remote sensing

MDPI

Communication

An Extensive Field-Scale Dataset of Topsoil Organic Carbon Content Aimed to Assess Remote Sensed Datasets and Data-Derived Products from Modeling Approaches

Ahmed Laamrani [1,2,*], Paul R. Voroney [3], Daniel D. Saurette [3,4], Aaron A. Berg [2], Line Blackburn [3], Adam W. Gillespie [3] and Ralph C. Martin [5]

1 Center for Remote Sensing Applications (CRSA), Mohammed VI Polytechnic University (UM6P), Ben Guerir 43150, Morocco
2 Department of Geography, Environment & Geomatics, University of Guelph, Guelph, ON 1G 2W1, Canada
3 School of Environmental Sciences, University of Guelph, Guelph, ON N1G 2W1, Canada
4 Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA), Guelph, ON N1G 4Y2, Canada
5 Department of Plant Agriculture, University of Guelph, Guelph, ON 1G 2W1, Canada
* Correspondence: ahmed.laamrani@um6p.ma or alaamran@uoguelph.ca

Abstract: The geosciences suffer from a lack of large georeferenced datasets that can be used to assess and monitor the role of soil organic carbon (SOC) in plant growth, soil fertility, and CO_2 sequestration. Publicly available, large field-scale georeferenced datasets are often limited in number and design to serve these purposes. This study provides the first publicly accessible dataset of georeferenced topsoil SOC measurements (n = 840) over a 26-hectare (ha) agricultural field located in southern Ontario, Canada, with a sampling density of ~32 points per ha. As SOC is usually influenced by site topography (i.e., slope and landscape position), each point of the database is associated with a wide range of remote sensing topographic derivatives; as well as with normalized difference vegetation index (NDVI) based value. The NDVI data were extracted from remote sensing Sentinel-2 imagery from over a five-year period (2017–2021). In this paper, the methodology for topsoil sampling, SOC measurement in the lab, as well as producing the suite of topographic derivatives is described. We discuss the opportunities that the database offers in terms of spatially explicit and continuous soil information to support international efforts in digital soil mapping (i.e., SoilGrids250m) as well as other potential applications detailed in the discussion section. We believe that the database with very dense point location measurements can help in conducting carbon stocks and sequestration studies. Such information can be used to help bridge the gap between ground data and remotely sensed datasets or data-derived products from modeling approaches intended to evaluate field-scale rates of agricultural carbon accumulation. The generated topsoil database in this study is archived and publicly available on the Zenodo open-access repository.

Keywords: agricultural land; soil total carbon; database; topographic derivatives; NDVI; digital soil mapping; Zenodo open-access repository

check for updates

Citation: Laamrani, A.; Voroney, P.R.; Saurette, D.D.; Berg, A.A.; Blackburn, L.; Gillespie, A.W.; Martin, R.C. An Extensive Field-Scale Dataset of Topsoil Organic Carbon Content Aimed to Assess Remote Sensed Datasets and Data-Derived Products from Modeling Approaches. *Remote Sens.* **2022**, *14*, 5519. https://doi.org/10.3390/rs14215519

Academic Editors: Ruxandra Vintila and Frank Veroustraete

Received: 24 August 2022
Accepted: 31 October 2022
Published: 2 November 2022

1. Introduction

The level of soil organic carbon (SOC) in agricultural soils is one of the key factors regulating soil quality as it directly benefits soil physical, chemical, and biological properties. It determines the ability of soils to perform ecosystem services such as maintaining plant productivity, enhancing water and air quality, and cycling of water and nutrients [1]. SOC also helps to set emission standards of greenhouse gases [2,3]. However, the geosciences suffer from a lack of large, georeferenced datasets that can be used to assess and monitor the role of SOC in plant growth, soil fertility, and organic carbon sequestration; therefore, the need is greater for accurate, quantified information at high spatial resolution over large areas [4]. For instance, the existing digital soil reference maps are based on legacy data that are at coarse grid cells and often outdated [4]. In addition, publicly available, large field georeferenced datasets are often limited in sample numbers and designed to serve other purposes [5–7].

At the field scale, SOC measurements obtained within agricultural fields can be related to a range of agricultural concerns including the quantification of how management practices influence SOC accumulation and loss (e.g., [8]), and to improve productivity and nutritional quality of food crops (e.g., [9]); however, many challenges can be attributed to the nature of the datasets. For instance, challenges can typically include: (i) a lack of consistent spatiotemporal structure that is needed with soil sampling to characterize the current state of SOC content and to track the evolution of surface soil systems [10–12]; (ii) a lack of the complexity of complementary datasets related to agricultural systems with a relatively high number of potential variables that may influence SOC, and thus many variables may have to be considered at the same time [11,13,14]; and (iii) the number of samples in soil datasets is often limited in both space and time and may not be labelled and aligned with ground-truthing samples [10,11,13].

To address this data gap, large databases of SOC measurements and associated publicly available topographic derivatives stored in a georeferenced format are needed, as well as field multi-resolution SOC datasets which are often unavailable at varying spatial and temporal resolutions. To this end, this study paper provides the first publicly accessible training set of a georeferenced SOC measurement database containing 840 sampling locations (dataset is stored at: https://doi.org/10.5281/zenodo.6611475, accessed on 30 October 2022 [15]) over a 26-hectare (ha) agricultural field located in southern Ontario, Canada. Since SOC is influenced by site topography (i.e., slope, aspect, and landscape position), each point of the database has been associated with a wide range of topographic derivatives.

2. Materials and Methods

2.1. Study Area Specifications and Description

The study site is located in the Lake Erie basin of southwestern (SW) Ontario, Canada, more specifically in Wellington County (Figure 1A,B). According to a Statistics Canada 2016 report [16], SW Ontario has more than 1.2 million ha of annually harvested crop land, thus representing the most intense agricultural production region of Canada. The field site is typical of conventionally cultivated farmland with corn (*Zea mays*), soybean (*Glycine max*), and winter wheat (*Triticum aestivum*) as the dominant crops grown in a 4-year crop rotation, sometimes with a cover crop. Cover crops promote sustainability of crop production and soil health, protect fields from soil erosion and promote SOC accumulation.

Field topography is characterized by a combination of irregular, moderately sloping terrains, interspersed with steep depressions (Figure 1C). Soil surface texture ranges from fine sandy loam to sandy loam (i.e., soil series include the Hillsburgh fine sandy loam, Caledon fine sandy loam and Fox sandy loam); [17]). This field is managed by a private agri-business farming company and agricultural management decisions are similar to those of farms in the nearby area. Crops receive chemical fertilizers as recommended and are visually monitored by farm operators and their agronomy team throughout the growing season. Crops are generally planted using no-till practices with soybean, while wheat crops are planted with air seed drills and corn is planted with a modern corn planter.

The climate is characterized by warm, humid summers, and long moderate winters (November–April) with a mean annual temperature of 6.7 °C and total annual precipitation of 946 mm. The Köppen–Geiger climate classification subtype for the study area's climate is Warm Summer Continental Climate (i.e., labeled as "Dfb" on the the Köppen–Geiger climate classification map; [18]). Almost one third of the total annual precipitation falls during the peak vegetative growth period (i.e., May–August). The coldest months are December, January, and February with a mean temperature of −3.1 °C; June, July, and August are the warmest months with a mean temperature of 18.8 °C (Environment Canada

2011 [19]; these data are from the Fergus Shand Dam weather station, which is located 5 km NW of the study site).

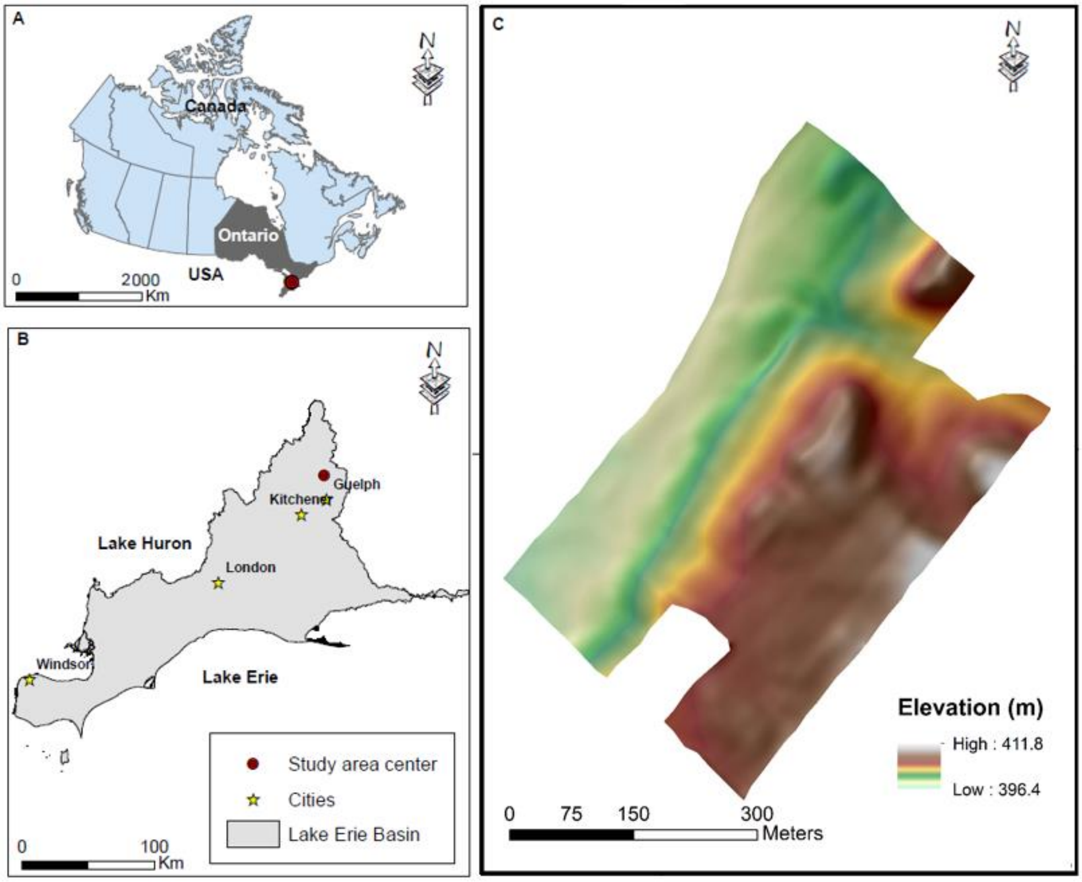

Figure 1. Maps showing the location of the study site in (**A**) southwestern Ontario within (**B**) the Lake Erie basin. (**C**) The study site with a topographic overview, where a hillshade raster was used beneath the elevation map to accentuate the topography.

2.2. Experimental Design, Sample Collection and Analysis

In spring 2018, an extensive field sampling campaign was conducted over the field to determine SOC levels which were subsequently linked to high resolution remote sensing imagery. As part of a larger project that dealt with the use of aerial sensors to assess SOC levels at the field-scale, our sampling design consisted of twenty-one parallel transects established across the field in a southwest to northeast direction. This spatial grid configuration provided a spatially continuous cross-sectional profile of topsoil SOC across the field. Georeferenced, systematic soil samples were acquired along the transects covering the field. A minimum of 20 m was maintained between transects, and soil samples were collected along each transect at 20-m intervals, producing a regular grid (for a total of 840 locations). At each location about 500 g of soil were collected from the topsoil horizon (0–15 cm) using a standard Dutch auger (Eijkelkamp, Giesbeek, The Netherlands). This agricultural soil was previously tilled, so the topsoil would have been well mixed prior to soil sampling. Field areas with irregular topography were sampled more intensely at various distances in areas of interest with soil accumulation (i.e., wet depressions). In addition to the grid data, samples along transects [20] were added to the database.

Soil samples were air-dried, passed through a 200-µm sieve, ground, and 300 mg were analyzed for total carbon (TC) via combustion at 1300 °C in a Leco CR-12 Carbon analyzer (Leco Corporation, St. Joseph, MI, USA; [21]). Each sample was analyzed twice and when the measurements agreed (±5% difference) with the measured carbon percent, they were averaged and used as the sample's datum (95% of the data; n = 801). When the two measurements disagreed (i.e., more than ±5% difference; 5% of the data, n = 39), a third measurement was initiated, and the estimate provided by the two measurements whose estimate was in closest agreement with the 5% pre-established rule was retained and averaged to represent the investigated sample. In total, 1719 samples (i.e., 801 × 2 replicates plus 39 × 3 replicates) were analyzed for total carbon (referred to henceforward as SOC as the inorganic carbon portion was negligeable in the investigated region [8]).

2.3. Topographic Derivatives Generation

In addition to SOC levels, our database contains a fine-resolution topographic dataset at 5 m spatial resolution that was upscaled from an 0.5 m initial dataset. This data can be used to analyze the relationships between local topographic variations and SOC changes. The topographic dataset was generated from a LiDAR-derived digital elevation model (DEM) that was created for the study site with the use of the publicly available LiDAR data collected by the Ontario Ministry of Agriculture, Food and Rural Affairs [22] and Ministry of Natural Resources and Forestry [23]. The LiDAR data was collected using a Leica Geosystems ADS100 sensor from an airborne platform over the studied field during 2017 and 2018 with a vertical accuracy of 0.09 m [23] and density of about 8 points/m^2. A bare earth DEM, referred to as a digital terrain model (DTM), was generated from the ground points of the classified LiDAR point cloud. The DTM was then used to create a set of 54 topographic derivatives (i.e., primary and secondary). Each topographic derivative was expressed as a raster and clipped to match the extent of the field. The 54 topographic variables are summarized in Table 1 in [24]. The DTM was first preprocessed using tools from WhiteboxTools [25,26] which were implemented in the Whitebox package [27] in the R coding environment (R Core Team, 2020; [28]). Terrain derivatives were then generated using the Whitebox [26,29] and RSAGA packages [30,31].

Positions of all field sampling measurements (i.e., along transects) were recorded using a Trimble GeoXT handheld GPS to provide 50 cm-level horizontal accuracy and to allow direct comparison with the DTM. The field SOC dataset was then superimposed upon the DTM and the 54 topographic derivatives to extract their values for each sampling location resulting in the final dataset of SOC measurements linked to topographic variables for the study site.

A list of environmental covariates generated from the LiDAR-derived DTM, their codes used in the database, reference to the algorithm used to derive each derivative, and the software used to compute them are provided with the database (i.e., Index sheet). This list has been adapted from our recent study on the same investigated field [24].

3. Results

3.1. Summary Statistics of the Topsoil SOC Measurements

The database combines 840 SOC measurements with their respective high precision GPS coordinates (database link: https://doi.org/10.5281/zenodo.6611475, accessed on 30 October 2022; [15]). This database has not been previously published, and it includes the SOC measurements for the georeferenced ground sampling locations and their corresponding topographic and derivative information (e.g., elevation, slope, aspect, TWI, etc.), as well as NDVI values. The SOC content values of the topsoil varied across the field with a mean ± standard deviation of 3.53 ± 1.02 %. These SOC values are ranging from a minimum of 1.91% to a maximum of 7.65%, and their distribution tended toward a slight positive skewing and a kurtosis of 3.04 based on the Jarque-Bera Normality Test.

3.2. SOC Mapping

To provide an overview of the spatial variability of SOC across the study site, kriging interpolation was applied to the dataset at a 20-m resolution (Figure 2). The kriging was conducted using ArcGIS and more details on how the advanced geostatistical kriging tools generate an estimated surface from a scattered set of points are provided on the ESRI website (i.e., [32,33]).

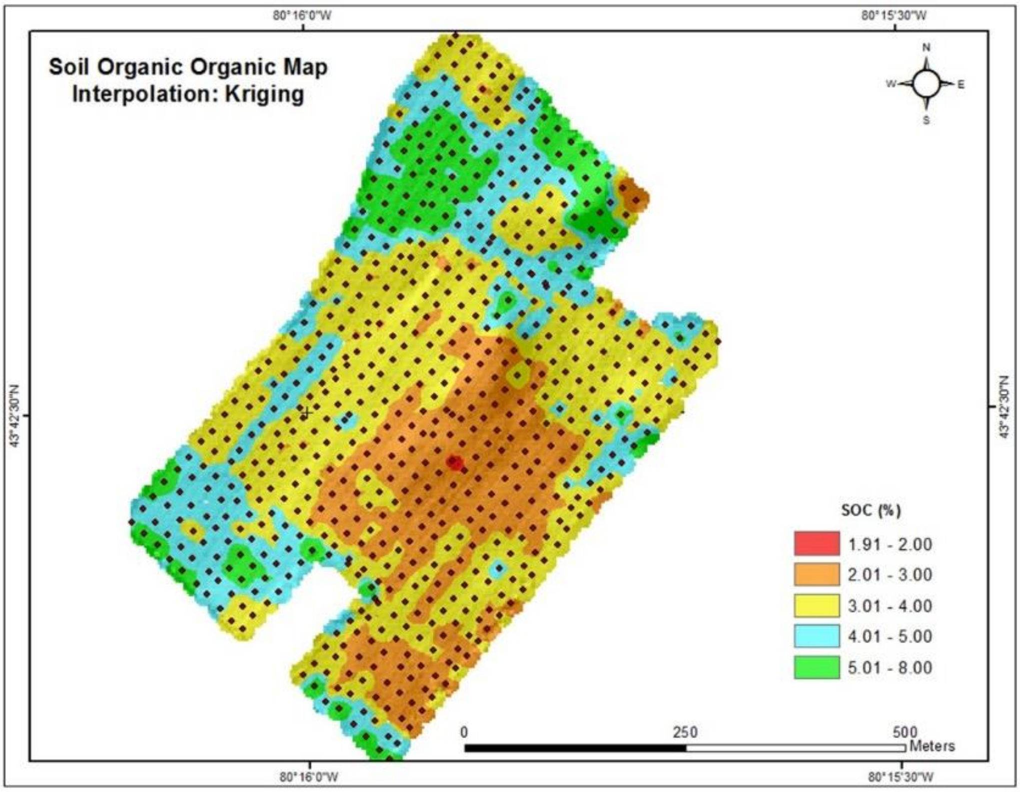

Figure 2. Map of the investigated site showing the spatial variability of SOC levels. The map was generated at 20-m resolution using a kriging interpolation.

4. Discussion of Potential Applications of the Created Database

To our knowledge, this study has provided the largest freely-available, field-scale database of SOC measurements in Canada. We address this spatial and temporal paucity of ground truthing by publishing the 840 SOC measurements combined with 54 topographical derivatives and NDVI data. It is provided in a georeferenced labelled form for comparison with the field SOC sampled locations. The spatially explicit, dense sampling, and continuous database achieved in this study is available to support national efforts in validating and calibrating SOC practices at the field-scale. For instance, the database can be useful for the following potential applications:

Frist, this study site is a subset of a large group of long-term study plots located in southern Ontario [24,34]. Our study provided a snapshot of the accumulated topsoil carbon with respect to current and past cropping and tillage systems. Therefore, our dataset should provide managers with a benchmark of SOC content together with site topographic indices. It could be used to monitor the effects of agricultural management practices through time with the original levels of agricultural productivity as well as SOC change, since organic carbon accumulation is a dynamic process that changes with time due to agricultural management practices and a changing climate [35].

Second, our sampling grid design (20 m × 20 m cells) provides a full coverage field distribution of SOC and topography that can be easily combined with lower resolution remote sensing imagery (i.e., Sentinel and Landsat with 20 m and 30 m resolutions, respectively). This is of great importance as data of such spatial resolution can assist with soil carbon modeling (i.e., the CENTURY model) simulation to be carried out for each cell/polygon in the field. In addition, this can help to correlate drivers of greenhouse gas emissions (i.e., topography, soil type, management practices) within a single cell. Indeed, a raster-based, spatially explicit modeling approach over our field which uses earth observation (EO) data (e.g., remote sensing products) can be a promising research tool [4]. Moreover, information on agricultural soils/crops and Enhanced Vegetation Indices could be derived from satellite imagery (e.g., Sentinel 2 with 20 m resolution) and inputted to the model to produce maps of SOC stocks on a raster basis [4].

Third, given the huge efforts required to build such a database, one can argue that remote sensing techniques using satellites can capture continuous soil cover data during the non-growing season (i.e., bare soil) over hundreds of square kilometers within minutes with regular, frequent revisit times (i.e., every 16 and 10 days) [36,37]. However, all remote sensing–based studies need calibration and validation as essential components, which can be done by collecting a limited number of ground measurements and relating them to remote sensing observations [38]. However, reported attempts to systematically determine the optimal number of ground measurements and the concomitant area sampled on the ground are limited [24]. Therefore, an essential challenge in using field data for calibration and validation in remote sensing–based studies for SOC mapping is to ensure that field measurements provide an appropriate and representative sample in support of mapping purposes [39]. Using an inappropriate number or type of measurement points could under- or overestimate the spatial variability and the accuracy of SOC estimates [40]. Therefore, one advantage of our large database will be to determine the number of soil measurements required to optimize the precision and accuracy of SOC estimations from various applications (i.e., SOC digital remotely sensed techniques using satellites). To this end, a recent study [24] has assessed the effects of sample numbers and covariate resolution on the prediction of SOC over our field using different machine learning algorithms, to identify the optimal sample numbers for various covariates. In another recent study [34], the grid-point generated on the investigated field was used to conduct a within-field yield prediction in cereal crops using LiDAR-derived topographic attributes. A detailed overview of how satellite imagery can be used to map topsoil organic carbon content over cultivated areas is provided in [41].

Finally, in addition to the terrain derivatives, we acquired Sentinel-2 data for the study area through Google Earth Engine (GEE) [42,43]. Our interest was in surface reflectance and indices during the growing season that could be linked to topsoil SOC measurements. As such, we queried Sentinel-2 imagery over a five-year period (2017–2021) during the growth season (i.e., from May 1 to September 1). For each image, we calculated the normalized difference vegetation index (NDVI; [44]); then summarized the NDVI data across the five-year period by calculating the median, maximum, and standard deviation of NDVI. In the context of the investigated field, we believe that NDVI can be useful to help differentiate bare soil from crop, detect periods where plants are under stress, crop stages, and potential yield and production.

5. Conclusions

This study represents a first attempt to develop a judicious training dataset for SOC and topographic information in Canada at the field scale. We used SOC measurements from topsoil samples (840 sampled locations) over a single 26 ha field to generate this large, open-access database together with topographic data. This large dataset can be used as a test case (i.e., for training and validation) for many applications that relate to agricultural SOC. The work presented here is a first step to overcoming some of the fundamental limitations of applying these techniques to mapping other soil attributes. It also provides valuable insights on the huge efforts required to build such a database. As topographic characteristics are often field-specific, the range of the topographic derivatives used in this study might limit the applicability to fields with similar topographic characteristics.

Author Contributions: A.L., P.R.V., A.A.B. and R.C.M. were involved in the study conceptualization; A.L. was involved in the preparation and writing the first draft of the manuscript. A.L., P.R.V., A.A.B., R.C.M., A.W.G., D.D.S. and L.B. were involved in reviewing the submitted version of the manuscript. R.C.M., P.R.V. and A.A.B. were involved in the developing of the proposal and funding

acquisition. A.L., P.R.V., L.B. and A.A.B. were involved in designing, running the field experiment, and conducting the laboratory carbon analyses. A.L., D.D.S. and L.B. prepared and conceived the georeferenced database. All authors have read and agreed to the published version of the manuscript.

Funding: Funding was provided by the Agricultural Greenhouse Gases Program-Agriculture and Agri-Food Canada (AGGP-015) and the Canada First Research Excellence Fund: Food from Thought program of the University of Guelph.

Data Availability Statement: The dataset supporting reported results are stored in the Zenodo general-purpose open repository developed under the European OpenAIRE program and operated by CERN. Here is the link to the dataset: https://doi.org/10.5281/zenodo.6611475, accessed on 30 October 2022.

Acknowledgments: We acknowledge the support from the many undergraduate and graduate students from the University of Guelph for assistance with field work. We thank all the graduate students who helped with the field data collection and preliminary analyses; and special thanks go to Mike March and Riley Eyre. Thanks to Alexander (Sandy) McLaren who helped with all the field logistic and remote sensing data collection during the project duration. The authors also thank the editor and three anonymous reviewers for their helpful comments which greatly improved an earlier version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Villarino, S.H.; Studdert, G.A.; Laterra, P. How does soil organic carbon mediate trade-offs between ecosystem services and agricultural production? *Ecol. Indic.* **2019**, *103*, 280–288. [CrossRef]
2. Lal, R. Soil carbon sequestration impacts on global climate change and food security. *Science* **2004**, *304*, 1623–1627. [CrossRef]
3. Laamrani, A.; Voroney, P.R.; Gillespie, A.W.; Chehbouni, A. Development of a Land Use Carbon Inventory for Agricultural Soils in the Canadian Province of Ontario. *Land* **2021**, *10*, 765. [CrossRef]
4. Tziolas, N.; Tsakiridis, N.; Chabrillat, S.; Demattê, J.A.M.; Ben-Dor, E.; Gholizadeh, A.; Zalidis, G.; van Wesemael, B. Earth Observation Data-Driven Cropland Soil Monitoring: A Review. *Remote Sens.* **2021**, *13*, 4439. [CrossRef]
5. Hengl, T.; de Jesus, J.M.; Heuvelink, G.B.; Gonzalez, M.R.; Kilibarda, M.; Blagotić, A.; Shangguan, W.; Wright, M.N.; Geng, X.; Bauer-Marschallinger, B.; et al. SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE* **2017**, *12*, e0169748. [CrossRef] [PubMed]
6. Georgiou, K.; Jackson, R.B.; Vindušková, O.; Abramoff, R.Z.; Ahlström, A.; Feng, W.; Harden, J.W.; Pellegrini, A.F.A.; Polley, H.W.; Soong, J.L.; et al. Global stocks and capacity of mineral-associated soil organic carbon. *Nat. Commun.* **2022**, *13*, 3797. [CrossRef] [PubMed]
7. Chen, S.; Arrouays, D.; Leatitia Mulder, V.; Poggio, L.; Minasny, B.; Roudier, P.; Libohova, Z.; Lagacherie, P.; Shi, Z.; Hannam, J.; et al. Digital Mapping of GlobalSoilMap Soil Properties at a Broad Scale: A Review. *Geoderma* **2022**, *409*, 115567. [CrossRef]
8. Laamrani, A.; Voroney, P.R.; Berg, A.A.; Gillespie, A.W.; March, M.; Deen, B.; Martin, R.C. Temporal Change of Soil Carbon on a Long-Term Experimental Site with Variable Crop Rotations and Tillage Systems. *Agronomy* **2020**, *10*, 840. [CrossRef]
9. Antil, R.S.; Narwal, R.P.; Singh, B.R. Raising Soil Organic Matter to Improve Productivity and Nutritional Quality of Food Crops in India. In *Soil Organic Matter and Feeding the Future: Environmental and Agronomic Impacts*, 1st ed.; Lal, R., Ed.; CRC Press: Boca Raton, FL, USA, 2021.
10. Kempen, B.; Brus, D.; Stoorvogel, J. Three-dimensional mapping of soil organic matter content using soil type–specific depth functions. *Geoderma* **2011**, *162*, 107–123. [CrossRef]
11. McBratney, A.B.; Stockmann, U.; Angers, D.A.; Minasny, B.; Field, D.J. Challenges for Soil Organic Carbon Research. In *Soil Carbon*; Springer: Cham, Switzerland, 2014; pp. 3–16.
12. Fathizad, H.; Taghizadeh-Mehrjardi, R.; Hakimzadeh Ardakani, M.A.; Zeraatpisheh, M.; Heung, B.; Scholten, T. Spatiotemporal Assessment of Soil Organic Carbon Change Using Machine-Learning in Arid Regions. *Agronomy* **2022**, *12*, 628. [CrossRef]
13. Piccini, C.; Marchetti, A.; Francaviglia, R. Estimation of Soil Organic Matter by Geostatistical Methods: Use of Auxiliary Information in Agriculturaland Environmental Assessment. *Ecol. Ind.* **2014**, *36*, 301–314. [CrossRef]
14. Angelopoulou, T.; Tziolas, N.; Balafoutis, A.; Zalidis, G.; Bochtis, D. Remote Sensing Techniques for Soil Organic Carbon Estimation: A Review. *Remote Sens.* **2019**, *11*, 676. [CrossRef]
15. Laamrani, A.; Voroney, P.; Saurette, D.D.; Berg, A.; Blackburn, L.; Gillespie, A.; Martin, R.C. Large dataset of soil organic carbon and topographic derivatives [Data set]. *Zenodo* **2022**. [CrossRef]
16. Statistics Canada. Census of Agriculture. 2016. Available online: https://www150.statcan.gc.ca/n1/daily-quotidien/170510/dq170510a-eng.htm (accessed on 22 August 2022).
17. Hoffman, D.W.; Matthews, B.C.; Wicklund, R.E. *Soil Associations of Southern Ontario*; Report No. 30 of the Ontario Soil Survey; Canada Department of Agriculture: Ottawa, ON, Canada, 1964. Available online: https://sis.agr.gc.ca/cansis/publications/surveys/on/on30/on30_report.pdf (accessed on 22 August 2022).

18. Beck, H.E.; Zimmermann, N.E.; McVicar, T.R.; Vergopolan, N.; Berg, A.; Wood, E.F. Present and future köppen-geiger climate classification maps at 1-km resolution. *Sci. Data* **2018**, *5*, 180214. [CrossRef]
19. Environment Canada. Canadian Climate Normals 1981–2010: Fergus Shand Dam Weather Station. 2011. Available online: http://climate.weather.gc.ca/climate_normals/index_e.html (accessed on 22 August 2022).
20. Laamrani, A.; Berg, A.A.; Voroney, P.; Feilhauer, H.; Blackburn, L.; March, M.; Dao, P.D.; He, Y.; Martin, R.C. Ensemble Identification of Spectral Bands Related to Soil Organic Carbon Levels over an Agricultural Field in Southern Ontario, Canada. *Remote Sens.* **2019**, *11*, 1298. [CrossRef]
21. Wang, D.; Anderson, D.W. Direct measurement of organic carbon content in soils by the Leco CR-12 carbon analyzer. *Commun. Soil Sci Plan.* **1998**, *29*, 15–21. [CrossRef]
22. Ontario Ministry of Agriculture, Food and Rural Affairs. Ontario Classified Point Cloud (Lidar-Derived). 2019. Available online: https://geohub.lio.gov.on.ca/datasets/mnrf::ontario-classified-point-cloud-lidar-derived (accessed on 22 August 2022).
23. Ontario Ministry of Natural Resources and Forestry. Ontario Elevation Mapping Program. 2019. Available online: https://geohub.lio.gov.on.ca/pages/ontario-elevation-mapping-program (accessed on 22 August 2022).
24. Saurette, D.; Berg, A.; Laamrani, A.; Heck, R.J.; Gillespie, A.W.; Voroney, P. Effects of sample size and covariate resolution on field-scale predictive digital mapping of soil carbon. *Geoderma* **2022**, *425*, 116054. [CrossRef]
25. Lindsay, J.B. Whitebox GAT: A case study in geomorphometric analysis. *Comput. Geosci.* **2016**, *95*, 75–84. [CrossRef]
26. Newman, D.R.; Lindsay, J.B.; Cockburn, J.M.H. Evaluating metrics of local topographic position for multiscale geomorphometric analysis. *Geomorphology* **2018**, *312*, 40–50. [CrossRef]
27. Wu, Q. Whitebox: "WhiteboxTools" R Frontend. 2019. Available online: https://R-Forge.R-project.org/projects/whitebox (accessed on 22 August 2022).
28. Team, R. *Core. R: A Language and Environment for Statistical Computing*; R foundation for Statistical Computing: Vienna, Austria, 2013.
29. Lindsay, J. WhiteboxTools User Manual. User Manual, University of Guelph. 2018. Available online: https://www.uoguelph.ca/~{}hydrogeo/WhiteboxTools/index.html (accessed on 22 August 2022).
30. Conrad, O.; Bechtel, B.; Bock, M.; Dietrich, H.; Fischer, E.; Gerlitz, L.; Wehberg, J.; Wichmann, V.; Böhner, J. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci. Model Dev. Discuss.* **2015**, *8*, 2271–2312. [CrossRef]
31. Brenning, A.; Bangs, D.; Becker, M. RSAGA: SAGA geoprocessing and terrain analysis. R package version 1.3.0. 2018. Available online: https://CRAN.R-project.org/package=RSAGA (accessed on 22 August 2022).
32. Esri—ArcGIS Pro 2.9. How Kriging works? Available online: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/how-kriging-works.htm#GUID-08AA4C59-A05E-4F9F-A18D-E30B36C7523A (accessed on 22 August 2022).
33. Oliver, M.A. Kriging: A Method of Interpolation for Geographical Information Systems. *Int. J. Geogr. Inf. Syst.* **1990**, *4*, 313–332. [CrossRef]
34. Eyre, R.; Lindsay, J.; Laamrani, A.; Berg, A. Within-Field Yield Prediction in Cereal Crops Using LiDAR-Derived Topographic Attributes with Geographically Weighted Regression Models. *Remote Sens.* **2021**, *13*, 4152. [CrossRef]
35. Mandal, A.; Majumder, A.; Dhaliwal, S.S.; Toor, A.S.; Mani, P.K.; Naresh, R.K.; Gupta, R.K.; Mitran, T. Impact of Agricultural Management Practices on Soil Carbon Sequestration and Its Monitoring through Simulation Models and Remote Sensing Techniques: A Review. *Crit. Rev. Environ. Sci. Technol.* **2020**, *52*, 1–49. [CrossRef]
36. Laamrani, A.; Joosse, P.; McNairn, H.; Berg, A.A.; Hagerman, J.; Powell, K.; Berry, M. Assessing Soil Cover Levels during the Non-Growing Season Using Multitemporal Satellite Imagery and Spectral Unmixing Techniques. *Remote Sens.* **2020**, *12*, 1397. [CrossRef]
37. Mzid, N.; Castaldi, F.; Tolomio, M.; Pascucci, S.; Casa, R.; Pignatti, S. Evaluation of Agricultural Bare Soil Properties Retrieval from Landsat 8, Sentinel-2 and PRISMA Satellite Data. *Remote Sens.* **2022**, *14*, 714. [CrossRef]
38. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46. [CrossRef]
39. Somarathna, P.; Minasny, B.; Malone, B.P. More data or a better model? Figuring out what matters most for the spatial prediction of soil carbon. *Soil Sci. Soc. Am. J.* **2017**, *81*, 1413–1426. [CrossRef]
40. Khaledian, Y.; Miller, B.A. Selecting appropriate machine learning methods for digital soil mapping. *Appl. Math. Model.* **2019**, *81*, 401–418. [CrossRef]
41. Vaudour, E.; Gholizadeh, A.; Castaldi, F.; Saberioon, M.; Borůvka, L.; Urbina-Salazar, D.; Fouad, Y.; Arrouays, D.; Richer-de-Forges, A.C.; Biney, J.; et al. Satellite Imagery to Map Topsoil Organic Carbon Content over Cultivated Areas: An Overview. *Remote Sens.* **2022**, *14*, 2917. [CrossRef]
42. Zhao, Q.; Yu, L.; Li, X.; Peng, D.; Zhang, Y.; Gong, P. Progress and Trends in the Application of Google Earth and Google Earth Engine. *Remote Sens.* **2021**, *13*, 3778. [CrossRef]
43. Yu, L.; Gong, P. Google Earth as a virtual globe tool for Earth science applications at the global scale: Progress and perspectives. *Int. J. Remote Sens.* **2012**, *33*, 3966–3986. [CrossRef]
44. Rouse, J.W.; Haas, R.W.; Schell, J.A.; Deering, D.W.; Harlan, J.C. *Monitoring vegetation systems in the Great Plains with ERTS. Third Earth Resources Technology Satellite–1 Syposium*; Freden, S.C., Mercanti, E.P., Becker, M., Eds.; Technical Presentations, NASA SP-351; NASA: Washington, DC, USA, 1974; Volume I, pp. 309–317.