Amalgamation and harmonization of soil survey reports into a multi-purpose database

Authors: Furze, Shane, and Arp, Paul

Source: Canadian Journal of Soil Science, 101(2) : 222-247

Published By: Canadian Science Publishing

URL: https://doi.org/10.1139/cjss-2020-0078

NRC Research Press

ARTICLE

Amalgamation and harmonization of soil survey reports into a multi-purpose database

Shane Furze and Paul Arp

Abstract: There is a growing demand for standardized, easily accessible, and detailed information pertaining to soil and its variability across the landscape. Typically, this information is only available for selected areas in the form of local or regional soil surveys reports which are difficult, and costly, to develop. Additionally, soil surveying protocols have changed with time, resulting in inconsistencies between surveys conducted over different periods. This article describes systematic procedures applied to generate an aspatial, terminologically, and unit-consistent, database for forest soils from county-based soil survey reports for the province of New Brunswick, Canada. The procedures involved (*i*) amalgamating data from individual soil surveys following a hierarchical framework, (*ii*) summarizing and grouping soil information by soil associations, (*iii*) assigning correct soil associates to each association, with each soil associate distinguished by drainage classification, (*iv*) assigning pedologically correct horizon sequences, as identified in the original soil surveys, to each soil associate, (*v*) assigning horizon descriptors and measured soil properties to each horizon, as outlined by the Canadian System of Soil Classification, and (*vi*) harmonizing units of measurement for individual soil properties. Identification and summarization of all soil associations (and corresponding soil associates) was completed with reference to the principal soil-forming factors, namely soil parent material, topographic surface expressions, soil drainage, and dominant vegetation type(s). This procedure, utilizing 17 soil surveys, resulted in an amalgamated database containing 106 soil associations, 243 soil associates, and 522 soil horizon sequences summarizing the variability of forest soil conditions across New Brunswick.

Key words: soil surveys, amalgamation, harmonization, soil database.

Résumé : De plus en plus, on demande des informations normalisées, faciles d'accès et suffisamment détaillées, sur le sol et la manière dont il varie avec le relief. Habituellement, les informations de ce genre ne sont disponibles que pour certains endroits et prennent la forme de levés pédologiques régionaux difficiles à élaborer, donc coûteux. Par ailleurs, les protocoles employés pour effectuer les levés ont évolué avec le temps, si bien qu'on note des incohérences dans les rapports produits à diverses époques. L'article que voici décrit la procédure méthodique utilisée pour créer une base de données non asservie à l'espace, cohérente sur les plans de la terminologie et de la métrologie, pour les sols forestiers du Nouveau-Brunswick (Canada). La base de données s'appuie sur les levés pédologiques des comtés de la province. La procédure était la suivante : (*i*) amalgamer les données des différents levés selon un cadre hiérarchique; (*ii*) synthétiser et assembler les données par groupes de sols; (*iii*) attribuer les sols d'association adéquats à chaque groupe, chaque sol d'association se distinguant par sa classe de drainage; (*iv*) attribuer à chaque sol d'association la séquence d'horizons qui correspond à la pédologie du levé original; (*v*) attribuer des descripteurs et les valeurs des propriétés du sol à chaque horizon, conformément au Système canadien de classification des sols; (*vi*) harmoniser les unités de mesure de chaque propriété. Tous les groupes pédologiques (et les sols d'association correspondants) ont été identifiés et résumés d'après les principaux facteurs constitutifs, c'est-à-dire les matériaux originels, le modelé topographique, le drainage et la végétation dominante. Cette procédure, appliquée à 17 levés, a débouché sur une base de données générale comprenant 106 groupes pédologiques, 243 sols d'association et 522 séquences d'horizons. Elle illustre la variabilité des sols forestiers et de leur condition au Nouveau-Brunswick. [Traduit par la Rédaction]

Mots-clés : levés pédologiques, amalgamation, harmonisation, base de données sur les sols.

Received 10 June 2020. Accepted 28 September 2020.

S. Furze and P. Arp. Faculty of Forestry and Environmental Management, University of New Brunswick, Fredericton, NB E3B 5A3, Canada.

Corresponding author: Shane Furze (email: furze.shane@gmail.com).

Can. J. Soil Sci. **101**: 222–247 (2021) dx.doi.org/10.1139/cjss-2020-0078

Introduction

Soils are an integral component of the natural environment, consisting of complex interactions between organic and inorganic constituents as affected by soil-forming factors, namely geology, climate, topography, organisms, and time (Jenny 1941; Birkeland 1999; Adhikari et al. 2012). As such, soils vary vertically with increasing depth (soil profile) and laterally by spatial location. These changes are generally summarized in the form of soil survey reports which have long been utilized as the dominant format for summarizing and mapping soils in Canada, with the first soil survey conducted over a century ago (McKeague and Stobbe 1978). For the Province of New Brunswick, Canada, these surveys have been conducted on a county-by-county basis over the past six decades, with the most recent survey conducted in 2000 (Michalica et al. 2000; Anderson and Smith 2011).

Soil surveys completed within NB summarize soils via landform- and lithology-defined soil associations, and further divide these into drainage-explicit soil associates. Each soil associate is typically assigned one or more soil profiles (horizon sequence from forest floor to parent material) summarizing the depth, chemical, and physical properties of each horizon (McKeague and Stobbe 1978; Anderson and Smith 2011). This information is typically outlined via three sections, with the third section often represented as an appendix at the end of the survey [e.g., description of Poitras soil associate retrieved from Langmaid et al. (1980)]:

Section 1: It comprises an overview of each soil associate within the surveyed area, including information pertaining to parent soil association and soil-forming factors, namely landform, lithology, vegetation, drainage, and topographic surface expression.

Section 2: It provides profile description for sampled soil associates including field-based, horizon-specific identification and measurements including depth, texture, coarse fragment (CF) content, structure, root presence, mottling (if applicable), and pH (Fig. 1).

Section 3: It outlines the laboratory-measured physical and chemical properties, by horizon, for each soil associate [e.g., horizon description with depth, % carbon, % sand, silt and clay, bulk density (D_b), field capacity (FC), permanent wilting point (PWP)] (Fig. 2).

Soils vary with changing soil-forming factors which vary continuously over landscapes, whereas individual soil surveys are limited geographically, often by administrative boundaries (a provincial county) that are not related to soil-forming factors. Although individual soil surveys are useful as stand-alone documents, they describe the range of soil properties within a limited geographic area that may not adequately capture the full range of variability of soil properties within a soil type. In addition, soils surveys have been collected over decades, a time frame that spans significant changes in analytical procedures, protocols, and methods (Subcommittee on Methods of Analysis of Canada Soil Survey Committee 1978; Guertin et al. 1984; Carter and Gregorich 2006). Therefore, it is important to combine soil surveys into harmonized databases which provide the full variability of soil properties found across environmental gradients while ensuring consistent methods and units of measurement for individual soil properties.

Compiling available soil survey reports for NB revealed numerous inconsistencies in terms of (*i*) naming of the same soil associations and subsequent associates, (*ii*) labelling horizon descriptors, and (*iii*) methods, and units, of measurement for analyzed soil properties. This is, in part, due to changing soil classification and mapping protocols over the past six decades, with the first published soil survey for NB conducted in 1940 (Stobbe 1940).

The objective of this article was to introduce and describe a framework applied to develop a seamless and terminologically consistent aspatial forest soils database by amalgamating and harmonizing existing soil survey reports for NB as a case study. As such, this study aims to outline the step-by-step process in which the database was created as a framework for application in other geographic locations, whether at a regional, provincial, national, or international scale. This objective was accomplished by

1. compiling existing soil survey information into a single initial database,
2. unifying the classifications and descriptions assigned to all surveyed soil associations, soil associates, and soil-forming factors (including drainage regime and soil classification),
3. standardizing the soil associate names within each soil association,
4. ensuring consistent soil horizon classifications, and both methods and units of measurement, for the physical and chemical properties of each horizon, notably horizon depth, soil texture, soil organic matter (SOM) content, CF content, D_b, and soil moisture retention at both FC and PWP.

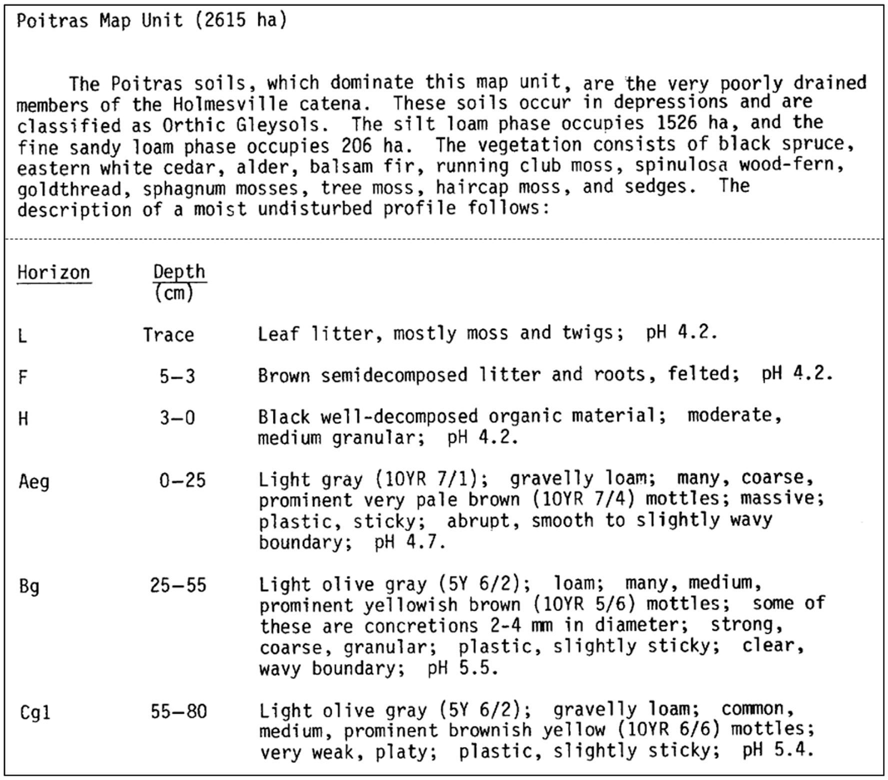
Materials and Methods

Survey report amalgamation

The amalgamation of soil survey reports and harmonization of soil attributes was guided by

1. the standardized soil surveying terminology of the Mapping System Working Group (1981) and the Expert Committee on Soil Survey (1982),
2. sampling and analytical techniques as described by Sub Comiittee on Methods of Analysis (1978) and Guertin et al. (1984),

Fig. 1. Example of information obtained from soil surveys and utilized in developing the database, including general information (Section 1) (top paragraph) and field-based measurements for each horizon (Section 2) (bottom descriptions) separated by dotted line. The example provided represents the Poitras soil associate retrieved from Langmaid et al. (1980).

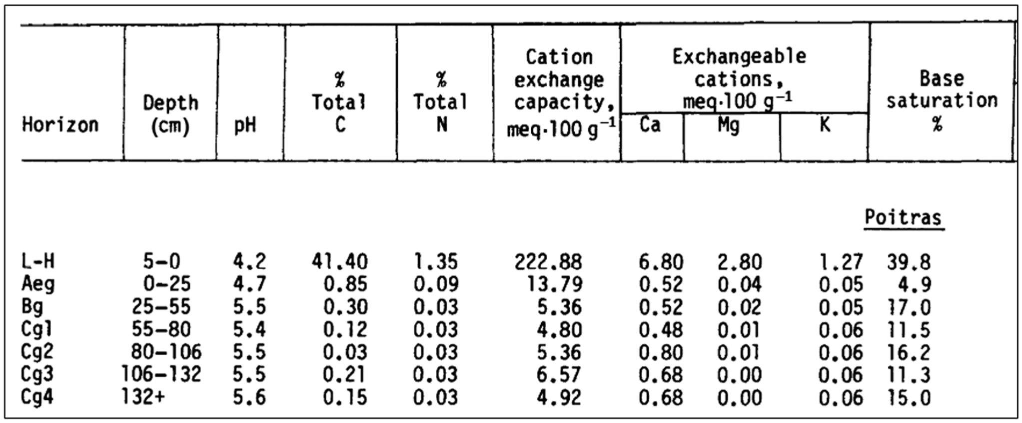


Poitras Map Unit (2615 ha)

The Poitras soils, which dominate this map unit, are the very poorly drained members of the Holmesville catena. These soils occur in depressions and are classified as Orthic Gleysols. The silt loam phase occupies 1526 ha, and the fine sandy loam phase occupies 206 ha. The vegetation consists of black spruce, eastern white cedar, alder, balsam fir, running club moss, spinulosa wood-fern, goldthread, sphagnum mosses, tree moss, haircap moss, and sedges. The description of a moist undisturbed profile follows:

Horizon	Depth (cm)	
L	Trace	Leaf litter, mostly moss and twigs; pH 4.2.
F	5–3	Brown semidecomposed litter and roots, felted; pH 4.2.
H	3–0	Black well-decomposed organic material; moderate, medium granular; pH 4.2.
Aeg	0–25	Light gray (10YR 7/1); gravelly loam; many, coarse, prominent very pale brown (10YR 7/4) mottles; massive; plastic, sticky; abrupt, smooth to slightly wavy boundary; pH 4.7.
Bg	25–55	Light olive gray (5Y 6/2); loam; many, medium, prominent yellowish brown (10YR 5/6) mottles; some of these are concretions 2-4 mm in diameter; strong, coarse, granular; plastic, slightly sticky; clear, wavy boundary; pH 5.5.
Cg1	55–80	Light olive gray (5Y 6/2); gravelly loam; common, medium, prominent brownish yellow (10YR 6/6) mottles; very weak, platy; plastic, slightly sticky; pH 5.4.

Fig. 2. Example of information obtained from soil surveys and utilized in developing the database, including laboratory-measured, horizon-specific soil physical and chemical properties. The example provided represents the Poitras soil associate retrieved from Langmaid et al. (1980).



Horizon	Depth (cm)	pH	% Total C	% Total N	Cation exchange capacity, meq·100 g^{-1}	Exchangeable cations, meq·100 g^{-1}			Base saturation %
						Ca	Mg	K	
									Poitras
L-H	5–0	4.2	41.40	1.35	222.88	6.80	2.80	1.27	39.8
Aeg	0–25	4.7	0.85	0.09	13.79	0.52	0.04	0.05	4.9
Bg	25–55	5.5	0.30	0.03	5.36	0.52	0.02	0.05	17.0
Cg1	55–80	5.4	0.12	0.03	4.80	0.48	0.01	0.06	11.5
Cg2	80–106	5.5	0.03	0.03	5.36	0.80	0.01	0.06	16.2
Cg3	106–132	5.5	0.21	0.03	6.57	0.68	0.00	0.06	11.3
Cg4	132+	5.6	0.15	0.03	4.92	0.68	0.00	0.06	15.0

3. The National Soils Database ["NSDB", spatial coverage with summary documentation, Canadian Soil Information Service (2000)],
4. the two province-wide soil summary documents and associated distribution maps:
   a. "Soils of New Brunswick: the Second Approximation" ["SNB", Fahmy et al. (2010)],
   b. "Forest Soils of New Brunswick" ["FSNB", Colpitts et al. (1995)].

Soil survey reports for NB, in addition to the NSDB, SNB, and FSNB documentation, were retrieved from the publications section of the Canadian Soil Information System (Canadian Soil Information Service 2012). Each survey was assessed to determine the extent of data availability and spatial coverage (Table 1; Fig. 3). Soil survey reports excluded from the database were excluded either due to the omission of horizon-specific data, or insufficient information pertaining to horizon-

Table 1. Overview of available soil surveys for New Brunswick, Canada, with relevant information regarding utility for inclusion into the database.

Soil Survey	Year	Scale	Horizon-specific (Y/N)	Utilized	Digitized (Y/N)	Spatial coverage (% of NB)	Source
Fredericton – Gagetown	1940	95 040	N	—	Y	4.3	Stobbe 1940
Woodstock Area	1944	63 360	Y	X	N	—	Stobbe and Aalund 1944
Southeastern New Brunswick	1950	126 720	N	—	N	—	Aalund and Wicklund 1950
Southwestern New Brunswick	1953	156 720	N	—	Y	12.9	Wicklund and Langmaid 1949
Andover – Plaster Rock	1963	63 360	Y	X	Y	4.6	Millette and Langmaid 1964
Southern Northumberland County	1964	31 680	N	—	N	—	Langmaid et al. 1964
Mount Carleton Provincial Park	1972	—	N	—	N	—	High Country Research and Development 1972
Lepreau Provincial Park	1973	—	Y	—	N	—	Langmaid and Losier 1973
Northern Victoria County	1976	63 360	Y	X	Y	4.6	Langmaid et al. 1980
Havelock Parish	1980	10 000	Y	X	Y	0.4	MacMillan and Chalifour 1980
Madawaska County	1980	50 000	Y	X	Y	4.8	Langmaid et al. 1980
St. Quentin – Kedgwick	1982	50 000	Y	X	Y	0.6	Dube 1982
Rogersville – Richibucto Region	1983	50 000	Y	X	Y	5.8	Wang and Rees 1983
Senator Herve J. Michaud Experimental Farm Agriculture Canada, Buctouche	1983	3 000	N	—	N	—	Rees and Fahmy 1983
Agriculture Canada Research Station, Fredericton	1984	4 800	N	—	N	—	Rees and Fahmy 1984
Sussex Area	1986	20 000	Y	X	N (GR)	0.3	Holmstrom 1986
Woodstock – Florenceville (Vol. 1)	1989	20 000	Y	X	Y	0.1	Fahmy and Rees 1989
Woodstock – Florenceville (Vol. 2)	1992	20 000	Y	X	Y	0.1	Fahmy and Rees 1992
Chipman – Minto – Harcourt Region	1992	—	Y	X	Y	10.0	Rees et al. 1992
Agriculture Canada Benton Ridge Potato Breeding Substation	1992	5 000	N	—	N	—	Rees and Fahmy 1992
Black Brook Watershed	1993	10 000	Y	X	N	—	Mellerowicz et al. 1993
Moncton Parish	1993	20 000	Y	X	N (GR)	1.1	Rees et al. 1995
Woodstock – Florenceville (Vol. 3)	1996	20 000	Y	X	Y	0.4	Fahmy and Rees 1996
Shediac and Botsford Parishes	1996	20 000	Y	X	Y	1.0	Rees et al. 1996
Dorchester Parish	1998	20 000	Y	X	Y	0.1	Rees et al. 1998
Acadian Peninsula	2000	20 000	Y	X	Y	1.5	Michalica et al. 2000
Woodstock – Florenceville (Vol. 4)	2001	20 000	Y	X	Y	1.0	Fahmy et al. 2001
Central and Northern New Brunswick	2005	250 000	N	—	N	—	Rees et al. 2005

Note: Some reports do not provide a spatial data layer but do provide images. These have been georeferenced (GR) but not digitized.

specific soil-forming processes, e.g., "A_1" instead of "Ae" or "Ah". Also excluded were survey reports specifically dealing with small sections of agricultural lands due to the influence of ploughing and tillage on soil properties.

The compilation of soil survey data within the database was conducted manually on a row-by-row basis following a hierarchical framework (Table 2) which first separates the database by the original soil survey from which the data was retrieved. Following survey source, the database was separated by soil association then soil-forming factors for each association, including dominant vegetation type, topography (surface expression, slope position, slope steepness, and aspect), and soil parent materials (lithology and mode of deposition).

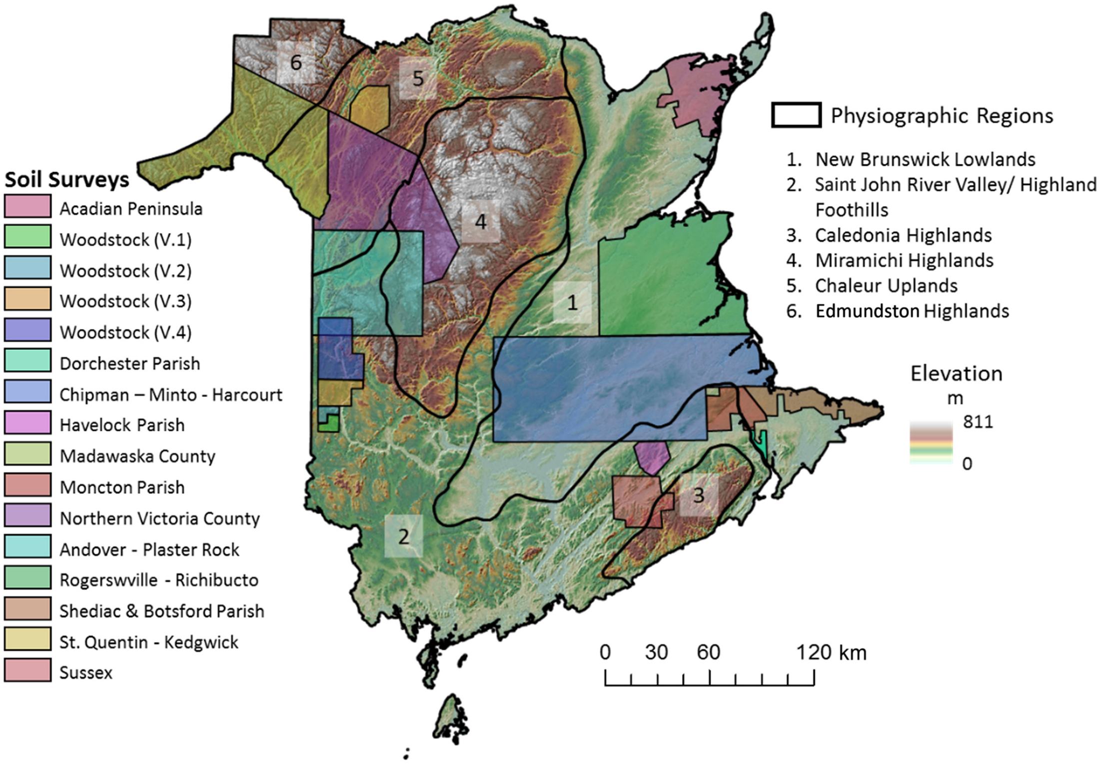


Fig. 3. Spatial coverage of soil surveys utilized for developing the database. Also included is elevation and physiographic regions to represent how coverage varies in each region. Physiographic regions retrieved from Colpitts et al. (1995), elevation with shaded relief retrieved from Furze et al. (2017), and survey outlines retrieved from Canadian Soil Information Service (2012). Figure created in ArcGIS (Environmental Systems Research Institute 2018).

Next, the database was separated by drainage, stoniness, and rockiness, with individual soil associates assigned to each drainage class. Each soil associate was provided with either single or multiple horizon sequences depending on how often that associate was sampled within each survey. Each soil horizon was also assigned depths for which the horizon begins and ends in addition to measured soil properties for each horizon.

This compilation resulted in an amalgamated database consisting of 522 soil profiles with 2490 rows of horizon-specific data. Inspecting the amalgamated database, however, revealed inconsistencies between surveys, including

1. naming of soil associations and soil associates,
2. variability and incompleteness in soil drainage and soil order classifications,
3. labelling, and incompleteness, of soil-forming factor entries, mainly parent material, topographic surface expression, vegetative cover, and
4. changes in soil survey methods such as sampling strategies, laboratory analyses, and quantitative units for reporting results (Table 3) (Subcommittee on Methods of Analysis of Canada Soil Survey Committee 1978; Mapping System Working Group 1981; Expert Committee on Soil Survey 1982; Guertin et al. 1984).

Correcting soil naming, drainage, and classification inconsistencies

Naming

Naming inconsistencies of both soil associations and associates were resolved utilizing the FSNB and SNB reports as guiding authority as follows:

1. Some of the soil associations were referred to as complexes between two associations, e.g., "Baie du Vin – Galloway", "Barrieau – Buctouche", and "Parleeville – Tobique" due to similar soil-forming factors and soil properties. For these instances, only one name was retained based on descriptions of parent material lithology and mode of deposition. For example, the Baie du Vin – Galloway complex was assigned a sandstone lithology and a glaciomarine mode of deposition. From this, Galloway was assigned as the soil association because it occurs on glaciomarine deposits, whereas Baie du Vin occurs on glaciofluvial and marine deposits.
2. In some situations, only the names of the soil associations, instead of individual soil associates, were provided for each profile although explicit drainage classes were assigned to each. For these, new drainage-related soil associate names were assigned based on table 6 ("Correlation of New Brunswick Soil Series/Associations with Forest Soil

Table 2. Visual example of database layout for each association within the database utilizing one of the Holmesville association entries.

						Depth (cm)		Physical properties				Chemical properties		
Association	Lithology	Landform	Drainage class	Associate	Horizon	From	To	Sand (%)	Silt (%)	Clay (%)	D_b ($g{\cdot}cm^{-3}$)	SOM (%)	pH	CEC
Holmesville	Sandstone and quartzite with some argillite and slate and shale and schist	Basal Till	Well	Holmesville	LFH	−5	0	—	—	—	—	—	—	—
					Ae	0	5	41.9	42.8	15.3	0.76	4.18	3.50	9.77
					Bhf_1	5	6	31.8	44.9	23.3	0.66	15.0	3.90	61.34
					Bhf_2	6	9	26.2	50.1	23.7	0.67	10.97	4.20	48.17
					Bf_1	9	15	42.4	42.6	15	1.04	5.57	4.60	23.84
					Bf_2	15	30	48.6	37.7	13.7	1.33	2.39	4.70	10.08
					BC	30	46	45.9	38.3	15.8	1.63	0.88	4.80	6.93
					C	46+	—	54.2	32.3	13.5	1.98	0.36	5.20	3.48
			Imperfect	Johnville	Aeg	0	5	22.5	53.6	23.9	1.1	3.69	3.80	13.05
					Bfgj	5	9	19.7	53.4	26.9	1.38	7.57	4.30	31.72
					BCg_1	9	22	40.3	45.5	14.2	1.4	2.08	4.80	12.22
					BCg_2	22	56	34.2	49.5	16.3	1.83	2.65	4.80	14.02
					Cg_1	56	80	34.4	42.9	22.7	1.98	0.41	4.90	5.2
					Cg_2	80	121	32.3	42	25.7	1.83	0.21	5.00	5.75
					C	121+	—	31.1	41.1	27.8	1.93	0.26	5.00	6.20
			Poor	Poitras	Aeg	0	25	36.9	44.3	18.8	1.81	1.46	4.70	13.79
					Bg	25	55	55	33.7	11.3	1.82	0.52	5.50	5.36
					Cg_1	55	80	41.4	37.6	21.0	1.9	0.21	5.40	4.80
					Cg_2	80	106	38.8	43.3	17.9	1.92	0.05	5.50	5.36
					Cg_3	106	132	36.8	46	17.2	1.95	0.36	5.50	6.57
					Cg_4	132+	—	55.8	30.1	14.1	2.01	0.26	5.60	4.92

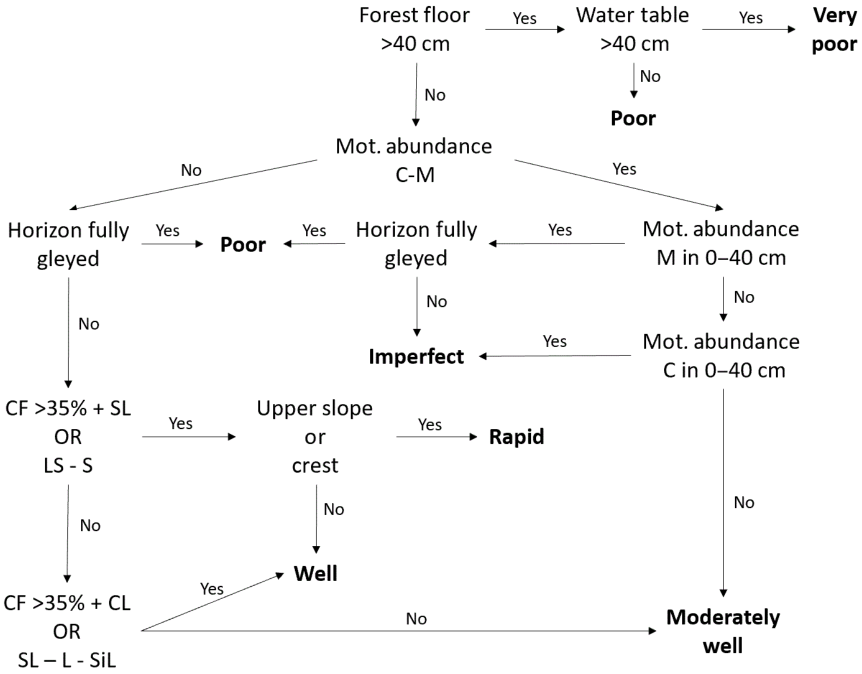
Note: Not all site- and horizon-specific properties are provided in this example. D_b, bulk density; SOM, soil organic matter; CEC, cation exchange capacity.

Table 3. Overview of variability in units of measurements for select soil properties from soil surveys utilized in developing the database.

													Water retention					
	CF			Texture		pH			SOM				FC			PWP		
Survey	%	Range	Desc.	%	Class	H_2O	$CaCl_2$	Unknown	C %	OM %	LOI 450°	D_b ($g{\cdot}cm^{-3}$)	33 kPa	1/3 atm	1/3 bar	1500 kPa	1/3 atm	1/3 bar
Acadian Peninsula	—	X	X	X	X	X	—	—	—	X	—	X	X	—	—	X	—	—
Black Brook Watershed	X	—	—	X	X	—	X	—	X	—	—	X	X	—	—	X	—	—
Woodstock – Florenceville (Vol. 1)	X	—	—	X	X	X	X	—	X	—	—	X	—	—	X	—	—	X
Woodstock – Florenceville (Vol. 2)	X	—	—	X	X	—	X	—	X	—	—	X	—	—	—	—	—	—
Woodstock – Florenceville (Vol. 3)	X	—	—	X	X	—	X	—	X	—	—	X	—	—	—	—	—	—
Woodstock – Florenceville (Vol. 4)	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Chipman – Minto – Harcourt Region	X	—	X	X	X	X	X	—	X	—	—	X	—	X	—	—	X	—
Dorchester Parish	X	—	—	X	X	—	X	—	X	—	—	X	X	—	—	X	—	—
Havelock Parish	—	—	X	X	X	—	—	X	X	—	—	X	—	X	—	—	—	—
Madawaska County	—	—	X	X	X	—	—	X	X	—	—	X	X	—	—	X	—	—
Moncton Parish	—	X	—	X	X	X	X	—	X	—	—	X	X	—	—	X	—	—
Northern Victoria County	—	—	X	X	X	X	—	—	X	—	X	X	X	—	—	X	—	—
Andover – Plaster Rock	—	—	X	X	X	X	—	—	X	—	X	X	X	—	—	—	—	—
Rogersville – Richibucto Region	X	—	—	X	X	X	X	—	X	—	—	X	—	—	—	—	—	—
Shediac and Botsford Parishes	X	—	—	X	X	X	X	—	X	—	—	X	X	—	—	X	—	—
St. Quentin – Kedgwick	—	—	X	X	X	—	—	X	X	—	—	X	X	—	—	X	—	—
Sussex Area	—	—	X	X	—	X	X	—	—	X	—	X	—	—	—	—	—	—

Note: "X" denotes the units of measurement provided. CF, coarse fragment; SOM, soil organic matter; LOI, loss on ignition; D_b, bulk density; FC, field capacity, PWP, permanent wilting point.

Fig. 4. Visual representation of model developed to assign drainage regime to aspatial database. Resulting drainage classes are bolded. Note that “Mot.” Is abbreviation for mottle, “CF” for coarse fragments, “C” for coarse, “M” for moderate, “LS” for loamy sand, “S” for sand, “SL” for sandy loam, “L” for loam, “SiL” for silty loam.



Units”) of the SNB report, resulting in 97 soil profiles with updated soil associate names.

3. Within some reports, horizon and depth specifications by soil associate (Section 2 of Fig. 1) were inconsistent with their listing at the end of the reports (Section 3, Fig. 2). This was most prevalent in the Northern Victoria and St. Quentin soil surveys. In these situations, the depths provided with the property measurements (Section 3) were retained.

Also, inconsistent was the quantity of data provided for each soil associate. For example, only general information was provided for some soil associates (Sections 1 and 2), whereas measured soil properties (Section 3) were omitted. This resulted in some soil associates lacking horizon-specific property measurements altogether. This was the case for 150 soil profiles within the database.

Drainage

Soil drainage classifications ranged from very poor (wetlands and organic soils) to rapidly and excessively drained (coarse-textured, upper slope positions), as outlined by the Expert Committee on Soil Survey (1982). The procedure in Fig. 4 was used to determine if soil drainage was correctly classified for each soil associate in terms of soil horizon sequence, as well as properly assigning drainage classifications to horizon sequences lacking this information.

This procedure ensured that (*i*) the soil associate names within each association were consistent with the drainage expectations based on table 4 of the SNB report, titled “New Brunswick mineral soil catenas”, and (*ii*) drainage and slope positions were congruent such that well- to rapidly drained members occur on upper slope and hill-crest positions, imperfect- to moderately well-drained members occur on the lower slopes, and very poor to poorly drained members along toe slopes and in depressions. The derived drainage classifications were generally consistent with the original drainage assignments. In cases where the original drainage classifications provided broad ranges, the middle drainage class was assigned as a median, e.g., “very poor – imperfect” reassigned to “poor”.

Soil classification

Each soil profile was classified by way of the hierarchical Canadian Soil Classification System context by specifying its belonging to a soil order, great group, and subgroup (Soil Classification Working Group 1998). Classifications for each soil profile were assessed by comparing the provided horizon sequences to those found in the Canadian Soil Classification System. The database includes the Podzolic, Brunisolic, Regosolic, Luvisolic, Gleysolic, and Organic orders.

Correcting soil-forming factors

Parent material

Soil parent materials are classified by mode of deposition and primary lithology. Differences in these descriptions were observed within the same soil association names when cross-referencing the database entries to table 6 of the SNB, tables 2 and 5 of the FSNB, and the NSDB reports. The discrepancies were corrected by comparing the modes of deposition outlined in the SNB, FSNB, and NSDB documents. If three or more sources (including soil surveys, SNB, FSNB, and NSDB) provided the same mode of deposition for an individual soil association, then that mode of deposition was assigned to that association. Any remaining inconsistencies were addressed by determining the likely mode of deposition by surface expressions (topography), CF content, and horizon sequences. Together, this cross-referencing resulted in 21 unique modes of deposition with some associations having two distinct modes of deposition overlaying each other (i.e., glaciomarine/basal). In such cases, the top parent material will have the dominant influence on soil formation and development.

Inconsistencies also occurred with respect to primary lithology specifications. For example, the parent material of the Baie du Vin association was labelled as "*acidic GLFL or MA sand, petrologically similar to underlying sandstone bedrock, and rich in biotite*". The Galloway soil associations, stated to have the same lithology, were labelled as "*acidic, petrologically similar to the underlying sandstone bedrock and rich in biotite*". These inconsistencies were addressed through re-labelling and by updating lithology via (*i*) dominant rock types, (*ii*) dominant grain sizes, and (*iii*) mineral hardness (based on Mohs hardness scale). This was followed by (*i*) providing binary descriptors for sedimentary, igneous, and metamorphic parent materials per soil association, and (*ii*) the ranking of rock type weatherability and inherent fertility. Weatherability and fertility specifications were based on table 4 of the FSNB report. Some soil associations (e.g., Bellefleur, Bottomland, Bransfield, Chockpish, Gulquac, Lower Ridge, St. Charles, and Wakefield) could not be identified as forest soil associations in either FSNB or SNB reports, resulting in absent lithological classifications for these soils. Table A1 summarizes the relationships between soil associates, associations, and parent material mode of deposition and lithology.

Topography

Topographic surface expression descriptions for each soil association, when provided, also varied by survey report, ranging from flat (or domed) for organic soils to strongly rolling and hilly on dense igneous parent materials in the New Brunswick Highlands. Although included in the database, little emphasis was placed on topographic expressions because the expressions vary by resolution, intensity and frequency of changing topographic positions, slopes, and geographic regions, resulting in inconsistent descriptions between reports. A consistent measure for each association referred to average slope position and slope percent, but this information was only provided for 40% of the database.

Vegetation

Some surveys listed the presence of dominant overstory species, generally within the vicinity of the soil sampling points. These specifications were entered into the database in the form of binary fields referring to dominance of shade tolerant hardwoods, softwoods, and mixed woods within the overstory canopy. This was available for 46% of the database. Also, where forest floor data were provided, forest floor thickness was assigned to each horizon sequence (available for 59% of the database).

Correcting horizon-specific properties

Horizon descriptions and depths

Considerable effort was placed on ensuring that the soil profile and individual horizon classifications within the surveys were consistent with those outlined in Soil Classification Working Group (1998). All horizons were generalized by master horizon (forest floor, A, B, and C) and by the first subscript for each master horizon, e.g., Ae, Ah, Bf (represents the dominant process influencing the soil). When generalizing by the initial subscript, dominant suffixes followed by the "j" subscript were replaced with the "m" subscript. For example, a "Bf_j" was replaced with "Bm". Additionally, dominant horizon specifications such as g, c, x, j, and t, were entered into the database in their own individual columns as values ranging from 0 to 1 depending on prominence (0 = absent, 0.5 = partial ("j" suffix), and 1 = present). For example, the presence of gleying in a Bg horizon would receive a value of 1, Bgj a value of 0.5 and B a value of 0 for the "g" column.

Profiles where soil horizon descriptors were marked by "?" or "or" were re-labelled through cross-referencing with other similar soil profiles. It was also ensured that horizon depths (many originally measured in inches) were entered into the database in metric format. Additionally, some C horizons were provided depths with a "+" sign because the bottom of the horizon was not reached (Table 2 — Holmesville soil associate) and, for such cases, these were retained. A new depth column was added to the database, and the depth to the center of each horizon was provided. When the depth to the bottom of the horizon was unknown, the center depth was left blank.

Soil physical properties

For each mineral horizon, measured values for soil physical properties, namely soil texture, CF content, structure, SOM content, D_b, and water retention (at both FC and PWP) were entered into the database utilizing specific procedures for each property, as outlined below.

Soil texture information, was entered into the database in two forms:

1. texture classes assigned from the soil texture ternary diagram, as outlined in Soil Classification Working Group (1998) (Fig. 5), and

Table 4. Logical rule statements applied in ascending order to determine proper soil texture class based on texture ternary diagram as outlined in Soil Classification Working Group (1998).

Rule	Output class	Output abbreviation
Clay ≥ 60	Heavy Clay	HC
(60 – Sand) < Clay ≤ 40 and Sand ≤ 45	Silty Clay	SiC
Clay ≥ 40 and Sand < 45 and Clay ≥ (60 – Sand)	Clay	C
Clay ≥ 35 and Sand ≥ 45	Sandy Clay	SC
Clay ≥ 28 and Sand ≤ 20	Silty Clay Loam	SiCL
27.5 < Clay < 40 and 20 < Sand < 45	Clay Loam	CL
Clay ≤ 12 and Clay < (20 – Sand)	Silt	Si
Clay < (50 – Sand)	Silt Loam	SiL
Clay ≤ [2·(Sand – 70)]	Sand	S
Clay ≤ (Sand – 70)	Loamy Sand	LS
Clay ≤ 20 and Sand ≥ 53 OR Clay ≤ 7	Sandy Loam	SL
Clay ≥ (73 – Sand)	Sandy Clay Loam	SCL
Else	Loam	L

2. proportions of sand, silt, and clay within the fine-earth fraction, measured as percentages with the summation equaling 100% (although not always the case within the database).

Some texture descriptions provided broad ranges, e.g., “SiL-LS” (silty loam – loamy sand). For these cases, both texture class and percentage of sand, silt, and clay were re-assigned by choosing the midpoint within these classes on the texture ternary diagram. Additionally, some texture classifications did not fall within the realm of the ternary diagram, e.g., “G” (gravel) and “SG” (sandy gravel). This had occurred within 37 samples to which sand, silt, and clay contents remained absent. Although not in the texture triangle, these classifications were retained and placed in a separate column. Also, most texture classifications were assigned including a modifier, e.g., “vfSL” (very fine sandy loam). These modifiers were retained, but texture classes without modifiers were placed in a separate column.

For some horizons, only a texture class was provided, this was typically the case when specific horizon properties (Section 3) were absent from the survey. With these horizons, the percentage of sand, silt, and clay was left absent. For the cases where only the percentage of sand, silt, and clay were present (without assigned texture class), an automated model (Table 4) was derived to determine the texture class based on these percentages. This was used to fill in the voids where the texture classes were absent.

Coarse fragment content entries within the database varied in format, including ranges, qualitative descriptions, e.g., “few” or “some”, and specific measurements (e.g., 10%). Where ranges were provided, the mean values were assigned. Additionally, CF content was also included as part of the horizon texture description, e.g., “gravelly sandy loam”. For these cases, the suggestions of the Expert Committee on Soil Survey (1982) were adopted as follows: descriptions including the adjective “Non” were assigned <15% CF by volume, descriptions with no adjective were assigned 15%–35% CF content, descriptions including the adjective “very” were assigned 30%–60%, and descriptions including the adjective “extremely” were assigned CF contents >60%. Parent material modes of deposition were reviewed, via pivot tables, to determine their applicability in assigning CF contents, but some modes of deposition lacked CF measurements altogether (Table 5). As a result, some soil associations lack CF content measurements.

Soil structure was provided as a description with three components, shape, size, and distinctness, with all three components assigned to the database. An overview of soil structures can be found in Expert Committee on Soil Survey (1982). It was noticed that terminology for structureless soils were used interchangeably, namely “single grain”, “loose”, “amorphous”, and “massive”; therefore, these were grouped into two classes (“massive” for amorphous and massive descriptions and “single grain” for the remaining two). With this, binary columns were developed for each structure class and assigned 0 or 1 if absent or present, respectively. Samples assigned as transitions between two structure classes (e.g., subangular blocky – platy) were provided a value of 1 in the columns for each structure class.

Soil organic matter content was provided in four formats, % SOM, % carbon, and loss on ignition (LOI) at 450 °C and at 850 °C. Soil surveys for Plaster Rock and Northern Victoria Counties provided both % carbon and LOI at 450 °C (328 samples, 13% of database). Kent County was the only report to record LOI at 850 °C (11 samples, 0.4% of database) and did not record % carbon for comparison. Due to the lack of samples and omission of carbon values for comparison, readings for LOI at 850 °C were omitted. The % carbon readings were converted to % organic matter via eq. 1.

$$\%\text{SOM} = \%\text{C} \cdot 1.72 \quad (1)$$

Table 5. Overview of coarse fragment (CF) content for each parent material mode of deposition found within database.

	CF content				
Mode of deposition	Minimum	Maximum	Mean	SD	Sample size
Residual	10	40	24.2	12.8	6
Residual + colluvium	—	—	—	—	0
Ablation	5	80	24.7	16.6	76
Ablation/basal	0	30	11.9	7.5	94
Ablation/residual	10	70	47.1	16.0	24
Alluvium	0	0	0	0	14
Alluvium + glaciofluvial	—	—	—	—	0
Ancient alluvium	0	0	0	0	6
Basal	1	60	15.5	9.8	371
Basal/residual	—	—	—	—	0
Colluvium + water-reworked till	—	—	—	—	0
Glaciofluvial	1	80	35	22.9	45
Glaciofluvial + marine	0	35	4.6	7.4	106
Glaciolacustrine	—	—	—	—	0
Glaciomarine	0	12.5	3.6	3.4	55
Glaciomarine/basal	0	45	4.7	8.6	54
Glaciomarine/marine	0	0	0	0	8
Lacustrine	0	1	0.3	0.3	14
Marine	0	0	0	0	11

Note: Some modes of deposition lack CF content data, whereas others have large variations in values. SD, standard deviation.

where %SOM is % soil organic matter, %C is % carbon, and 1.72 is the conversion factor because SOM is composed of 58% carbon (Romano and Palladino 2002; Pollacco 2008; Chaudhari et al. 2013; Poggio et al. 2013). Standardizing these measurements into %SOM resulted in 1202 samples with measurements (48.3% of database).

Soil density measurements were provided for both particle density (D_p) and D_b. Box plots for D_b were used per horizon label to determine the extent to which outliers were skewing the dataset (Fig. 6). For each outlier, the original report was reviewed to determine if a data entry mistake had occurred. It was ensured that the ranges of the density values were generally consistent with soil texture, SOM, and soil depth expectations, with additional considerations to distinguish density in compacted versus non-compacted soils. Data entry with obvious data errors (soils with D_b greater than densities for silicate rocks) were deleted ($n = 4$).

Water retention measurements were provided in bars (bar), atmospheres (atm), and kilopascals (kPa) with values measured in both volumetric and gravimetric forms. Together, 10 reports provided water retention measurements in gravimetric form, whereas five reports provided measurements in volumetric form. Two reports did not specify whether measurements were gravimetric or volumetric. Gravimetric water retention values were recorded and converted to volumetric form via multiplication with D_b, where applicable. The different units of measurement were then amalgamated and adjusted to represent water retentions in kPa: −33 kPa for water retention at FC, and −1500 kPa for water retention at PWP. Additional moisture measurements included water % at 0 cm, water-holding capacity, maximum water-holding capacity, and water retention at saturation, 10 cm, 50 cm, 100 cm, −100 kPa, −400 kPa, hygroscopic moisture, available water, and moisture percentage. Emphasis was placed on moisture retention at FC and PWP due to the influence of these pressures on rooting. Once combined, water retention at FC had 836 samples measured (34.0% of database), whereas PWP had 743 samples measured (29.8%).

Results

Naming

The amalgamation and harmonization efforts have resulted in a database which highlights the variability and range in conditions found within different soil associations across NB. Correcting for naming inconsistencies resulted in summarized data for 106 soil associations and 243 drainage-explicit soil associates.

Drainage classifications

Incorrect drainage entries were addressed via the framework highlighted in Fig. 4 resulting in biased representations of different drainage classes throughout the database (Table 6). From this, poorly drained soils represent 16.63% of the database, whereas imperfectly drained soils represent 21.76%, moderately well-drained represent 16.47%, well-drained represent 34.06%, excessively drained represent 10.16%, and <1% with absent drainage classifications.

Fig. 5. Texture composition of soil samples for each soil survey utilized in developing the database overlain on texture ternary diagram. Figure was derived in R Studio (RStudio Team 2020) using packages “ggplot2”, “plyr”, and “ggtern” (Wickham 2011, 2016; Hamilton and Ferry 2018).

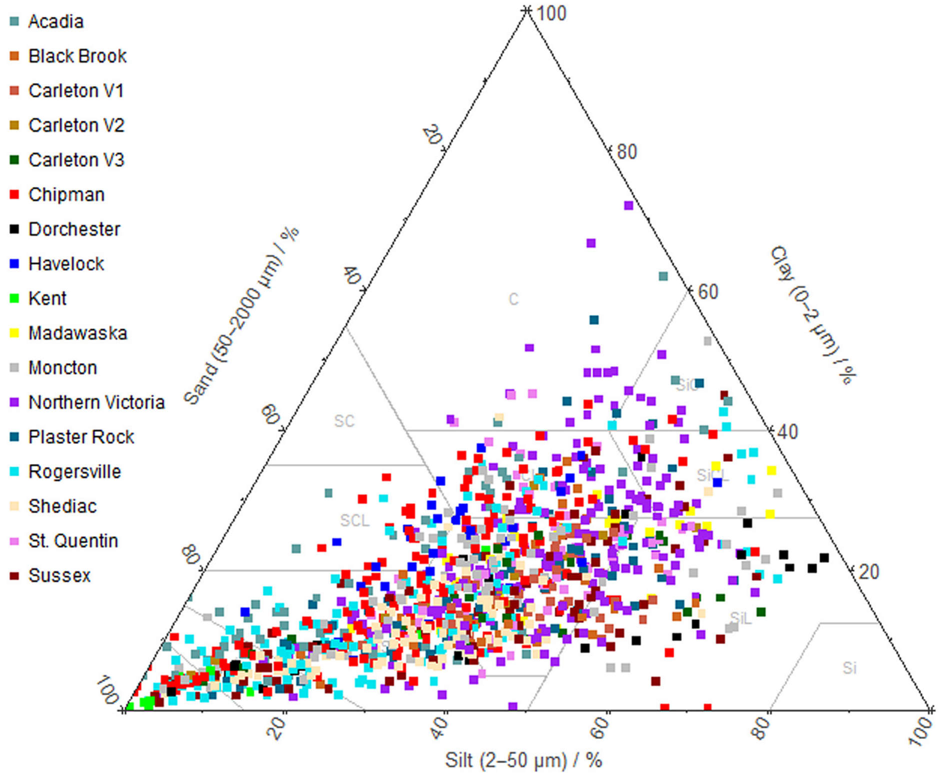


Table 6. Representation of drainage classes assigned to soil associates within the database with separations highlighting the dominant drainage classes.

Drainage class	Count	% Representation
Very poor (VP)	182	7.3
Very poor – poor (VP – P)	38	1.5
Poor (P)	194	7.8
Poor – imperfect (P – I)	301	12.1
Imperfect (I)	240	9.6
Imperfect – moderately well (I – MW)	8	0.3
Moderately well (MW)	341	13.7
Moderately well – well (MW – W)	61	2.5
Well (W)	839	33.7
Well – rapid (W – R)	9	0.4
Rapid (R)	89	3.6
Rapid – excessive (R – EX)	150	6.0
Excessive (EX)	14	0.6
—	24	1.0

Soil profile classifications

The variability in soil classifications assigned to soil profiles within the database is presented in Table 7. Podzols represent nearly half of the database (46.9% of database) followed by Luvisols (18.8%), Brunisols and Gleysols (both at 12.8%, respectively), Organic (6.4%), then Regosols (2.3%). Once all soil classifications were assessed and updated, abbreviations and rankings for soil classifications, stoniness, rockiness, and drainage were assigned to every soil associate (where applicable).

Soil parent material

Soil parent materials were updated highlighting the variability in both modes of deposition (Table 8) and primary lithology between soil associations. From this, lithology was used to determine the mineral hardness, dominant grain size, and dominant rock type of each lithological class (Table 9).

Soil horizon classifications

The variability in soil horizon classifications is outlined in Table 10. Summarizing the variations by master horizon and dominant subscript for the mineral soil horizons (excluding forest floor) represents the range in environmental conditions, and processes, represented by the database (Table 11). From this, B horizons represent 41.78% of the database followed by C horizons (27.60%), A horizons (20.17%), then BC horizons (10.16%).

Soil physical properties

Correcting inconsistencies within the texture entries (both percentages of sand, silt, and clay, and texture classes) resulted in 86.7% of the database having texture measurements. The variability of texture between soil survey reports is outlined in Fig. 5. Most reports tend to

Table 7. Overview of soil orders separated by great group and subgroup within the database with overall representation of great groups provided.

Order	Great group	Subgroup	No. of profiles	% Total
Podzol	Ferro-Humic	Gleyed	1	4.8
		Gleyed Fragic	1	
		Orthic	23	
	Humic	Orstein	1	0.2
	Humo-Ferric	Fragic	4	42.0
		Gleyed	29	
		Gleyed Luvic	1	
		Gleyed Mini	1	
		Gleyed Orthic	18	
		Gleyed Sombric	1	
		Mini	1	
		Orstein	7	
		Orthic	155	
		Sombric	2	
Brunisol	Dystric	Eluviated	11	6.5
		Gleyed	7	
		Gleyed Eluviated	7	
		Orthic	9	
	Eutric	Gleyed Eluviated	3	0.8
		Orthic	1	
	Melanic	Gleyed Eluviated	1	1.5
		Gleyed	4	
		Orthic	3	
	Sombric	Gleyed Eluviated	1	4.0
		Gleyed	7	
		Orthic	13	
	Humic Regosol	Gleyed	3	0.6
Regosol	Regosol	Cumulic	1	1.7
		Gleyed Cumulic	2	
		Gleyed	4	
		Orthic	2	
Luvisol	Gray	Brunisolic	7	18.4
		Dark	5	
		Gleyed Brunisolic	10	
		Gleyed	8	
		Gleyed Podzolic	20	
		Orthic	3	
		Podzolic	43	
	Gray Brown	Brunisolic	1	0.4
		Gleyed	1	
Gleysol	Gleysol	Fera	3	4.0
		Orthic	16	
		Rego	2	
	Humic	Orthic	20	4.2
		Rego	2	
	Luvic	Fera	5	4.6
		Fragic	2	
		Humic	1	
		Orthic	16	

Table 7 (*concluded*).

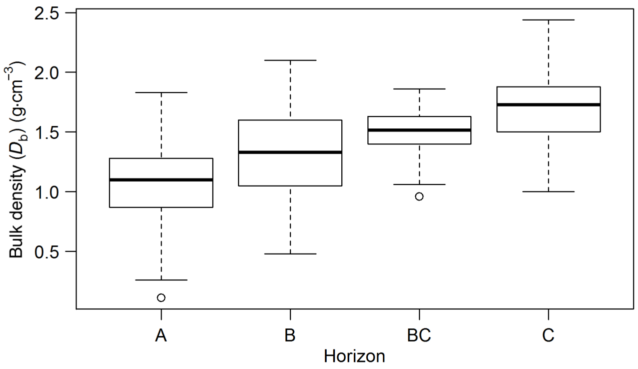
Order	Great group	Subgroup	No. of profiles	% Total
Organic	Fibrisol	Typic	5	2.5
		Terric Mesic	3	
		Terric Humic	2	
		Mesic	1	
		Terric	2	
	Mesisol	Typic	6	2.1
		Terric	3	
		Terric Fibric	1	
		Terric Humic	1	
	Humisol	Typic	3	1.7
		Terric	2	
		Terric Fibric	2	
		Terric Mesic	2	

Note: Naming follows framework outlines in Canadian Soil Classification (Soil Classification Working Group 1998).

Table 8. Summary of updated parent material modes of deposition (landforms) within the database including the quantity of associations within each mode of deposition.

Mode of deposition	No. of associations	% of total
Residual	4	3.6
Residual and colluvium	1	0.9
Colluvium and water re-worked till	4	3.6
Ablation/residual	9	8.1
Ablation	14	12.6
Ablation/basal	2	1.8
Basal	29	26.1
Basal/residual	1	0.9
Glaciomarine/basal	5	4.5
Glaciomarine	5	4.5
Glaciomarine/marine	1	0.9
Marine	3	2.7
Marine/basal	1	0.9
Glaciofluvial and marine	6	5.4
Glaciofluvial	10	9.0
Alluvium and glaciofluvial	2	1.8
Ancient alluvium	1	0.9
Alluvium	3	2.7
Lacustrine	1	0.9
Glaciolacustrine	1	0.9
Organic	8	7.2

Fig. 6. Visual comparison of range in D_b values associated with each master horizon to assess presence (or absence) of outliers. Plot derived in R Studio (RStudio Team 2020).



sample soils which fall within the center of the texture ternary diagram (texture class of loam). On the contrary, heavy clays (>80% clay), sandy clays, and pure silts remain unsampled.

CF content entries within the database remained separate from one another depending on the format of the original measurements. For example, measurements provided as a specific percentage were entered into the database apart from those provided as ranges. Combining CF measurements resulted in only 35.5% of the database having CF measurements with many soil profiles and soil associations lacking any CF measurements.

Soil structure information was provided for 73.6% of the database. Of this, structureless soils dominated (1527 samples, 61.3%) followed by granular (643 samples, 25.8%), subangular blocky (442 samples, 17.8%), platy (404 samples, 16.2%), then prismatic (six samples, 0.2%). With this, some samples were labelled as transitions between two structure classes (e.g., platy — subangular blocky), thus, these samples were provided with two structure classes.

D_b measurements within the database were generally sparse (937 samples, 38% of database). These measurements followed theoretical expectations for the most part in that (*i*) D_b for mineral soil horizons <2.4 g·cm^{-3} and (*ii*) D_b typically increased with increasing depth (Fig. 6).

Of the physical properties assessed, Table 12 summarizes the average values associated with each master

Table 9. Overview of lithology (dominant rock types) within the database and associated mineral hardness, and dominant grain size and rock type classifications.

Lithology	Hardness	Grain size	Rock type
Sandstone, gritstone, shale	6.75	C	Sandstone, gritstone, shale
Mafic volcanics, gabbro, diorite	6.6	M–C	Mafics, gabbro, diorite
Granite and metamorphic	6.5	M–C	Granite, metamorphic
Granite, some quartzite, sandstone	6.5	M–C	Granite
Sandstone	6.5	M–C	Sandstone
Sandstone, some granite, quartzite, gneiss	6.5	M–C	Sandstone
Sandstone/undifferentiated	6.5	M–C	Sandstone
Metamorphosed rhyolite, andesite, schist, slate, granite	6.4	VF–C	Rhyolite, andesite, schist, slate, granite
Granite, gneiss, basalt, felsite	6.3	F–C	Granite, gneiss, basalt, felsite
Calcareous sandstone and quartzite, some argillite, shale	6.0	M–C	Sandstone and quartzite
Granite, quartzite, gneiss, argillite, volcanics, some sandstone	6.0	F–C	Granite, quartzite, gneiss, argillite, volcanics
Non to weakly calcareous sandstone, shale, quartzite	6.0	M–C	Sandstone, shale, quartzite
Non-calcareous sandstone and quartzite, some argillite, slate, shale, schist	6.0	M–C	Sandstone, quartzite
Sandstone/sandstone and shale or siltstone	5.8	M–C	Sandstone, shale, siltstone
Sandstone/clay from calcareous shale	5.5	M–C	Sandstone and shale
Strongly metamorphosed slate, quartzite, volcanics	5.3	VF–C	Slate, quartzite, volcanics
Highly calcareous shale, quartzite, argillite, sandstone	5.0	VF	Shale, quartzite, argillite, sandstone
Quartz, schist, igneous	5	M	Quartz, schist, igneous
Igneous, slate	4.8	VF–F	Igneous, slate
Calcareous sandstone and shale	4.5	VF–C	Sandstone and shale
Non-calcareous shale, some sandstone	4.5	VF	Shale
Sandstone and conglomerate (calcareous if un-weathered)	4.5	C	Sandstone and conglomerate
Sandstone and conglomerate (some quartz)	4.5	C	Sandstone and conglomerate
Weakly calcareous sandstone and shale	4.5	VF–C	Sandstone and shale
Metamorphosed non to weakly calcareous slate, quartzite, argillite, sandstone	4.2	M–C	Slate, quartzite, argillite, sandstone
Sandstone, conglomerate, shale	4.0	F–C	Sandstone, conglomerate, shale
Slate and argillite and sandstone and schist	4.0	VF–M	Slate, argillite, sandstone, schist
Weakly to calcareous shale, slate, quartzite, some sandstone	4.0	VF–M	Shale, slate, quartzite
Sandstone, conglomerate, mudstone	3.8	F–C	Sandstone, conglomerate, mudstone
Sandstone, shale, mudstone	3.8	VF–C	Sandstone, shale, mudstone
Calcareous shale and slate with/without limestone, argillite	3.5	M–C	Shale and slate
Slate	3.5	F	Slate
Calcareous shale, slate, quartzite, argillite	3.3	M–C	Shale, slate, quartzite, argillite
Calcareous Shale	3.0	VF	Shale
Weakly calcareous shale, mudstone	2.75	VF	Shale, mudstone
Clay	2.0	VF	Clay

Note: VF, very fine; F, fine; M, moderate; and C, coarse.

horizon with dominant subscript to provide a broad representation of the variability of soil properties. From these, it is apparent that clay content is higher in gleyed and Bt horizons, CF content and D_b increase with increasing depth, SOM content decreases with increasing depth, is lower in eluviated horizons and higher in illuviated horizons (as expected), decreasing D_b (increasing SOM) increases FC and PWP and that, of the soil colloids, SOM has a stronger impact on FC and PWP than clay content.

Discussion

The rationale for a province-wide compilation of soil survey reports stemmed from the need for

Table 10. Variability in soil horizon classification encountered within the database, separated by master horizons, followed by primary subscripts, resulting in 180 unique soil horizons.

Master horizon	Primary subscript	Variations
O (Organic)	Of	Of, "Of, Om", Of$_1$, Of$_2$
	Om	Om, Om$_1$, Om$_2$, Om$_3$, Om$_4$, Om$_5$
	Oh	Oh, Oh$_1$, Oh$_2$, Oh$_3$, Oh$_4$
	Ol	Of
	Oco	Oco
LFH (Forest Floor)	L	L, LF, LF$_2$, LFH
	F	F, F$_1$, F$_2$, FH
	H	H, H$_2$, HC
A	Ae, Aeg	Ae, Ae$_1$, Ae$_2$, Aeh, Aej, Aeg, Aeg$_1$, Aeg$_2$, Aegj, 2Aeg, Aexjg, Aejg
	Ah, Ahg	Ah, Ah$_1$, Ah$_2$, Ahb, Ahg, Ahgj
	Ahe, Aheg	Ahe, Aheg, Ahejg, Ahegj
	Ap, Apg	Ap, Apg, Apgj
B	Bf, Bfg	Bf, Bft, Bftg, Bfc, "Bf, Bfj", "Bf, Bm", Bf$_1$, Bf$_2$, Bf$_3$, Bf$_4$, Bfg, Bfg$_1$, Bfg$_2$, Bfgj, Bfcg, Bfcjg, Bfgcj
	Bfh, Bfhg	Bfh, Bfht, Bfhc, Bfh$_2$, Bfhg, Bfhgj, Bfhgj$_2$
	Bh	Bh, Bhcg
	Bg	Bg, Bg$_1$, B$_2$, Bgj, 2Bg, Bgc, Bgx, Bgf, Bgfcc
	Bhf, Bhfg	Bhf, Bhfg, Bhfgj, Bhfg$_2$, Bhfjg
	Bm, Bmg	Bm, Bm$_1$, Bm$_2$, 2Bm, Bmg, Bmx, Bfj, Bfjg, Bfjg$_2$, Bfjg$_3$, Bfjgj, Bfjg, Bfjgc, Bfjgjc, Btj, 2Btj, 2Btjgj, Btjg, Btjg$_1$
	Bt, Btg	Bt, "Bt, C", Bt$_1$, Bt$_2$, Bt$_3$, Bt$_4$, 2Bt, 2Bt$_2$, Btg, Btg$_1$, Btg$_2$, Btgj, "Btgj or Cgj", "Btgj, Cgj", Btgj$_1$, Btgj$_2$, Btgj$_3$, Btgk, Btgj$_2$, 2Btg, 2Btg2, 2Btgj, Btxg, Btxjgj
BC		BC, BCgj, BCx
C	C, Cg	C, C$_1$, C$_2$, C$_3$, C$_4$, C$_5$, 2C, 2C$_1$, 2C$_2$, 2C$_3$, 3C, Cg, Cg$_1$, Cg$_2$, Cg$_3$, Cg$_4$, Cg$_5$, Cgj, Cgj$_1$, Cgj$_2$, Cgj$_3$, 2Cg, 2Cg$_1$, 2Cg$_2$, 2Cg$_3$, 2Cgj, 3Cg, 3Cgj, 4Cg, 4Cgj, 5Cg, 6Cg,
	Ck, Ckg	Ck, Ckg, Ckg$_1$, Ckg$_2$, Ckg$_3$, Ckgj, 2Ckgj
	Cx, Cxg	Cx, 2Cxj, Cxgj
R	R	R

Table 11. Variability in master horizons and dominant subscripts for mineral soil horizons within the database.

Master horizon	Dominant subscript(s)	Count	% Representation
A	Ae	254	10.2
	Aeg	123	4.9
	Aeb	11	0.4
	Ah	68	2.7
	Ahb	1	0.0
	Ahe	37	1.5
B	Bf	371	14.9
	Bfg	72	2.9
	Bg	72	2.9
	Bfh	103	4.1
	Bh	2	0.1
	Bhf	90	3.6
	Bm	179	7.2
	Bt	98	3.9
	Btg	81	3.3
BC	BC	185	7.4
	BCg	62	2.5
C	C	379	15.2
	Cg	260	10.4
	Ck	42	1.7

understanding how underlying soils vary, both locally and regionally, for better land and resource management. Having one harmonized database for all soil information, connected to a spatial map defining soil boundaries, supersedes that of having to review and compare multiple stand-alone soil surveys conducted over a broad period. Such a database allows users to quickly gain access to all available soil information found within the original soil surveys without having to access each survey individually. In addition, compiled soil information can be utilized for modeling the relationships between soil properties and soil-forming factors instead of having to manually compile the data, significantly reducing the pre-processing time required to complete analyses. Finally, with growing concerns around climate change, a harmonized database allows for the determination of carbon storage by soil type at a much larger scale than county-by-county (Carré et al. 2007; Aksoy et al. 2009; Grimm and Behrens 2010; Poggio et al. 2013).

The interest in amalgamating soil survey data into a harmonized database also stemmed from the past efforts in combining soil information into harmonized databases

Table 12. Summary of average soil property values associated with each individual master horizon and master horizon with dominant subscript(s).

Master horizon	Dominant subscript(s)	Depth (cm)	Sand (%)	Silt (%)	Clay (%)	CF (%)	D_b (g·cm^{-3})	SOM (%)	FC (%)	PWP (%)
A	**A**	8.8	42.9	39.5	17.6	8.5	1.1	3.8	32.5	8.1
	Ae	9.1	44.6	38.9	16.5	7.7	1.2	2.0	28.6	6.9
	Aeg	6.2	29.3	49.4	21.6	—	1.3	2.2	29.3	10.3
	Aeb	33.8	56.0	28.9	15.1	12.5	1.6	0.6	18.3	5.4
	Ah	4.4	30.0	49.7	20.3	19.4	0.6	15.1	58.5	13.4
	Ahb	26.5	—	—	—	—	—	—	—	—
	Ahe	7.3	31.6	37.2	31.2	5.0	0.7	10.3	40.6	18.7
B	**B**	**28.1**	**46.2**	**34.1**	**19.7**	**11.8**	**1.3**	**3.5**	**27.3**	**12.0**
	Bf	25.1	53.9	30.6	15.5	12.2	1.2	3.1	26.8	10.3
	Bfg	14.0	29.3	46.0	24.8	5.0	1.3	4.7	32.4	13.3
	Bg	30.9	44.1	35.5	20.4	9.3	1.6	1.0	23.6	10.2
	Bfh	16.0	45.7	37.2	17.2	—	1.0	6.7	35.4	16.0
	Bh	11.6	84.2	7.5	8.3	—	1.1	6.9	—	—
	Bhf	9.5	40.1	40.6	19.3	10.1	0.8	14.9	37.0	26.1
	Bm	27.4	47.5	38.0	14.6	10.1	1.5	1.4	25.6	7.6
	Bt	49.3	33.5	36.0	30.5	13.0	1.7	0.5	22.2	11.5
	Btg	6.3	25.9	49.3	24.9	—	1.2	1.2	25.5	5.4
BC	**BC**	**44.5**	**58.2**	**29.2**	**12.5**	**20.6**	**1.5**	**1.0**	**18.9**	**5.9**
	BC	44.7	58.6	29.0	12.5	20.6	1.5	0.9	18.8	5.9
	BCg	33.5	37.3	47.5	15.3	—	1.6	2.4	22.0	7.7
C	**C**	**74.2**	**47.5**	**34.2**	**18.3**	**18.7**	**1.7**	**0.5**	**19.8**	**6.7**
	C	74.8	47.6	34.1	18.3	18.7	1.7	0.5	19.8	6.8
	Cg	36.8	40.1	39.6	20.3	—	1.9	0.3	19.8	5.0
	Ck	7.5	35.5	26.3	38.2	—	1.8	—	22.0	14.1

Note: The bolded rows represent the average soil property values for the master horizons (e.g., "A" for all A horizons). CF, coarse fragment; SOM, soil organic matter; FC, field capacity, PWP, permanent wilting point.

for numerous applications. For example, Velmurugan et al. (2009), Dobos et al. (2010), Sulaeman et al. (2013), Kristensen et al. (2019), and Lark et al. (2019) utilized soil profile data from different sources to develop harmonized soil profile databases for application in digital soil mapping (DSM). Alternatively, the Soil Survey Staff within the United States Department of Agriculture's Natural Resource Conservation Center harmonized soil survey information into two databases, Soil Survey Geographic (SSURGO) Database and Web Soil Survey, each with online spatial applications. Such databases allow users to have full access to all soil survey information for any geographic location (Soil Survey Staff 2020*a*, 2020*b*). Additionally, the International Institute for Applied Systems Analysis (IIASA) and the United Nations' Food and Agriculture Organization (FAO) addressed the need for a global, standardized database representing soils from around the world (Nachtergaele et al. 2010). These studies demonstrate the need and applicability of harmonized soils databases for identifying soils for different land uses and DSM research.

Although the framework introduced in this study is straightforward, it presents a novel framework to harmonizing soil surveys into a multi-purpose database in a Canadian context where soil surveys have yet to be amalgamated into a single database for application. This framework can be applied to other provinces and territories or built upon by incorporating additional soil surveys to develop regional, or a national, database for available soil information. This process is more straightforward if the original soil surveys adhere to the Expert Committee on Soil Survey (1982) and the Soil Classification Working Group (1998). Table 13 outlines the availability of both detailed and reconnaissance soil surveys on a provincial/territorial basis across Canada. From this, it is apparent that there is a substantial amount of information that could be combined into harmonized and standardized databases with many applications.

The amalgamated database for NB supersedes that of the SNB and FSNB reports. The limitation with these original soil reports is that they aggregate various soil profile information into individual, generalized profile summaries for each soil associate in each soil association (or forest soil associations for FSNB). Doing so is counterintuitive because it loses much of the inherent variability in soil properties found across soil-forming factors as they vary across the province. Soils are complex in nature, and generalizing this information can lead to misleading interpretations of the data as well as broad spatial delineations of each soil association. Thus, much information is lost with this form of aggregation.

Table 13. Overview of available detailed and reconnaissance soil surveys for each Province/Territory across Canada with the range in vintages.

Province/Territory	No. of detailed soil surveys	No. of reconnaissance soil surveys	Range in vintage (years)
Alberta	37	34	1925–1996 (71)
British Columbia	23	28	1960–1992 (32)
Manitoba	73	24	1939–2013 (74)
New Brunswick	21	4	1940–2005 (65)
Newfoundland and Labrador	28	8	1957–2002 (45)
Northwest Territories and Nunavut	19	8	1953–2008 (55)
Nova Scotia	18	2	1954–1993 (39)
Ontario	71	4	1930–1998 (58)
Prince Edward Island	3	2	1950–1988 (38)
Quebec	79	5	1936–2013 (77)
Saskatchewan	201	0	1958–1998 (40)
Yukon	8	3	1943–1997 (54)

Unlike the SNB and FSNB reports, the NSDB maintains multiple soil profiles for each soil associate, but again, this information is summarized losing some of the inherent variability of the soil properties found within each soil associate. For example, the NSDB provides two soil profiles for the Johnville soil associate (imperfectly drained member of the Holmesville soil association), whereas the harmonized database provides 11 soil profiles for the Johnville soil associate. Unlike the SNB, FSNB, and NSDB data, the amalgamated database presented in this study maintains the variability found within each soil association and soil associate and presents the variability in an organized manner. This is important to the end users who, with all soil information made available, may be better able to interpret what is seen on the land base, whereas the aggregated and generalized summaries provided in the SNB and FSNB reports may prevent this.

All the outlined efforts demonstrate the need for standardized protocols for collecting, and recording, soil information. Systematically sampling soils, regardless of topographic position, land use, and drainage class, would provide a much more robust data set for use in future endeavours. With the standardization and harmonization processes applied, there are limitations with the database. First, it is apparent from these efforts that the type, and amount, of data collected at each sampled location remain inconsistent (Table 14). For example, slope steepness and position were only determined for 95 soil associates. Also, horizon-specific measurements were inconsistent in measurement frequencies. D_b measurements were made more frequently than CF content. Second, limitations arise when assigning the middle class to a range in values as well as assigning one value to a categorical description. For example, if the soil texture was assigned as "clay loam" the sand content can vary from 20% to 45%, whereas clay content can vary from 27% to 40% based on the texture ternary diagram. As such, assigning the middle point of 32% for sand and 32.5% for clay may result in misinterpretation of the class, or add error when utilizing the data for analyses. The same holds true for the drainage classes and CF content descriptions within this database. Finally, climate information was unavailable for each soil associate within the original surveys, thus, and although an important soil-forming factor, climate information is not included within the database. This is particularly important because the database houses soil profiles from across the province where climate varies from the lowlands in the east to the highlands in the north and south of the province (Pronk and Allard 2003).

An inconsistency which could not be addressed in this amalgamation and harmonization procedure is the inconsistencies in sampling size of different soil associates for developing soil surveys. Most soil associates have more than one profile described, depending on frequency of occurrence within different surveys. For example, the Holmesville soil association occurred in eight soil surveys. As a result, the well-drained soil associate (also called Holmesville) has 18 profiles within the database. Its moderatelywell-drained associate, Johnville, has 12 profiles within the database, followed by the poorly drained associate, Poitras, also with 12 profiles. It is common for some soil associations to occur within different surveys, and therefore, have multiple profiles within the database. In contrast, some of the less-common soil associations lacked a single soil profile altogether (e.g., Aulac, Babineau, Becaguimec, Belledune, Big Bald Mountain, Blackland, Caissie, Bottomland, Catamaran, Clearwater, Escuminac, Jacquet River, Kingston, Research Station, and Tetagouche). To have a fully comprehensive representation of forest soil conditions across NB, soil profiles are needed for these soil associations.

In addition to inconsistencies within the database, many issues became apparent with the spatial

Table 14. Overview of measured soil attributes, including general characteristics and horizon-specific soil properties, within amalgamated database with overall completeness.

Soil property		No. of horizons	% Complete
Overview	Attribute		
General	Surface expression	2490	100.0
	Slope	1142	45.9
	Soil association	2490	100.0
	Soil associate	2490	100.0
	Parent material	2490	100.0
	Drainage	2467	99.1
	Soil classification	2453	98.5
Horizon-specific	Horizon depth (cm)	2490	100.0
	Texture (% sand, silt, and clay)	1306	52.5
	Texture (class)	2158	86.7
	Structure	1832	73.6
	Coarse fragment content (%)	885	35.5
	pH (H_2O)	1535	61.6
	pH ($CaCl_2$)	647	26.0
	Bulk density ($g \cdot cm^{-3}$)	938	37.7
	Soil organic matter content (%)[a]	1202	48.3
	Base saturation (%)	466	18.7
	Cation exchange capacity	659	26.5
	Field capacity (−33 kPa)	846	34.0
	Permanent wilting point (−1500 kPa)	738	29.6
	Ca ($meq \cdot 100\ g^{-1}$)	976	39.2
	Mg ($meq \cdot 100\ g^{-1}$)	960	38.6
	K ($meq \cdot 100\ g^{-1}$)	973	39.1

[a] Counts for soil organic matter content includes converted %C readings.

representation of soil association boundary delineations across NB, including

1. incompleteness in terms of outlining the geographic locations of soil associates within soil association delineations, with these varying in scale and resolution, and with many focused solely on agricultural lands (Pitty 1979; Zhu and Mackay 2001; Adhikari et al. 2012; Odgers et al. 2014). In addition, past survey practices generally addressed soil variations at the 1:10 000 scale (and often coarser) (Table 1). As a result, higher resolution spatial pedological variations which influence crop and forest productivity, and root growth via nutrient and water retention, remain unrecognized (Parr et al. 1992; Southorn 2003; Keys 2007; Taylor et al. 2013).
2. implied differences in soil association conditions, and extent, across arbitrary and discrete survey boundaries. This form of recording and mapping assumes that soil properties abruptly change at the boundaries of each soil association, due to changes in soil-forming factors. However, in the field no such discrete boundaries exist as soil properties change along dynamic continua. In addition, boundaries, and delineations, of soil associations are often inconsistent with adjacent delineations of a neighboring surveys due to different parties conducting the soil surveys.
3. inconsistencies between surveyed soil associates and delineated soil association boundaries such that (*i*) not all spatially mapped soil associations occur within the original soil survey reports (e.g., Becaguimec, Big Bald Mountain, Catamaran, Jacquet River, Kingston, Popple Depot, and Tetagouche), (*ii*) conversely, some of the surveyed soil associations are not spatially represented in the existing soil association delineations for the province (Table 15).

The soil database, amalgamated from 17 soil surveys for NB, Canada, is intended to provide a comprehensive overview of forest soil conditions as well as provide a framework for amalgamating and harmonizing soil survey information. Through careful cross-referencing, all data entries were examined to ensure they coincided with soil association and horizon-specific expectations as outlined in the Soil Classification Working Group (1998). Although amalgamated and harmonized, the aspatial database remains incomplete in terms of measurement gaps for horizon-specific physical and chemical properties. Although additional soil properties included

Table 15. Surveyed soil associations not currently represented in the province-wide spatial soil association data set.

Aldouane
Anagance
Baie du Vin
Bellefleur
Benedict
Big Hole
Boston Brook
Bransfield
Bretagneville
Caraquet
Chockpish
Dorchester
Flemming
Fundy
Galloway
Green River
Green Road
Gulquac
Harquail
Island Lake
Jardine
Jeffries Corner
Kingsclear
Knightville
Kouchibouguac
Lord and Foy
Lower Ridge
Maliseet
Monquart
Mount Hope
Parsons Brook
Guimond River
Petitcodiac
Queenville
Quisbis
Richibucto
Riley Brook
Salem
Shemogue
St. Charles
Sussex
Tobique
Tormentine
Upper Caraquet
Violette
Wakefield

within the original surveys were also entered into the database, emphasis was placed on those which are known to have strong impacts on rooting as well as those typically used as predictors in developing pedotransfer functions (PTFs) (Table 14). Future work focuses on standardizing individual soil properties to (*i*) correct for inconsistent units of measurement and (*ii*) predict absent soil property values by way of PTF.

Of the available soil surveys (both included and excluded from the database), 53.5% of NB has survey coverage with the spatial coverage of the surveys utilized in this study only representing 39.2% of NB (28 000 km^2) (Fig. 3). Thus, additional efforts are required in updating soil representations across the province, either by soil association delineations or digital soil mapping of individual soil properties.

Conclusions

This study demonstrated means to amalgamate and harmonize soil information found within county-based soil surveys for the Province of New Brunswick, Canada, as a case study. This procedure, utilizing 17 soil surveys, has resulted in an amalgamated database containing 106 soil associations, 243 soil associates, and 522 soil horizon sequences (profiles). This framework demonstrates techniques which can easily be adapted to other locations in which soil surveys had been conducted over long periods of time. Such implementation allows for more consistent and standardized soil information at provincial, or even national, scales. This study expressed techniques to address and correct the prominent inconsistencies arising from amalgamating soil surveys. To summarize, this study corrected for

1. inconsistent labelling — soil associations, soil associates, soil-forming factors (landforms, lithology, topographic surface expressions, and vegetation), soil drainage, and soil classifications,
2. inconsistent methods of measurement — texture (classes and percentages), CF content (ranges, descriptions, and percentages),
3. inconsistent units of measurement — SOM content (%C, %SOM, and LOI), and moisture retention at FC and PWP (bar, atm, kPa in volumetric and gravimetric form),
4. inconsistent data recording (gaps) — drainage, soil classifications, landforms, lithology, CF content.

Creation of this database expedites PTF development for gap-filling and summarizing and quantifying both soil association and associates via similarities and differences. Once complete, this will enable spatially re-digitizing the updated database using already-existing soil association delineations, followed by revising these to ensure topographic mapping consistencies.

The amalgamated and harmonized database represents a way forward for large-scale soil studies and continuing with, otherwise discontinued, soil surveys. Many provinces house an array of soil surveys (Table 13) with which the amalgamation outlined in this study can be applied to develop either province-specific or an updated national soils database. Such databases can be used for applications in digital soil mapping, modeling soil property relationships, as well as spatial re-delineation of soil associations and associates, based on updated understanding of soil-forming factors. In combination, this information can assist users in understanding the spatial variability of soils, how soils vary with changing land uses, carbon stock prediction under different climate change scenarios, and asset management in terms of soil erosion and sedimentation modeling.

Acknowledgements

We would like to acknowledge the dedication and commitment of the soil surveyors whom conducted the original soil surveys throughout NB. In addition, we acknowledge CANSIS for the compilation of soil surveys for NB which are made freely available online in a userfriendly manner. Finally, we would like to thank the authors and developers of the NSDB, and SNB and FSNB reports. This data proved crucial in assisting with the harmonization of soil names and attributes within this study.

Conflicts of Interest

The authors have declared that no conflicts of interest exist with respect to this publication.

References

Aalund, H., Wicklund, R.E. 1950. Soil Survey Report of Southeastern New Brunswick, New Brunswick Soil Survey Report No. 3. Fredericton, New Brunswick.

Adhikari, K., Kheir, R.B., Greve, M.B., Bøcher, P.K., Malone, B.P., Minasny, B., et al. 2012. High-resolution 3-D mapping of soil texture in Denmark. Soil Sci. Soc. Am. J. **77**: 1–17.

Aksoy, E., Özsoy, G., and Dirim, M.S. 2009. Soil mapping approach in GIS using Landsat satellite imagery and DEM data. Afr. J. Agric. Res. **4**: 1295–1302. doi:10.2136/sssaj2012.0275.

Anderson, D.W., and Smith, C.A.S. 2011. A history of soil classification and soil survey in Canada: personal perspectives. Can. J. Soil Sci. **91**: 675–694. doi:10.4141/cjss10063.

Birkeland, P. 1999. Soils and geomorphology. 3rd ed. Oxford University Press, Oxford, UK.

Canadian Soil Information Service. 2000. National soils database.

Canadian Soil Information Service. 2012. Soil survey reports for New Brunswick. [Online]. Available from http://sis.agr.gc.ca/cansis/publications/surveys/nb/index.html [21 May 2019].

Carré, F., McBratney, A.B., Mayr, T., and Montanarella, L. 2007. Digital soil assessments: beyond DSM. Geoderma, **142**: 69–79. doi:10.1016/j.geoderma.2007.08.015.

Carter, M.R., and Gregorich, E.G. (eds.). 2006. Soil sampling and methods of analysis. 2nd ed. CRC Press, Boca Raton, FL, USA.

Chaudhari, P.R., Ahire, D.V., Ahire, V.D., Chkravarty, M., and Maity, S. 2013. Soil bulk density as related to soil texture, organic matter content and available total nutrients of Coimbatore soil. Int. J. Sci. Res. Publ. **3**: 1–8.

Colpitts, M.C., Fahmy, S.H., Macdougall, J.E., Ng, T.T.M., Mclnnis, B.G., and Zelazny, V.F. 1995. Forest soils of New Brunswick. Agriculture and Agri-Food Canada Research Branch, Natural Resources Canada Canadian Forest Service.

Dobos, E., Bialko, T., Micheli, E., and Kobza, J. 2010. Legacy soil data harmonization and database development. Pages 309–327 in J.L. Boettinger, D.W. Howell, A.C. Moore, A.E. Hartemink, and S. Kienast-Brown, eds. Digital soil mapping: progress in soil science. Springer, Dordrecht, the Netherlands.

Dube, R. 1982. St. Quentin – Kedgwick Soil Survey Report. Fredericton, New Brunswick.

Environmental Systems Research Institute. 2018. ArcGIS Desktop.

Expert Committee on Soil Survey. 1982. The Canada soil information system manual for describing soils in the field: 1982 revised. Expert Committee on Soil Survey, Ottawa, ON, Canada.

Fahmy, S.H., and Rees, H.W. 1989. Soils of the Woodstock–Florenceville Area, Carleton County, New Brunswick, New Brunswick Soil Survey Report No. 14, LRRC Contribution No. 88–85. Fredericton, New Brunswick.

Fahmy, S.H., and Rees, H.W. 1992. Soils of the Woodstock–Florenceville Area, Carleton County, New Brunswick, New Brunswick Soil Survey Report No. 14, CLBRR Contribution No. 92–33. Fredericton, New Brunswick.

Fahmy, S.H., and Rees, H.W. 1996. Soils of the Woodstock–Florenceville Area, Carleton County, New Brunswick, New Brunswick Soil Survey Report No. 14, CLBRR Contribution No. 96–02. Fredericton, New Brunswick.

Fahmy, S.H., Rees, H.W., and Soley, T.J. 2001. Soils of the Woodstock–Florenceville Area, Carleton County, New Brunswick, New Brunswick Soil Survey Report No. 14, PRC Contribution No. 01–01. Fredericton, New Brunswick.

Fahmy, S.H., Hann, S.W.R., and Jiao, Y. 2010. Soils of New Brunswick: the second approximation. Agriculture and Agri-Food Canada, Fredericton, NB, Canada.

Furze, S., Ogilvie, J., and Arp, P.A. 2017. Fusing digital elevation models to improve hydrological interpretations. J. Geogr. Inf. Syst. **9**: 558–575. doi:10.4236/jgis.2017.95035.

Grimm, R., and Behrens, T. 2010. Uncertainty analysis of sample locations within digital soil mapping approaches. Geoderma, **155**: 154–163. doi:10.1016/j.geoderma.2009.05.006.

Guertin, R.K., Schuppli, P.A., Sheldrick, B.H., Wires, K.C., and Zebchuk, W.D. 1984. Analytical methods manual 1984. Land Resource Research Institute, Research Branch, Agriculture Canada, Ottawa, ON, Canada.

Hamilton, N.E., and Ferry, M. 2018. ggtern: ternary diagrams using ggplot2. J. Stat. Softw. Code Snippets, **87**: 1–17. doi:10.18637/jss.v087.c03.

High Country Research and Development 1972. Mount Carleton Provincial Park, New Brunwick. Fredericton, New Brunswick.

Holmstrom, D.A. 1986. Soils of the Sussex Area of New Brunswick, New Brunswick Soil Survey Report No. 10, LRRC Contribution No. 83–38. Truro, Nova Scotia, Canada.

Jenny, H. 1941. Factors of soil formation: a system of quantitative pedology. McGraw-Hill, New York, NY, USA.

Keys, K. 2007. Forest soil types of Nova Scotia — identification, description, and interpretation. Nova Scotia Department of Natural Resources Manual FOR 2007-2. Dartmouth, NS, Canada.

Kristensen, J.A., Balstrom, T., Jones, R.J.A., Jones, A., Montanarella, L., Panagos, P., and Breuning-Madsen, H. 2019. Development of a harmonised soil profile analytical database for Europe: a resource for supporting regional soil management. Soil, **5**: 289–301. doi:10.5194/soil-5-289-2019.

Langmaid, K.K., and Losier, J.G. 1973. Soil Survey of Lepreau Provincial Park. Fredericton, New Brunswick.

Langmaid, K.K., MacMillan, J.K., Losier, J.G., and Millette, J.F.G. 1964. Description of sandy soils in cleared areas of coastal Kent and Southern Northumberland Counties, New Brunswick. Fredericton, New Brunswick.

Langmaid, K.K., MacMillan, J.K., and Losier, J.G. 1980. Soils of Madawaska County, New Brunswick. New Brunswick Soil Survey Report No. 8. Fredericton, NB, Canada.

Lark, R.M., Ander, E.L., and Broadley, M.R. 2019. Combining two national-scale datasets to map soil properties, the case of available magnesium in England and Wales. Eur. J. Soil Sci. **70**: 361–377. doi:10.1111/ejss.12743. PMID:30983873.

MacMillan, J.K., and Chalifour, A.P. 1980. The Land Base Study of Havelock Parish Phase 1. Fredericton, New Brunswick.

Mapping System Working Group. 1981. A soil mapping system for Canada: revised. Ottawa, ON, Canada.

McKeague, J.A., and Stobbe, P.C. 1978. History of soil survey in Canada 1914–1975. 11th ed. Research Branch, Canada Department of Agriculture, Ottawa, ON, Canada.

Mellerowicz, K.T., Rees, H.W., Chow, T.L., and Ghanem, I. 1993. Soils of the Black Brook Watershed, St. Andre Parish, Madawaska County, New Brunswick. Fredericton, New Brunswick.

Michalica, K., Rees, H.W., Dillon, M.J., Loro, P.J., Toner, P.M., Donald, R.G., and Parks, J.J. 2000. Soils of the Acadian Peninsula, Gloucester County, New Brunswick. New Brunswick Soil Survey Report No. 18. Fredericton, NB, Canada.

Millette, J.F.G., and Langmaid, K.K. 1964. Soil Survey of the Andover- Plaster Rock Area, New Brunswick, New Brunswick Soil Survey Report No. 5. Fredericton, New Brunswick.

Nachtergaele, F., van Velthuizen, H., Verelst, L., Batjes, N., Dijkshoorn, K., van Engelen, V., et al. 2010. The harmonized world soils database. 19th World Congress of Soil Science, Soil Solutions for a Changing World, Brisbane, Australia.

Odgers, N.P., Sun, W., McBratney, A.B., Minasny, B., and Clifford, D. 2014. Disaggregating and harmonising soil map units through resampled classification trees. Geoderma, **214**–**215**: 91–100. doi:10.1016/j.geoderma.2013.09.024.

Parr, J.F., Hornick, S.B., and Papendick, R.I. 1992. Soil quality: the foundation of a sustainable agriculture.

Pitty, A.F. 1979. Geography and soil properties. Methuen, London, UK.

Poggio, L., Gimona, A., and Brewer, M.J. 2013. Regional scale mapping of soil properties and their uncertainty with a large number of satellite-derived covariates. Geoderma, **209**–**210**: 1–14. doi:10.1016/j.geoderma.2013.05.029.

Pollacco, J.A.P. 2008. A generally applicable pedotransfer function that estimates field capacity and permanent wilting point from soil texture and bulk density. Can. J. Soil Sci. **88**: 761–774. doi:10.4141/CJSS07120.

Pronk, A.G., and Allard, S. 2003. Landscape map of New Brunswick.

Rees, H.W., and Fahmy, S.H. 1983. Soils of the Senator Herve J. Michaud Experimental Farm, Agriculture Canada, Buctouche, New Brunswick. Fredericton, New Brunswick.

Rees, H.W., and Fahmy, S.H. 1984. Soils of the Agriculture Canada Research Station, Fredericton, New Brunswick. Fredericton, New Brunswick.

Rees, H.W., and Fahmy, S.H. 1992. Soils of the Agriculture Canada Benton Ridge Potato Breeding Substation, Benton Ridge, New Brunswick. Fredericton, New Brunswick.

Rees, H.W., Langmaid, K.K., Veer, C., Wang, C., Wells, R.E., Fahmy, S.H., and Losier, J.G. 1992. Soils of the Chipman-Minto-Harcourt Region of New Brunswick, New Brunswick Soil Survey Report No. 11, CLBRR Contribution No. 84–60. Fredericton, New Brunswick.

Rees, H.W., Duff, J.P., Colville, S., and Chow, T.L. 1995. Soils of Selected Agricultural Areas of Moncton Parish, Westmorland County, New Brunswick, New Brunswick Soil Survey Report No. 15, CLBRR Contribution No. 92–24. Ottawa, Ontario, Canada.

Rees, H.W., Duff, J.P., Soley, T.J., Colville, S., and Chow, T.L. 1996. Soils of Selected Agricultural Areas of Shediac and Botsford Parishes, Westmorland County, New Brunswick, New Brunswick Soil Survey Report No. 16. Ottawa, Ontario, Canada.

Rees, H.W., Duff, J.P., Soley, T.J., and Chow, T.L. 1998. Soils of Selected Agricultural Areas of Dorchester Parish, Westmorland County, New Brunswick, New Brunswick Soil Survey Report No. 17. Fredericton, New Brunswick.

Rees, H.W., Fahmy, S.H., Wang, C., and Wells, R.E. 2005. Soils of Central and Northern New Brunswick, New Brunswick Soil Survey Report No. 12, PRC Contribution No. 05–01. Fredericton, New Brunswick.

Romano, N., and Palladino, M. 2002. Prediction of soil water retention using soil physical data and terrain attributes. J. Hydrol. **265**: 56–75. doi:10.1016/S0022-1694(02)00094-X.

RStudio Team. 2020. RStudio: integrated development for R.

Soil Classification Working Group. 1998. The Canadian system of soil classification. 3rd ed. National Research Council of Canada, Ottawa, ON, Canada.

Soil Survey Staff. 2020*a*. Web soil survey. [Online]. Available from http://websoilsurvey.sc.egov.usda.gov.

Soil Survey Staff. 2020*b*. Soil survey geographic (SSURGO) database.

Southorn, N. 2003. The role of soil quality criteria in assessing farm performance. 14th International Farm Management Congress, Perth, Australia. pp. 1–11.

Stobbe, P.C. 1940. Soil survey of the Fredericton-Gagetown area, New Brunswick. New Brunswick Soil Survey Report No. 1. Fredericton, NB, Canada.

Stobbe, P.C., and Aalund, H. 1944. Soil Survey of the Woodstock Area, New Brunswick, New Brunswick Soil Survey Report No. 2. Fredericton, New Brunswick.

Subcommittee on Methods of Analysis of Canada Soil Survey Committee. 1978. Manual on soil sampling and methods of analysis. 2nd ed. Canadian Society of Soil Science.

Sulaeman, Y., Minansy, B., McBratney, A., Sarwani, M., and Sutandi, A. 2013. Harmonizing legacy soil data for digital soil mapping in Indonesia. Geoderma, **192**: 77–85. doi:10.1016/j.geoderma.2012.08.005.

Taylor, J.A., Jacob, F., Galleguillos, M., Prévot, L., Guix, N., and Lagacherie, P. 2013. The utility of remotely-sensed vegetative and terrain covariates at different spatial resolutions in modelling soil and watertable depth (for digital soil mapping). Geoderma, **193**–**194**: 83–93. doi:10.1016/j.geoderma.2012.09.009.

Velmurugan, A., Krishan, G., Dadhwal, V.K., Kumar, S., Swarnam, T.P., and Saha, S.K. 2009. Harmonizing soil organic carbon estimates in historical and current data. Curr. Sci. **97**: 554–558.

Wang, C., and Rees, H.W. 1983. Soils of the Rogersville–Richibucto Region of New Brunswick, New Brunswick Soil Survey Report No. 9, LRRI Contribution No. 89. Ottawa, Ontario, Canada.

Wickham, H. 2011. The split-apply-combine strategy for data analysis. J. Stat. Softw. **40**: 1–29.

Wickham, H. 2016. ggplot2: eleganty graphics for data analysis. 2nd ed. Springer, New York, NY, USA.

Wicklund, R.E., and Langmaid, K.K. 1949. Soil Survey of Southwestern New Brunswick, New Brunswick Soil Survey Report No. 4. Fredericton, New Brunswick.

Zhu, A.-X., and Mackay, S.D. 2001. Effects of spatial detail of soil information on watershed modeling. J. Hydrol. **248**: 54–77. doi:10.1016/S0022-1694(01)00390-0.

Appendix A

Table A1. Overview of New Brunswick soil associations.

			Soil associate		
Landform	Lithology	Soil association	W-R	I-MW	VP-P
Residual	Weakly calcareous sandstone and shale	Undine	Undine	—	—
	Sandstone, gritstone, shale	Riley Brook	Riley Brook	—	—
		Cornhill	Cornhill	—	—
	Granite, gneiss, basalt, felsite	Big Bald Mountain	Big Bald Mountain	—	—
Residual and colluvium	Strongly metamorphosed slate, quartzite, volcanics	Serpentine	Serpentine	Jenkins	Adder
Ablation	Sandstone and Conglomerate (calcareous if un-weathered)	Anagance	Anagance	Dusinane	Dusinane
		Jeffries Corner	Jeffries Corner	—	—
		Aulac	Aulac	Tidnish	Tidnish
	Granite, quartzite, gneiss, argillite, volcanics, some sandstone	Irving	Irving	Goodfellow	Halls Brook
		Juniper	Juniper	Jummet Brook	McKiel
	Calcareous shale and slate with/without limestone, argillite	Caribou	Caribou	Carlingford	Washburn
		Thibault	Thibault	Guercheville	Lauzier
	Sandstone and quartzite, some argillite, slate, shale, schist	Monquart	Monquart	—	—
	Sandstone	Big Hole	Big Hole	Beaver Lake	—
		Fair Isle	Fair Isle	Black Brook	—
		Sunbury	Sunbury	Hoyt	Cork
	Highly calcareous shale, quartzite, argillite, sandstone	Jardine	Jardine	Nickel Mill	Five Fingers
	Sandstone, shale, mudstone	Becaguimec	Becaguimec	Snyder	Snyder
	Metamorphosed rhyolite, andesite, schist, slate, granite	Jacquet River	Jacquet River	—	—
Ablation/residual	Calcareous sandstone and shale	Harquail	Harquail	—	—
	Calcareous Shale	Erb Settlement	Erb Settlement	—	—
	Metamorphosed non to weakly calcareous slate, quartzite, argillite, sandstone	Glassville	Glassville	Temiscouata	Foreston
	Metamorphosed rhyolite, andesite, schist, slate, granite	Lomond	Lomond	—	—
	Sandstone and conglomerate (calcareous if unweathered)	Parleeville	Parleeville	Midland	Midland
		Queenville	Queenville	Deed	Deed
	Sandstone and quartzite, some argillite, slate, shale, schist	Quisbis	Quisbis	Dube	Big Spring
	Sandstone, gritstone, shale	Tobique	Tobique	—	—
	Slate and argillite and sandstone and schist	Boston Brook	Boston Brook	Skin Gulch	Yellow Brook
Alluvium	Undifferentiated	Interval	Interval	Waasis	East Canaan
		Sussex	Sussex	Hampton	—
	—	Bottomland	Bottomland	—	—

Table A1 (*continued*).

Landform	Lithology	Soil association	Soil associate		
			W-R	I-MW	VP-P
Ancient alluvium	Undifferentiated	Flemming	Flemming	Martial	Kelly
Alluvium and glaciofluvial	Calcareous shale, slate, quartzite, argillite	Maliseet	Maliseet	Wapske	Wapske
	Strongly metamorphosed slate, quartzite, volcanics	Benedict	Benedict	—	—
Basal	Calcareous sandstone and quartzite, some argillite, shale	Siegas	Siegas	Salmon	Bourgoin
	Calcareous shale	Kedgwick	Kedgwick	—	—
		Saltspring	Saltspring	Byrns	Byrns
	Granite, gneiss, basalt, felsite	Parry	Parry	Midway	Midway
	Granite, quartzite, gneiss, argillite, volcanics, some sandstone	Rogersville	Rogersville	Acadieville	Rosaireville
		Tuadook	Tuadook	Redstone	Lewis
	Granite, some quartzite, sandstone	Pinder	Pinder	Coronary	McAdam
		Catamaran	Catamaram	—	—
	Sandstone and conglomerate (calcareous if unweathered)	Wakefield	Wakefield	—	—
	Sandstone and quartzite, some argillite, slate, shale, schist	Holmesville	Holmesville	Johnville	Poitras
		Violette	—	Violette	Violette
	Sandstone, conglomerate, mudstone	Parsons Brook	Parsons Brook	—	—
	Sandstone, conglomerate, shale	Petitcodiac	Petitcodiac	Kings	Kings
		Salisbury	Salisbury	Harewood	Hicksville
		Salem	Salem	—	—
	Sandstone, shale, mudstone	Dorchester	Dorchester	—	—
		Knightville	Knightville	Byrns	Byrns
		Tracy	Tracy	Wirral	Rooth
	Strongly metamorphosed slate, quartzite, volcanics	Long Lake	Long Lake	Blue Mountain	Colter Mountain
	weakly calcareous Sandstone and shale	Shemogue	Shemogue	—	—
		Stony Brook	Stony Brook (Queens)	Blackville	Cambridge (Kings)
		Tormentine	Tormentine	Tidnish	Tidnish
	Weakly calcareous shale, mudstone	Kingsclear	Kingsclear	Plaster Rock	Nackawic
	Weakly to calcareous shale, slate, quartzite, some sandstone	Carleton	Carleton	Canterbury	Canterbury
		Green Road	Green Road	—	—
		Lower Ridge	Lower Ridge	—	—
	Mafic volcanics, gabbro, diorite	Kingston	Kingston	Deed	Deed
		Tetagouche	Tetagouche	—	—
	Metamorphosed rhyolite, andesite, schist, slate, granite	Popple Depot	Popple Depot	—	—
Basal/residual	Metamorphosed non to weakly calcareous slate, quartzite, argillite, sandstone	Green River	Green River	—	—
Colluvium and water reworked till	Granite, gneiss, basalt, felsite	Clearwater	Clearwater	Ogilvie Lake	Yellow Lake

Table A1 (*continued*).

			Soil associate		
Landform	Lithology	Soil association	W-R	I-MW	VP-P
Glaciofluvial	Metamorphosed non to weakly calcareous slate, quartzite, argillite, sandstone	McGee	McGee	Nason	Trafton
	Non to weakly calcareous sandstone, shale, quartzite	Victoria	Victoria	McCluskey	Cote
	Strongly metamorphosed slate, quartzite, volcanics	Britt Brook	Britt Brook	Babbit Brook	Portage Lake
	Granite	Gagetown	Gagetown	Geary	Penobsquis
		Island Lake	Island Lake	Pennfield	Pennfield
	Metamorphosed non to weakly calcareous slate, quartzite, argillite, sandstone	Grand Falls	Grand Falls	Sirois	Cyr
	Sandstone	Kennebecasis	Kennebecasis	Quispamsis	Nevers Road
		Lord and Foy	Lord and Foy	—	—
	Sandstone and conglomerate (some quartz)	Gulquac	Gulquac	—	—
	Sandstone, some granite, quartzite, gneiss	Guimond River	Guimond River	St. Oliver	St. Theodule
	Weakly to calcareous shale, slate, quartzite, some sandstone	Muniac	Muniac	Ennishore	Cyr
	—	Bransfield	Bransfield	—	—
	—	Chockpish	Chockpish	—	—
Glaciofluvial and marine	Sandstone	Aldouane	Aldouane	Marquant	—
		Baie du Vin	Baie du Vin	Napan	Fontaine
		Kouchibouguac	Kouchibouguac	Potters Mill	Vautour
		Caissie	Caissie	Robichaud	—
		Richibucto	Richibucto	Cap Limiere	Nevers Road
		Riverbank	Riverbank	Oromocto	Nevers Road
Glaciolacustrine	Undifferentiated/weakly calcareous clay	Bellefleur	Bellefleur	—	—
Glaciomarine	Calcareous sandstone and shale	Tracadie	Tracadie	Bouleau	Sheila
	Non-calcareous shale, some sandstone	Mount Hope	Mount Hope	Boland	Cambridge
	Sandstone	Babineau	Babineau	—	—
		Escuminac	Escuminac	Baie St. Anne	—
		Galloway	Galloway	Smelt Brook	Briggs Brook
Glaciomarine/basal	Sandstone/clay from calcareous shale	Upper Caraquet	Upper Caraquet	Little Shippegan	Shediac
	Sandstone/sandstone and shale or siltstone	Barrieau	Barrieau	Cote D'or	Shediac
		Buctouche	Buctouche	Michaud	Neguac
		Bretagneville	Bretagneville	—	—
	—	St. Charles	St. Charles	—	—
Glaciomarine/marine	Sandstone/clay from calcareous shale	Caraquet	Caraquet	Middle Caraquet	Neguac
Lacustrine	Clay	Fundy	—	Fundy	Canobie
Marine/basal	Undifferentiated	Research Station	Research Station	Baker Brook	—

Table A1 (*concluded*).

Landform	Lithology	Soil association	Soil associate		
			W-R	I-MW	VP-P
Marine	Undifferentiated	Belledune	Belledune	—	—
		Blackland	Blackland	—	—
		Acadia	Acadia	Acadia	Acadia
Ablation/basal	Sandstone/sandstone and shale or siltstone	Reece	Reece	Chipman	Pangburn
		Harcourt	Harcourt	Coal Branch	Grangeville
Organic	Undifferentiated	Lavilette	—	—	Lavilette
		Acadie Siding	—	—	Acadie Siding
		Bog	—	—	Bog
		Fen	—	—	Fen
		Swamp	—	—	Swamp
		Legaceville	—	—	Legaceville
		Chelmsford	—	—	Chelmsford
		St. Quentin	—	—	St. Quentin