are similar. At the 30–60 cm depth, clay was still positively skewed, but less so than at the surface depth, which was related to heavier materials in the subsoil

as described earlier, and the sand and silt distributions were still multimodal, but less so than at the surface depth. Again, the distributions between the training and test data sets were comparable, indicating both data sets were representative of the population of sampled sites.

Model evaluation — direct prediction of soil texture

Model evaluation metrics for the direct prediction of soil texture at the two depths showed fairly low overall accuracy and kappa scores (Table 4). For the 5–15 cm depth, an over-

Fig. 3. Comparison of sand, silt, and clay content for the training and test data sets at each of the 5–15 cm (*a*) and 30–60 cm (*b*) depths. [Colour online]
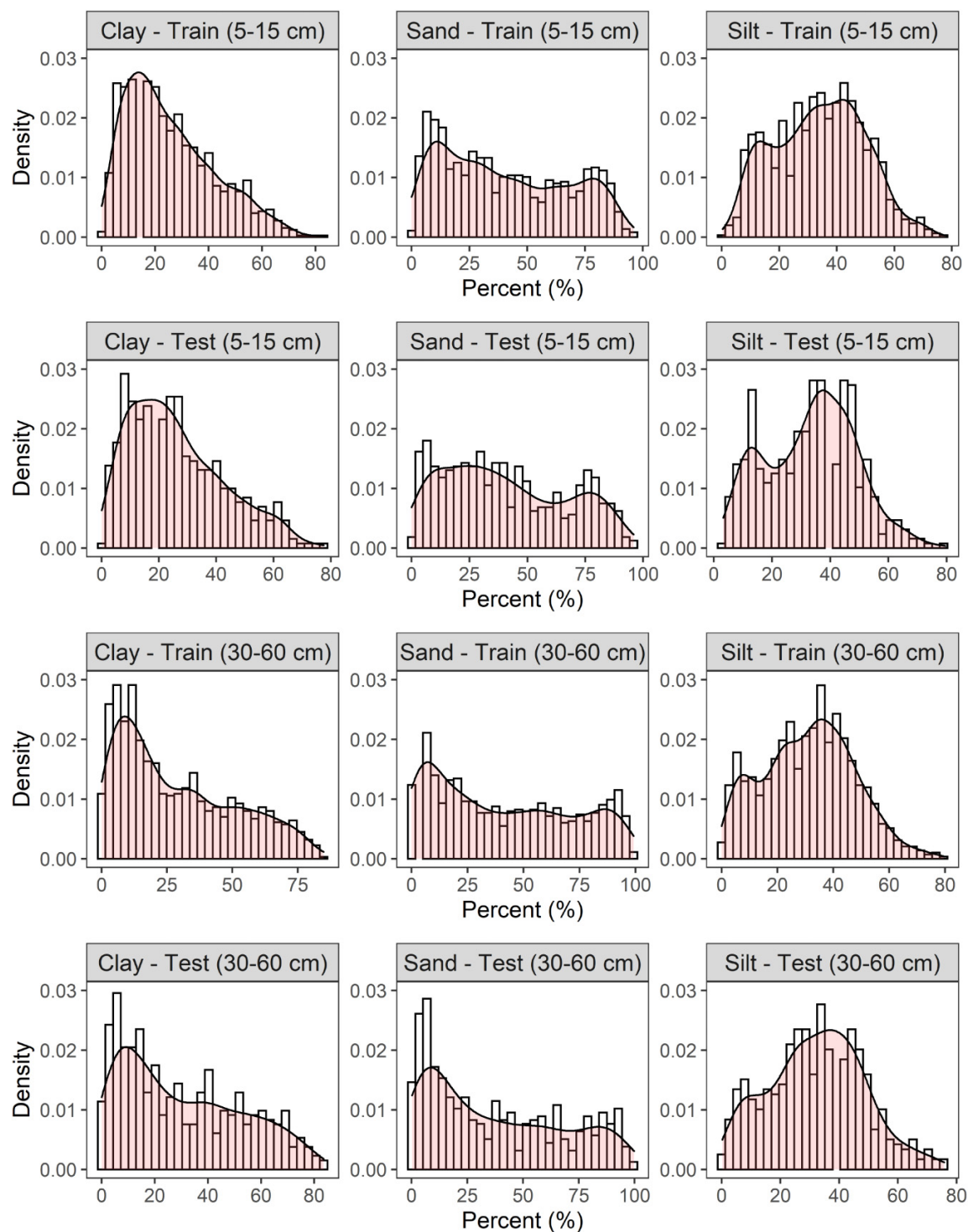
all accuracy of 36% and a kappa of 0.25, which is considered "fair" based on Landis and Koch (1977). The lower prediction depth, 30–60 cm, had marginally better results with an overall accuracy of 36% and kappa of 0.27 (Table 4). Producer's accuracy ranged from 0% for the sandy clay loam texture class to 61% for the sandy loam texture class in the 5–15 cm depth and from 0% for the sandy clay loam texture class to 66%

for the heavy clay and sand texture classes at the 30–60 cm depth. Similar ranges were seen for user's accuracy. Based on the confusion matrix for the 5–15 cm depth, it was clear that many of the misallocations were in an adjacent texture class (Supplementary Table S2). For example, for the sandy loam texture class, 30 of the 38 misallocations were either classified as loam (20) or loamy sand (10), both directly adjacent

Table 4. Overall accuracy, kappa, producer's accuracy, and user's accuracy for the direct and indirect prediction of soil texture class with the random forest model.

Depth interval (cm)	Overall accuracy	Kappa	Soil texture class										
			C	CL	HC	L	LS	S	SCL	SL	SiL	SiC	SiCL
			Producer's (user's) accuracy — direct prediction										
5–15	36	0.25	31 (27)	28 (38)	22 (46)	42 (34)	35 (34)	6 (34)	0 (0)	61 (40)	7 (40)	36 (34)	44 (38)
30–60	36	0.27	14 (25)	23 (22)	66 (49)	33 (28)	9 (24)	66 (50)	0 (0)	48 (31)	27 (32)	20 (29)	39 (52)
			Producer's (user's) accuracy — indirect prediction										
5–15	34	0.22	22 (22)	51 (29)	9 (100)	54 (35)	23 (43)	NA	6 (15)	46 (42)	0 (0)	18 (29)	38 (32)
30–60	28	0.19	20 (25)	26 (11)	20 (71)	54 (26)	17 (23)	27 (67)	0 (0)	39 (36)	0 (0)	46 (32)	27 (25)

Notes: HC, heavy clay; SiC, silty clay; C, clay; SiCL, silty clay loam; CL, clay loam; SCL, sandy clay loam; SiL, silt loam; L, loam; SL, sandy loam; LS, loamy sand; S, sand.

to the sandy loam texture class. Similarly, at the 30–60 cm depth, only eight correct predictions for clay texture were made, but 30 of the 50 misallocations were in adjacent texture classes (Supplementary Table S3). The same trend was apparent when reviewing the correlation matrices for the indirect prediction approach (Supplementary Tables S3 and S4).

In general, soils tended to become either finer in texture with depth in the areas of glaciomarine deposits or coarser in texture in the areas of moraimal and glaciofluvial deposits (Fig. 4, left). For example, in the northeast corner of the soil texture map, soils can be seen as silty clay to clay in the 5–15 cm depth interval and mostly as heavy clay in the deeper 30–60 cm interval, consistent with the glaciomarine deposits mapped in that area. The opposite trend was apparent in the central part of the study area, where soils grade from sandy loam to sand texture, and in the west and central areas, where materials tended to be coarser at depth, consistent with the moraimal and glaciofluvial materials.

Model evaluation — indirect prediction of soil texture

The random forest model performed well in predicting the continuous soil properties of sand, silt, and clay (Table 5). Lin's CCC was 0.74, 0.62, and 0.71 for the prediction of sand, silt, and clay, respectively, at the 5–15 cm depth, and 0.74, 0.59 and 0.67 at the 30–60 cm depth. Root mean square error ranged from 10.9% to 19.7% across all textures and both depths. Given that the prediction of the particle sizes was done as a composition using the ALR, for both depths, the sum of the bias was zero. In the case of the 5–15 cm depth, sand had a positive bias of 2.6%, while silt and clay had a negative bias of −2.1% and −0.4%, respectively. At the 30–60 cm depth, the clay had a positive bias (1.9%), while silt and sand had a negative bias (−1.7% and –0.2%, respectively). Particle size maps aligned well with known soil and material patterns in the study area (Fig. 5). Clay plains in the northeast, southeast, and northwest areas of the map had high silt and clay content, typical of deep water, glaciomarine deposits, while the south and west-central portions were dominated by sandier deposits, aligning with the shallow moraimal deposits over limestone bedrock.

Observed and predicted values for the validation data sets were plotted over the texture triangle (Fig. 6), commonly referred to as ternary plots. For both the 5–15 and 30–60 cm depth, it was clear that the random forest model "compresses" the distribution of the predictions towards the centre of the texture triangle (Fig. 6), as reported in Zhang et al. (2020). There were no observations in the sandy clay or silt texture classes in the 5–15 cm depth; however, there were two observations of sandy clay in the 30–60 cm depth. This was at odds with the training data used for the direct prediction approach and was a direct result of the depth harmonization of soil profiles using the "easpline" function. For the 5–15 cm depth, observations extended well into the heavy clay and the silt loam texture classes, but the predictions showed no observations classified as silt loam, and only a few classified as heavy clay (Fig. 6a). A similar trend was obvious in the ternary plots for the 30–60 cm depth where the observed values were spread widely throughout the texture triangle, whereas the predictions once again were in a fairly narrow band arching through the centre of the texture triangle (Fig. 6b).

Finally, after converting the sand, silt, and clay maps to texture class maps, the overall accuracy and kappa statistics were found to be worse than those of the direct prediction of texture class (Table 4). Overall accuracy was 34% for the 5–15 cm depth and 28% for the 30–60 cm depth, while kappa was 0.22 and 0.19 for the same two depths, respectively. Producer's and user's accuracies were similar to those of the direct prediction (Table 4).

Model evaluation — weighted accuracy metrics

Using the weight matrix (Table 2) computed from the Euclidean distance between texture classes significantly improved all accuracy metrics (Table 6). Overall accuracy for the direct prediction method increased to 82% and 78% for the 5–15 cm and 30–60 cm depths, representing increases of 128% and 117%, respectively. For the indirect prediction method, overall accuracy increased to 83% and 80% for the 5–15 and 30–60 cm depths, representing increases of 144% and 186%, respectively. Similar increases were seen for kappa, with increases from 0.25 to 0.53 and from 0.27 to 0.47 for the 5–15 and 30–60 cm depths for the direct prediction method, and from 0.22 to 0.56 and from 0.19 to 0.51 for the 5–15 and 30–60 cm depths for the indirect prediction method.

Fig. 4. Soil texture maps were generated from direct and indirect prediction techniques for the 5–15 and 30–60 cm depth intervals. HC, heavy clay; SiC, silty clay; C, clay; SiCL, silty clay loam; CL, clay loam; SCL, sandy clay loam; SiL, silt loam; L, loam; SL, sandy loam; LS, loamy sand; S, sand. [Colour online]
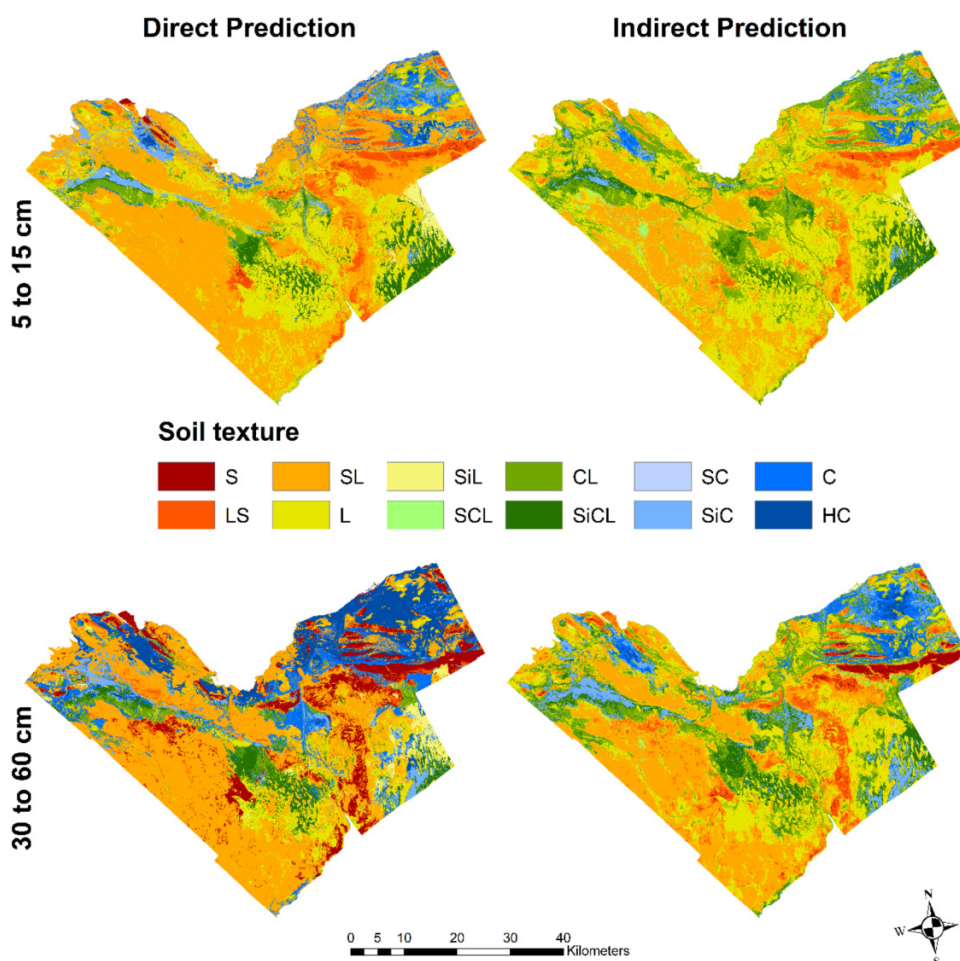
Table 5. Goodness of fit metrics from external validation of random forest models for sand, silt, and clay content.

Depth interval (cm)	Property	Adjusted R^2	CCC	RMSE	Bias
5–15	Sand (%)	0.59	0.74	17.1	2.6
	Silt (%)	0.45	0.62	11.6	-2.1
	Clay (%)	0.56	0.71	10.9	-0.4
30–60	Sand (%)	0.58	0.74	19.7	-0.2
	Silt (%)	0.40	0.59	12.6	-1.7
	Clay (%)	0.52	0.67	15.8	1.9

Notes: CCC, Lin's concordance correlation coefficient; RMSE, root mean square error.

## Comparison of direct and indirect texture class predictions

The spatial distributions of texture classes in the direct and indirect approaches were quite similar; however, there were noticeable differences (Fig. 4). For the 5–15 cm depth, areas of the loam texture class were more prominent in the southwest of the map for the indirect prediction when compared to the direct prediction. In the northeast corner of the study area, larger areas were predicted as sandy loam and loamy sand texture classes in the direct prediction approach. For the 30–60 cm depth, the differences were more evident than at the surface depth. For instance, the northeast corner of the study area was dominated by the heavy clay texture class in the direct approach, whereas the indirect approach yields clay and silty clay texture classes as dominant; in addition, the geographic extent of the clayey soils was larger in the direct approach. Soils predicted as sand texture class were more widespread in the direct prediction map, whereas the amount of loam texture class dominates the central portion of the map produced from the indirect prediction approach.

A more quantitative assessment can be discerned from the calculation of the Euclidean distance, on a pixel-by-pixel ba-

Fig. 5. Predicted maps of sand, silt, and clay content for the study area from the random forest model at the 5–15 and 30–60 cm depths. [Colour online]
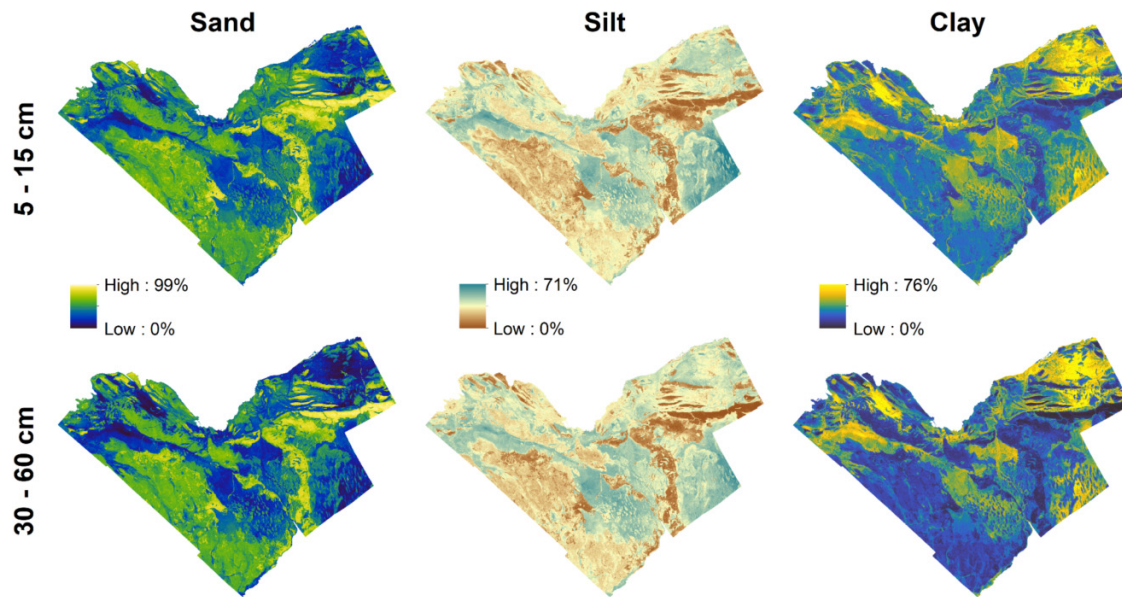
sis, between the maps generated from the two prediction approaches (Fig. 7, left). The mean Euclidean distance was smaller between the two maps at 5–15 cm (13.2) than that of the maps at 30–60 cm (17.6), indicating closer agreement between the two approaches at the surface depth compared to the subsoil depth. The spatial patterns of the Euclidean distance between maps were not consistent across the two depths of prediction. For example, along the northern edge of the study area in the centre of the map, the Euclidean distance was small at the 5–15 cm depth interval (<30); however, at the 30–60 cm interval, the Euclidean distance was much larger (>46). Other than this, most other areas saw an increase in the Euclidean distance when going from the 5–15 cm maps to the 30–60 cm maps. This same trend was apparent in the class separation between the maps generated from the direct and indirect approaches (Fig. 7, right). Overall, the texture class separation between the two approaches was dominantly 0 or 1 texture class; where the 5–15 cm depth interval was dominantly zero texture class difference, while the 30–60 cm depth was dominantly one texture class difference. For the 5–15 cm depth, 52.7% of the map was in perfect agreement, while another 42.1% was within one texture class, and less than 5% of the map was two texture classes apart (Table 7). For the 30–60 cm depth, 44.9% of the map was one texture class apart, 42.3% was in perfect agreement, and 11.4% had two texture classes difference (Table 7). Lastly, the weighted overall accuracy and kappa between the maps of the two approaches, an indication of the similarity between the maps, were 89% and 0.70 for the 5–15 cm depth and 85% and 0.63 for the 30–60 cm depth.

Discussion

In terms of texture class prediction, model performance metrics for the direct and indirect prediction approaches were quite similar and were at the low end of the range from results of similar studies. Maynard and Levi (2017) reported an overall accuracy of 67% and kappa of 0.53, while Gomez et al. (2019) reported an overall accuracy of 50% and kappa of 0.31, both studies achieving higher performance than our study using support vector machine models and time-series spectral data from Landsat TM and Sentinel-2, respectively. One reason for this might be the number of texture classes in the analysis; whereas our study included 10 texture classes, Maynard and Levi (2017) predicted six texture classes, while Gomez et al. (2019) predicted four texture classes in their respective study. Using random forest, Dharumarajan and Hegde (2022) achieved overall accuracies ranging from 50% to 65% and kappa scores ranging from 0.42 to 0.47 when predicting soil texture classes as per the GlobalSoilMap.net (Arrouays et al. 2014) depth interval specifications (0–5, 5–15, 15–30, 30–60, 60–100, and 100–200 cm). Our performance metrics were better aligned with Laborczi et al. (2016) who reported an overall accuracy of 29% and kappa of 0.15 (calculated from the confusion matrix reported in Laborczi et al. (2016)), with 11 texture classes.

With regard to the prediction of PSFs, R^2 ranged from 0.40 to 0.59 across the three PSFs and both depths. This is comparable to Poggio and Gimona (2017) who reported R^2 of 0.56–0.58 in their two-dimensional models of sand, silt, and clay, but better than their three-dimensional models for the 5–15 and 30–60 cm depths which ranged from 0.30 to 0.43. Root mean square error was within ranges from other studies (Liu et al. 2020; Malone and Searle 2021b; Pahlavan-Rad and Akbarimoghaddam 2018) for all three PSFs, but much higher than that reported in Amirian-Chakan et al. (2019), whose RMSE values ranged from 2.9% to 4.4%. This was almost an order of magnitude lower than these other studies, with the exception of Malone and Searle (2021a) who reported

Fig. 6. Ternary plots showing the distribution of observed and predicted values based on the indirect prediction approach in the 5–15 (*a*) and 30–60 cm (*b*) depths.
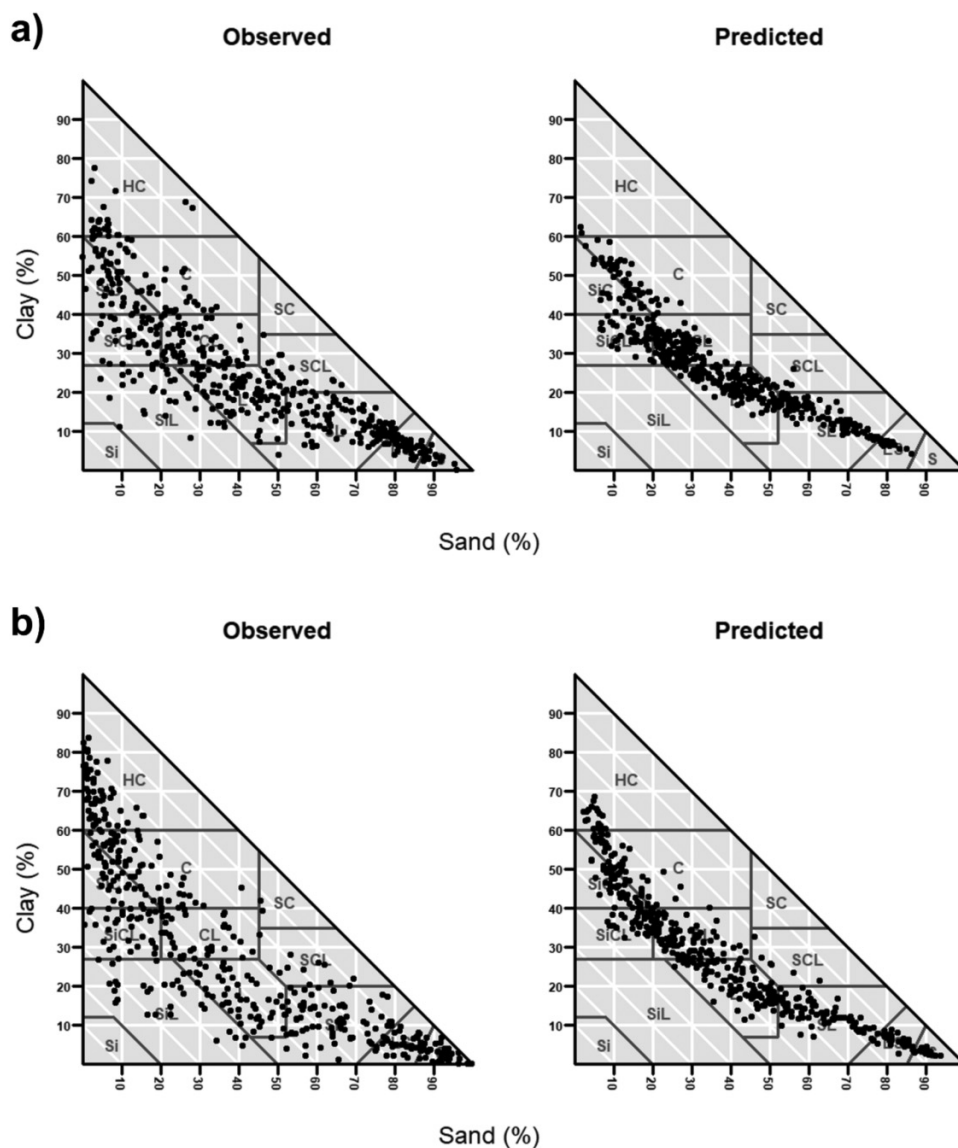
Table 6. Overall accuracy, kappa, producer's accuracy, and user's accuracy for the direct and indirect prediction of soil texture class with the random forest model adjusted based on the weight matrix calculated from the Euclidean distance between texture classes.

Depth interval (cm)	Overall accuracy	Kappa	Soil texture class											
			C	CL	HC	L	LS	S	SCL	SL	SiL	SiC	SiCL	SC
			Producer's (user's) accuracy — direct prediction											
5–15	82	0.53	77 (79)	83 (83)	75 (79)	85 (82)	84 (84)	79 (70)	0 (0)	89 (82)	70 (84)	82 (78)	79 (84)	NA
30–60	78	0.47	71 (75)	79 (80)	83 (80)	79 (77)	82 (63)	84 (86)	0 (0)	81 (77)	71 (77)	73 (72)	80 (86)	NA
			Producer's (user's) accuracy — indirect prediction											
5–15	83	0.56	80 (80)	91 (82)	73 (100)	90 (83)	81 (90)	NA	75 (82)	86 (86)	0 (0)	80 (80)	82 (82)	NA
30–60	80	0.51	79 (79)	82 (75)	72 (71)	89 (78)	82 (85)	82 (96)	66 (68)	82 (83)	0 (0)	87 (79)	81 (79)	69 (0)

Notes: HC, heavy clay; SiC, silty clay; C, clay; SC, sandy clay; SiCL, silty clay loam; CL, clay loam; SCL, sandy clay loam; SiL, silt loam; L, loam; SL, sandy loam; LS, loamy sand; S, sand.

Fig. 7. Euclidean distance and soil texture class separation maps for the 5–15 and 30–60 cm depth intervals. [Colour online]
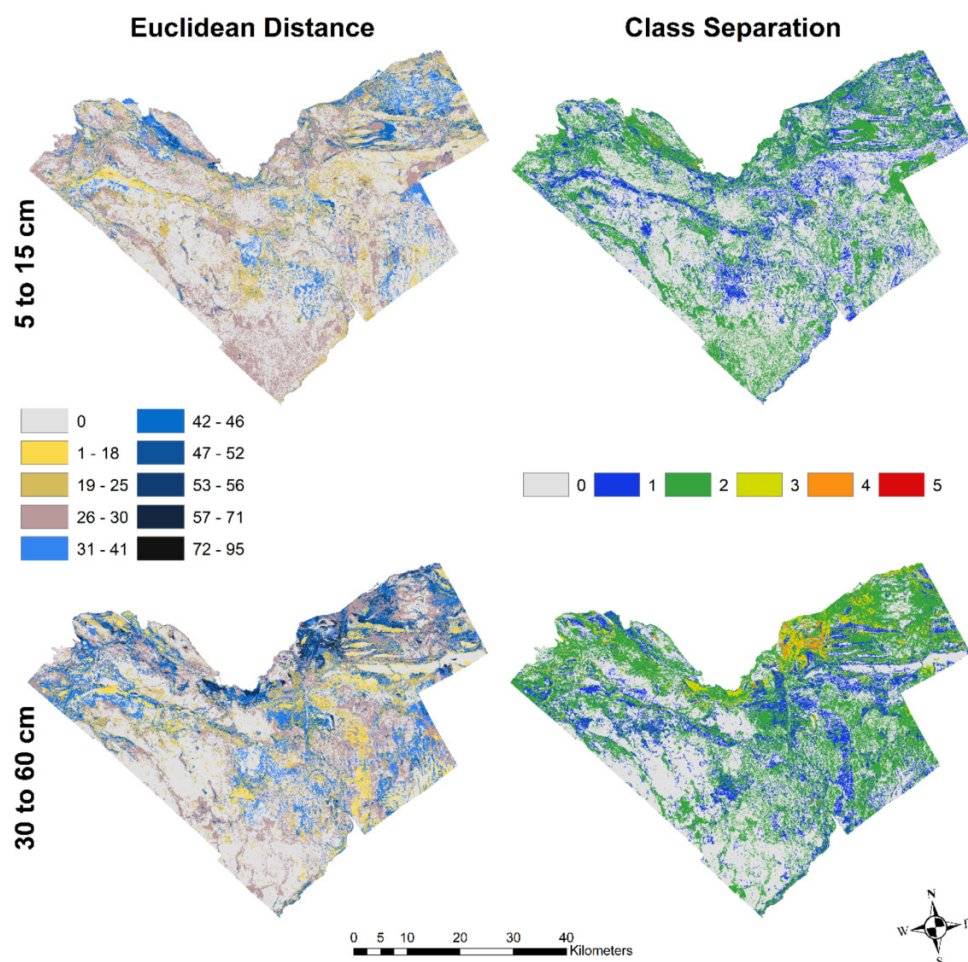
Table 7. Summary statistics for the class separation between the direct and indirect prediction approaches at the 5–15 and 30–60 cm depths

Texture class separation	5–15 cm			30–60 cm		
	Raster cells	ha	%	Raster cells	ha	%
0	14 810 858	148 109	52.7	11 893 001	118 930	42.3
1	11 834 023	118 340	42.1	12 610 127	126 101	44.9
2	1377 216	13 772	4.9	3214 655	32 147	11.4
3	46 787	468	0.2	347 823	3478	1.2
4	16 259	163	0.1	19 119	191	0.1
5				418	4	0.0
Total	**28 085 143**	**280 851**	**100.0**	**28 085 143**	**280 851**	**100.0**

RMSE from 5.2% to 7.7% for silt. However, this may be explained by the tight range of silt content in Malone and Searle (2021a) which shows most observations in their study ranging from 0% to 30% silt. With regard to CCC, a more robust measure of the fidelity between observed and predicted values that corrects for bias, our model validation suggested better performance ranging from 0.59 to 0.74, when compared to Malone and Searle (2021a) and Liu et al. (2020). Our model performance metrics consistently declined with increasing soil depth, which was also the case in many other studies

(Adhikari et al. 2013; Liu et al. 2020; Poggio et al. 2021; Poggio and Gimona 2017) and has been attributed to higher variability in training data, fewer observations available at depth, and the use of covariates which are reflective of surface conditions and therefore with a weaker relationship with deeper soil layers. The latter two were likely in this study, given the reduced number of sites with data at the 30–60 cm depth due to shallow soils over bedrock, and the reliance on environmental covariates mostly derived from a digital terrain model and spectral indices.

One advantage of the indirect prediction of texture class reported by Zhang et al. (2020) was that conversion of PSFs to texture class has the potential to add more texture classes since, in the direct prediction, the models are limited to those classes present in the training data. This advantage, however, is not likely to be realized. As noted in Zhang et al. (2020), and observed in our study, the contraction of the predictions towards the centre of the texture triangle would likely prevent the indirect methods from adding classes not already present in the training data.

Intuitively, model confusion is likely to be higher between classes that are more closely related, and lower between classes that are distant, as is the case with texture classes. Dharumarajan and Hedge (2022) identified higher confusion between adjacent texture classes. For example, at the 5–15 cm depth, only 10% of the loam observations were correctly classified whereas the majority were predicted in adjacent classes of clay loam and sandy clay loam; and at the 30–60 cm depth, none of the loam observations were correctly classified, but they were all classified in adjacent classes. Zhang et al. (2020) noted the highest confusion between sandy loam and loamy sand texture classes in their study area. Our study reveals the same trend, where the misallocations are biased towards adjacent classes (Supplementary Tables S1–S4). In our evaluation metrics for classification models, overall accuracy and kappa, all incorrect classifications were of the same magnitude and treated as equally serious. Rossiter et al. (2017) posited that misallocations are not all equally serious and that evaluation metrics could be adjusted based on the taxonomic distance between classes, using soil classification as an example. This can be applied to soil texture classification, where the distance between texture classes can be determined based on the distance between texture classes (Laborczi et al. 2016, 2019). Despite accounting for misallocations using a weight matrix based on the distance between texture classes of the USDA texture triangle, Laborczi et al. (2016) achieved low overall accuracy and kappa, 29% and 0.15, respectively, but did not report the unweighted results. Comparatively, the overall accuracy of our study ranged from 78% to 83% and kappa ranged from 0.47 to 0.56; however, without the unweighted overall accuracy and kappa from the previous study, a comparison is not possible. Laborczi et al. (2019) assessed the similarity between texture class predictions that were using different modelling approaches using both unweighted and weighted accuracy metrics. They saw significant improvements in overall accuracy and kappa: for example, for one comparison, they noted overall accuracy and kappa increased from 39% to 85% and from 0.27 to 0.42, respectively. Although not the same application, the magnitude of the increase in the metrics was similar to those from our study.

It is interesting that the overall accuracy and kappa were initially lower for the indirect prediction than for the direct prediction in our study, but that after computing the weighted scores, the indirect prediction approach performance metrics indicate this approach was superior. This means that overall the direct prediction of texture class had fewer misallocations, but that those misallocations were more serious (i.e., farther away from the observed texture class) than those of the indirect method, and this was not at all reflected in the unweighted performance metrics. This was unexpected but could prove useful when comparing different modelling approaches in the future.

There were some differences in the calculation of the weight matrix used in the two studies. First, we used the Canadian texture triangle, which has an additional texture class, heavy clay, which would explain a difference in the weights between the clay texture class and all other classes from those computed by Laborczi et al. (2016, 2019). In addition, instead of calculating the two-dimensional distances between the centroids of the texture classes of the texture triangle, we computed the Euclidean distance between the centroids of the texture classes using their sand, silt, and clay values. This generally increased the distance between texture classes that had a significant amount of silt and the other texture classes. For example, the largest distance in the matrix from Laborczi et al. (2016, 2019) was 83, between sand and silt, whereas the Euclidean distance we calculated was 118. Distances between the silt texture class and silt loam, loam, sandy loam, loamy sand, and sand texture classes increased by 23%, 40%, 30%, 42%, and 42% when using the Euclidean distance compared to the distances calculated directly from the two-dimensional texture triangle, which was then also reflected in the weight matrix, meaning that our approach applied a higher cost to misallocation of texture class.

A comparison of texture class maps predicted using different approaches can also be achieved using the computed distance matrix. Laborczi et al. (2019) compared a soil texture class map generated from classifying PSF maps directly from sand, silt, and clay predictions for 0–30 cm depth interval, to a map generated by calculating weighted average sand, silt, and clay for 0–5, 5–15, and 15–30 cm predictions and then classifying to soil texture. The authors report that 68% of the maps were in perfect agreement (predicted the same class), while the remaining 32% of the maps showed a mostly minor difference, and small isolated areas showed a major difference; however minor and major difference categories were not defined. Furthermore, when comparing the two predictions, the validation points showed a weighted overall accuracy of 93% and kappa of 0.76. In terms of perfect agreement between the maps, our predictions showed 53% and 42% agreement for the 5–15 and 30–60 cm depths, slightly lower than the previous study. Finally, in terms of weighted performance metrics, overall accuracy was 89% and 85%, and kappa was 0.70 and 0.63 for the 5–15 and 30–60 cm depths, which was also a bit lower than the results from Laborczi et al. (2019), but within the same range.

The comparison maps we produced, Euclidean distance and texture class separation, we feel, were more useful in interpreting the similarities and differences between two categorial maps of soil texture. The Euclidean distance map could be useful when interpreting the models to understand prediction uncertainty. Although not done as part of this assessment, prediction uncertainty was likely related to the areas with a larger Euclidean distance between the two maps. Furthermore, these maps could be useful for model diagnostics in future work by evaluating variable importance to understand which covariates are being leveraged the most by

the two approaches. The class separation map was certainly the most intuitive way to assess the differences between the maps, converting a continuous scale of distance to a class easily understood by a pedologist. We decided to assign a class separation value between the texture classes for creating these maps; however, another approach might be to calculate the mean Euclidean distance between all texture class pairs with one class of separation and reclassify the Euclidean distance map using the calculated mean. This would partially account for the relative uneven texture class sizes of the triangle. For example, the silt loam, heavy clay, clay, and sandy loam texture classes are quite large, and when combined they occupy 57% of the texture triangle; the relative size of the texture classes could also be considered in future studies.

Using a weight matrix to compute adjusted model performance metrics, which almost guarantees improvements by providing partial scores for misallocations, might seem like a self-fulfilling exercise for the producer of the map(s). Rossiter et al. (2017) pointed out that from the producer's perspective, it might provide a more realistic assessment of their work, and from the user's perspective, a refined assessment of the reliability of the map. In our study, when comparing the unweighted performance metrics from the direct categorial prediction (e.g., overall accuracy and kappa ranging from 28% to 36% and from 0.19 to 0.27) to the continuous predictions of sand, silt, and clay (CCC ranging from 0.59 to 0.74), although not directly comparable, one would note the relatively poor performance of the categorial models. The use of the weight matrix in essence allows the producer to leverage continuous data (Euclidean distance) to provide a more realistic evaluation for the user, improving the usefulness and interpretability of the final map.

Conclusion

In this study, we compared direct and indirect approaches for the prediction of soil texture classes using the random forest machine learning algorithm. Both approaches resulted in soil texture maps that were quite similar and had small subtle but important differences. With regard to validation of the approaches, both yielded similar overall accuracy and kappa scores. In both cases, however, performance metrics were quite low compared to most studies. We then demonstrated the use of a weight matrix, based on the Euclidean distance between the centroids of the texture classes of the texture triangle, to account for the magnitude of misallocations in the texture class predictions. Interestingly, the indirect approach performance metrics improved more than those of the direct approach, meaning the direct approach had fewer misallocations overall, but those misallocations were more serious than those of the indirect approach. In general, the adjusted performance metrics provide a better estimate of the reliability of the texture map being generated. Finally, we showed how the distance matrix can be used to compare multiple soil texture class maps by either calculating the Euclidean distance between the maps or by converting to class separation maps, which is more intuitive for soil scientists. Based on the analysis, it appears that the indirect approach is a superior option for predicting texture class, which is not surprising since it is a more data-rich approach that uses sand, silt, and clay data as opposed to simply using a texture class, and results in a texture class map that aligns with the predicted maps of sand, silt, and clay. However, in the absence of particle size data, the direct prediction of texture class is a suitable alternative.

Acknowledgements

Article information

History dates

Notes

This paper is part of a Collection entitled "Advances in Soil Survey & Classification in Canada".

Copyright

Data availability

Primary research data may be requested from the corresponding author.

Author information

Author ORCIDs

Daniel D. Saurette https://orcid.org/0000-0002-1971-1238

Author notes

Author Daniel D. Saurette served as a Guest Editor at the time of manuscript review and acceptance; peer review and editorial decisions regarding this manuscript were handled by Angela Bedard-Haughn.

Competing interests

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

Supplementary material

Supplementary data are available with the article at https://doi.org/10.1139/cjss-2022-0040.

References

Adhikari, K., Kheir, R.B., Greve, M.B., Bøcher, P.K., Malone, B.P. Minasny, B., et al. 2013. High-Resolution 3-D mapping of soil texture in Denmark. Soil Sci. Soc. Am. J. **77**: 860–876. doi:10.2136/sssaj2012.0275.

Amirian-Chakan, A., Minasny, B., Taghizadeh-Mehrjardi, R., Akbarifazli, R., Darvishpasand, Z., and Khordehbin, S. 2019. Some practical aspects of predicting texture data in digital soil mapping. Soil Tillage Res. **194**: 104289. doi:10.1016/j.still.2019.06.006.

Arrouays, D., Grundy, M.G., Hartemink, A.E., Hempel, J.W., Heuvelink, G.B.M. Hong, S.Y., et al. 2014. Chapter three - GlobalSoilMap: toward a fine-resolution global grid of soil properties, *In* Advances in agronomy. *Edited by* D.L. Sparks. Academic Press, New York. pp. 93–134. doi:10.1016/B978-0-12-800137-0.00003-0.

Bélanger, J.R., Moore, A., Prégent, A., and Richard, H. 1995. Surficial geology - Ottawa, Ontario-Quebec (31 G/5) (No. 1506A). Geological Survey of Canada, Ottawa, Ontario.

Bishop, T.F.A., McBratney, A.B., and Laslett, G.M. 1999. Modelling soil attribute depth functions with equal-area quadratic smoothing splines. Geoderma, **91**: 27–45. doi:10.1016/S0016-7061(99)00003-8.

Breiman, L. 2001. Random forests. Mach. Learn. **45**: 5–32. doi:10.1023/A:1010933404324.

Brenning, A., Bangs, D., and Becker, M. 2018. RSAGA: SAGA geoprocessing and terrain analysis. R package version 1.3.0.

Cohen, J. 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. **20**: 37–46. doi:10.1177/001316446002000104.

Congalton, R.G. 2001. Accuracy assessment and validation of remotely sensed and other spatial information. Int. J. Wildland Fire. **10**: 321–328. doi:10.1071/WF01031.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E. Gerlitz, L., et al. 2015. System for automated geoscientific analyses (SAGA) v.2.1.4. Geosci. Model Dev. **8**: 1991–2007. doi:10.5194/gmd-8-1991-2015.

Craney, T.A., and Surles, J.G. 2002. Model-dependent variance inflation factor cutoff values. Qual. Eng. **14**: 391–403. doi:10.1081/QEN-120001878.

Dharumarajan, S., and Hegde, R. 2022. Digital mapping of soil texture classes using random forest classification algorithm. Soil Use Manag. **38**: 135–149. doi:10.1111/sum.12668.

Gomez, C., Dharumarajan, S., Féret, J.-B., Lagacherie, P., Ruiz, L., and Sekhar, M. 2019. Use of Sentinel-2 time-series images for classification and uncertainty analysis of inherent biophysical property: case of soil texture mapping. Remote Sens, **11**: 565. doi: 10.3390/rs11050565.

Greenacre, M. 2021. Compositional data analysis. Annu. Rev. Stat. Its Appl. **8**: 271–299. doi:10.1146/annurev-statistics-042720-124436.

Hijmans, R.J. 2021. raster: geographic data analysis and modeling. R package version 3.4-13.

Jamagne, M. 1967. Bases et techniques d'une cartographie des sols. Ann. Agron. **18**: 142.

Krzic, M., Walley, F.L., Diochon, A., Paré, M.C., and Farrell, R.E.(Eds.) 2021. Digging into Canadian soils: An introduction to soil science. Canadian Society of Soil Science, Pinawa, MB.

Kuhn, M. 2018. caret: classification and regression training.

Kuhn, M. 2019. The caret Package.

Kuhn, M., and Johnson, K. 2013. Applied predictive modeling. Springer, New York.

Laborczi, A., Szatmári, G., Kaposi, A.D., and Pásztor, L. 2019. Comparison of soil texture maps synthetized from standard depth layers with directly compiled products. Geoderma, **352**: 360–372. doi:10.1016/j.geoderma.2018.01.020.

Laborczi, A., Szatmári, G., Takács, K., and Pásztor, L. 2016. Mapping of topsoil texture in Hungary using classification trees. J. Maps. **12**: 999–1009. doi:10.1080/17445647.2015.1113896.

Landis, J.R., and Koch, G.G. 1977. The measurement of observer agreement for categorical data. Biometrics, **33**(1): 159–174. doi:10.2307/2529310.

Levi, M.R. 2017. Modified centroid for estimating sand, silt, and clay from soil texture class. Soil Sci. Soc. Am. J. **81**: 578–588. doi:10.2136/sssaj2016.09.0301.

Lin, L.I.-K. 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics, **45**: 255. doi:10.2307/2532051. PMID: 2720055.

Lindsay, J. 2018. WhiteboxTools user manual (user manual). University of Guelph, Guelph, Ontario.

Liu, F., Zhang, G.-L., Song, X., Li, D., Zhao, Y. Yang, J., et al. 2020. High-resolution and three-dimensional mapping of soil texture of china. Geoderma, **361**: 114061. doi:10.1016/j.geoderma.2019.114061.

MacDonald, G., and Harrison, J.E. 1979. Generalized bedrock geology, Ottawa-Hull, Ontario and Quebec.

Malone, B. 2018. ithir: Soil data and some useful associated functions.

Malone, B., and Searle, R. 2021a. Updating the Australian digital soil texture mapping (Part 1∗): re-calibration of field soil texture class centroids and description of a field soil texture conversion algorithm. Soil Res. **59**: 419–434. doi:10.1071/SR20283.

Malone, B., and Searle, R. 2021b. Updating the Australian digital soil texture mapping (Part 2∗): spatial modelling of merged field and lab measurements. Soil Res. **59**: 435–451. doi:10.1071/SR20284.

Maynard, J.J., and Levi, M.R. 2017. Hyper-temporal remote sensing for digital soil mapping: characterizing soil-vegetation response to climatic variability. Geoderma, **285**: 94–109. doi:10.1016/j.geoderma.2016.09.024.

McBratney, A., Mendonça Santos, M.L., and Minasny, B. 2003. On digital soil mapping. Geoderma, **117**: 3–52. doi:10.1016/S0016-7061(03)00223-4.

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

Minasny, B., and McBratney, A.B. 2007. Incorporating taxonomic distance into spatial prediction and digital mapping of soil classes. Geoderma, **142**: 285–293. doi:10.1016/j.geoderma.2007.08.022.

Natural Resources Canada 2019. Geoscience data repository for geophysical data. Magnetic-Radiometric-EM Datasets.

Neter, J., Wasserman, W., and Kutner, M.H. 1983. Applied linear regression models. Richard D Irwin, Inc., Honeywood, Illinois.

O'brien, R.M. 2007. A caution regarding rules of thumb for variance inflation factors. Qual. Quant. **41**: 673–690. doi:10.1007/s11135-006-9018-6.

Odeh, I.O.A., Todd, A.J., and Triantafilis, J. 2003. Spatial prediction of soil particle-size fractions as compositional data. Soil Sci., **168**. doi:10.1097/01.ss.0000080335.10341.23

Ontario Geological Survey 2010. Surficial geology of Southern Ontario, miscellaneous release - Data-128-REV.

Ontario Geological Survey 2011. 1:250,000 scale bedrock geology of Ontario. MISCELLANEOUS RELEASE - DATA 126 - Revision 1.

Ontario Ministry of Agriculture, Food and Rural Affairs. 2019. Ontario Soil Survey Complex. Ontario Ministry of Agriculture, Food and Rural Affairs, Guelph, Ontario.

Pahlavan-Rad, M.R., and Akbarimoghaddam, A. 2018. Spatial variability of soil texture fractions and pH in a flood plain (case study from eastern Iran). Catena, **160**: 275–281. doi:10.1016/j.catena.2017.10.002.

Poggio, L., and Gimona, A. 2017. 3D mapping of soil texture in Scotland. Digit. Soil Mapp. Globe **9**: 5–16. doi:10.1016/j.geodrs.2016.11.003.

Poggio, L., de Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., and Rossiter, D. 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. Soil. **7**: 217–240. doi:10.5194/soil-7-217-202 1.

Pourghasemi, H.R., Yousefi, S., Kornejady, A., and Cerdà, A. 2017. Performance assessment of individual and ensemble data-mining techniques for gully erosion modeling. Sci. Total Environ. **609**: 764–775. doi:10.1016/j.scitotenv.2017.07.198. PMID: 28763673.

R Core Team 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Richer-de-Forges, A.C., Arrouays, D., Chen, S., Román Dobarco, M., Libohova, Z., Roudier, P., et al. 2022. Hand-feel soil texture and particle-size distribution in central France. Relationships and implications. Catena, **213**: 106155. doi:10.1016/j.catena.2022.106155.

Rossiter, D.G., Zeng, R., and Zhang, G.-L. 2017. Accounting for taxonomic distance in accuracy assessment of soil class predictions. Geoderma, **292**: 118–127. doi:10.1016/j.geoderma.2017.01.012.

Saurette, D.D. 2021. onsoilsurvey: Making PDSM in Ontario Better.

Schut, L.W., and Wilson, E.A. 1987. The soils of the regional municipality of Ottawa-Carleton (No. 58). Ontario Institute of Pedology. Research Branch, Agriculture and Agri-Food Canada. Ontario Ministry of Agriculture and Food. Department of Land Resource Science, University of Guelph., Guelph, Ontario.

SCWG, (Soil Classification Working Group) 1998. The Canadian System of Soil Classification, 3rd ed. NRC Research Press, Ottawa, Ontario.

Sheldrick, B.H., and Wang, C. 1993. Particle size distribution, *In*: Soil sampling and methods of analysis. Lewis Publishers, Canadian Society of Soil Science, pp. 499–507.

Soil Science Division Staff 2017. Soil survey manual, USDA Handbook 18. Government Printing Office, Washington, D.C.

Thompson, J.A., Roecker, S., Grunwald, S., and Owens, P.R. 2012. Chapter 21 - Digital Soil mapping: interactions with and applications for hydropedology, *In*: Hydropedology. *Edited by* H. Lin. Academic Press, Boston, pp. 665–709. doi:10.1016/B978-0-12-386941-8.00021-6.

Wadoux, A.M.J.-C., Minasny, B., and McBratney, A.B. 2020. Machine learning for digital soil mapping: applications, challenges and suggested solutions. Earth-Sci. Rev. **210**: 103359. doi:10.1016/j.earscirev.2020.103359.

Wang, Z., Shi, W., Zhou, W., Li, X., and Yue, T. 2020. Comparison of additive and isometric log-ratio transformations combined with machine learning and regression kriging models for mapping soil particle size fractions. Geoderma, **365**: 114214. doi:10.1016/j.geoderma.2020.114214.

Wu, Q. 2019. whitebox: "WhiteboxTools" R Frontend.

Zhang, M., Shi, W., and Xu, Z. 2020. Systematic comparison of five machine-learning models in classification and interpolation of soil particle size fractions using different transformed data. Hydrol. Earth Syst. Sci. **24**: 2505–2526. doi:10.5194/hess-24-2505-2020.