Estimate Soil Organic Carbon from Remote Sensing

Subjects: Remote Sensing

Contributor: Marko Pavlovic , Slobodan Ilic , Neobojša Ralevic , Nenad Antonic , Dylan Warren Raffa , Michele Bandecchi , Dubravko Culibrk

Monitoring soil organic carbon (SOC) typically assumes conducting a labor-intensive soil sampling campaign, followed by laboratory testing, which is both expensive and impractical for generating useful, spatially continuous data products.



1. Introduction

Soil represents a complex mixture of organic and inorganic constituents with different physical and chemical properties, which vary significantly between locations and even within a single field ^[1]. It is a key component of terrestrial ecosystems, as it facilitates the circulation of energy and materials between the atmosphere and the biosphere ^[2].

Soil health can be defined as the ability of the soil to function effectively as a component in a thriving ecosystem ^[3]. In order to ensure effective monitoring and enable adequate assessment of the condition of the soil, one needs to select appropriate indicators of its condition. The indicators should meet certain criteria: they should be accepted by experts as valid; their measurement should be carried out routinely and on a large scale; they need to be understood and accepted by the general population in order to achieve a global impact ^[4].

Soil organic carbon (SOC) content is a widely accepted indicator of soil quality, as SOC plays a central role in various soil functions ^[5]. SOC measurement is a common component of soil property analysis. Furthermore, carbon as an element is well known and recognized by the global population ^[6]. All of this makes SOC a valuable indicator for assessing and monitoring changes in soil health.

The amount and quality of SOC are closely related to key soil functions, including nutrient mineralization, aggregate stability, air and water permeability, water retention, and flood control ability ^[5]. These soil functions are, in turn, related to a wide range of ecosystem attributes. For example, high SOC levels in mineral soils tend to correlate with high plant productivity, which has a positive effect on wildlife habitat, distribution, and population size

^[Z]. Through the protection and increase of stored SOC, one can protect or increase soil fertility, reduce soil erosion, and reduce habitat conversions ^[8].

In addition to its importance for the soil, SOC has the potential to help neutralize the negative effects of increasing concentrations of CO_2 in the atmosphere (which significantly contribute to global warming and climate change ^[9]) and help ensure food security around the wold ^[10].

While SOC plays a key role in mitigating climate change by acting as a carbon sink, the historical loss of carbon from this pool ^[11] has been significant, and the potential for future accelerated loss under warming scenarios is a serious threat ^{[12][13]}.

As a natural solution to fight climate change, strategies that involve conserving existing SOC stocks (avoidance of losses) and replenishing stocks in carbon-depleted soils ^[14] can be used as a means of achieving the United Nations Sustainable Development Goals (UNSDG), the goals of the United Nations Framework Convention on Climate Change (UNFCCC), and the United Nations Convention on Combating Desertification (UNCCD) ^[15].

Despite the scientific consensus about the potential and myriad benefits that can be brought about by the development and application of soil organic carbon storage and sequestration techniques, they remain limited in practice. A fundamental issue affecting the adoption of such methodologies is the lack of accurate and cost-effective ways of measuring SOC content in the top layer of the soil (as this is most affected by land use, agricultural practices, etc.).

When it comes to measuring global SOC stocks, many estimates have been published over the past decades, and most studies report a global SOC estimate of approximately 1500 Pg of carbon (Pg C), but there is considerable variation among estimates (ranging from 504 to 3000 Pg C) ^[16].

The large variation in the estimates of global SOC stocks arises from differences in the sampling period, the intensity and spatial resolution of soil profile databases, as well as from differences in approaches to calculating the estimates themselves ^[17]. The uneven distribution of georeferenced soil profiles around the world is another reason for such a large variation in the estimates ^[18]. In addition, there is no consensus when it comes to including inorganic carbon, different levels of rock content ^[19], and the effects of natural or anthropogenic phenomena (such as flooding, erosion, fire, soil fertilization, and plowing ^[20]) in carbon stock assessments.

2. Monitoring Soil Organic Carbon Based on Remote Sensing

recent years, remote sensing has emerged as a particularly effective method for tracking agricultural and environmental changes ^{[21][22][23][24]}. The technology relies on diverse sensors and platforms, such as satellite constellations and Unmanned Aerial Systems (UAS) to gather data, which are then typically processed using advanced algorithms, often in the realm of machine learning (ML) and deep learning (DL) ^[25].

Deep learning represents a specialized subset of machine learning that excels at learning from large, unstructured datasets using complex, layered neural networks. While traditional ML algorithms work well with smaller, structured datasets and often require manual feature selection, DL algorithms automatically extract features and patterns, especially from data like images and speech. This makes deep learning more powerful for certain applications, but it requires more computational resources and is often less interpretable than conventional machine learning techniques.

The ongoing advancements in remote sensing represent a promising alternative to traditional SOC monitoring. Toth and Jóžków provide a fairly recent review of different remote sensing platforms and sensors available today ^[26].

In the research presented here, the focus is on inferring SOC content from satellite data only. Most studies focused on determining SOC, however, rely on data (spectrograms) collected from hand-held sensors. While the accuracy achieved in this way is typically higher than using satellite imagery, such approaches can hardly be scaled to enable continuous monitoring of carbon stocks on a global level.

Gomez et al. ^[27] presented an early, albeit limited study (based on just 146 soil samples), which compared the results that can be achieved applying ML methods to in-the-field Vis–NIR measurements vs. applying them to hyperspectral satellite imagery. The images were obtained from the Hyperion sensor on the EO-1 satellite, which is, unfortunately, no longer functional, and there is no longer an active hyperspectral satellite that captures images in the VNIR–SWIR region, making it hard to replicate their work. In addition to trying to model the whole dataset used in the study, the authors tried focusing on specific land cover classes (cropping soils, pasture soils) and opted for a partial least-squares regression as their SOC predictor. Gomez et al. observed that the SOC in their cropping soils ranged between 0.54% and 1% and was lower than in the pastures, where SOC was in the 1.08% to 5.1% range. They evaluated their methodology based on R² and the Root-Mean-Squared Error (RMSE). The models based on satellite imagery did not perform well for cropping soils (R² of 0.04 and RMSE of 0.11) and lagged significantly behind the hand-held-sensor-based models in terms of R² (R² of 0.16 and RMSE of 0.1). However, when evaluated on pastures and the whole dataset, the two approaches achieved comparable and much better performance. The approach based solely on satellite data at their native resolution achieved an R² of 0.51, but the RMSE was quite high (0.73% SOC). Thus, the study showed that land cover is very important, when it comes to modeling and estimating SOC remotely.

More recently, Wang et al. ^[28] tried to use ML techniques to estimate SOC stock in the semi-arid rangelands of eastern Australia through the application of different machine learning techniques, with a focus on evaluating the impact of considering seasonal fractional cover on model performance. These features were used to extend other hand-crafted features derived from satellite imagery, as well as other remotely sensed climate features such as rainfall and temperature and data about lithology. They trained and evaluated their models using a limited amount of soil samples (705). They used random forests (RF) ^[29], Boosted Regression Trees (BRT) ^[30], and support vector machines (SVM) ^[31] to model their data. The RF approach performed the best and achieved an R² of 0.47 on their dataset.

Several studies tried to evaluate the effectiveness of hyperspectral data obtained from airborne sensors and extended their findings to evaluate the expected performance of sensors expected to be deployed in the future ^[32] ^[33]. While researchers focus on multispectral data in the study presented here, it is worth noting that, albeit relying on a very limited set of soil samples (81) obtained for a 7 km² area in Luxembourg, 40% of which were used as a test set, Steinberg et al. achieved a relatively high R² (0.74) and an RMSE of 0.22% for SOC using autoPSLR applied to hyperspectral data from an airborne sensor ^[33]. Once sufficient hyperspectral data are available, the methodology researchers propose can easily be adapted to that domain, leading to even better performance.

Over the last decade, deep learning has revolutionized the area of machine learning and artificial intelligence and has become the dominant paradigm in the domain. The crucial advance over previously used methods is that the approach relies on end-to-end learning, which allows the ML models to learn the features on which to make their decisions and estimated directly from the raw input data, instead of relying on human-engineered features ^[34].

Yuan et al. provided an overview of the applications of both classical neural networks and DL models to the monitoring of environmental parameters using remote sensing data ^[35]. They showed that DL outperformed traditional ML models and has led to significant improvements in many applications, including land cover mapping, vegetation parameter, soil moisture, evapotranspiration, agricultural yield prediction, etc. The authors correctly highlighted the limitation of the DL approaches, which is related to the relatively limited amounts of training data available, as well as the potential to apply transfer learning to circumvent this problem. They mentioned two types of transfer learning: region-based and data-based. The first relates to pretraining on a geographical region for which ample data are available and adjusting the model to a different region with limited data available. In the ML community, this is usually referred to as fine-tuning. The latter is more in line with what the meaning of transfer learning is in the ML domain and relates to transferring the models trained on data obtained from a sensor or a group of sensors to other sensors. In the study presented here, we use a third kind of transfer learning, common in the computer vision community ^[36], where the initial model is trained on the same type of input data (Sentinel-2), but for a different visual task (land cover classification), and is used as a feature extractor for the final model (which performs SOC estimation in researchers' case).

While the first application that Yuan et al. discussed was land cover, no approaches to estimating SOC were mentioned in this study. In addition, while approaches based on different DNN architectures were discussed (most relying on convolutional neural networks), none were identified in the study that use the U-Net model.

Rakhlin et al., however, successfully applied U-Net with Lovász softmax loss for land cover classification using RGB data made available as part of the DeepGlobe Challenge ^[37].

Yang et al. used a CNN to try to infer SOC for a central location based on input data that covered the surrounding region ^[38]. The input of their model was environmental variables combined with MODIS MCD12Q2 phenology variables. They trained and evaluated their approach on a limited set of 733 samples, collected in Anhui Province of China. This limited the complexity of the CNN they could use, since no transfer learning was used in the study, but the CNN fared better than a random forest model, achieving a modest R² of 0.26.

Emadi et al. ^[39] focused on Northern Iran and used a large number of input features (105). Most were humancrafted indices extracted from Landsat-8 and MODIS satellite imagery, but their input also included topologyrelated parameters, such as curvature, slope, etc. Using a dataset of 1879 composite soil samples and relying on 10-fold cross-validation, they compared the performance of several traditional ML algorithms (support vector machines, multi-layer-perceptron, regression decision trees, random forests, and extreme gradient boosting) with a DL model when predicting SOC. The DL model that showed the best results in the study was a fairly simple fully connected neural net, with seven hidden layers and 50 neurons in each of them, but it still outperformed the other methods tested. The authors reported a comparatively large R² value of 0.65, with an RMSE of 0.75% SOC.

In a recent study, Castaldi et al. ^[40] evaluated the capability of Sentinel-2 time series to estimate soil organic carbon and clay content at local scale in croplands. The pipeline they proposed relies heavily on human engineering, both in terms of the features they derived from Sentinel-2 imagery (NDVI, NBR2, BSI, S2WI), as well as in terms of how they were used to create the input to their machine learning models. In terms of modeling, they did not opt for deep neural networks, but the Quantile Regression Forest (QRF) algorithm, QRF with added longitude and latitude as covariates, and a hybrid approach, the Linear Mixed-Effect Model (LMEM), which included the spatial autocorrelation of the soil properties. While the latter takes spatial information into account up to a point, their approach is essentially pixel based, which differs from the one proposed here. In addition, the authors of the research aimed to assess the capability of their approach in a very limited scenario, by creating and evaluating models for each of their test sites separately. No attempt was made to create a single model that could be applied globally, or at least for a large part of the Earth's surface. Thus, the results they achieved could be viewed as a sort of "blue-sky-performance", which could be reached by a global model using Sentinel-2 images as the input. The R² of the best of Castaldi et al.'s models ranged from 0.09 to 0.22 and was 0.152 on average.

References

- Jandl, R.; Rodeghiero, M.; Martinez, C.; Cotrufo, M.F.; Bampa, F.; Van Wesemael, B.; Harrison, R.B.; Guerrini, I.A.; Richter, D.d., Jr.; Rustad, L.; et al. Current status, uncertainty and future needs in soil organic carbon monitoring. Sci. Total Environ. 2014, 468, 376–383.
- Zhang, L.; Liu, Y.; Li, X.; Huang, L.; Yu, D.; Shi, X.; Chen, H.; Xing, S. Effects of soil map scales on simulating soil organic carbon changes of upland soils in Eastern China. Geoderma 2018, 312, 159–169.
- Schoenholtz, S.H.; Van Miegroet, H.; Burger, J. A review of chemical and physical properties as indicators of forest soil quality: Challenges and opportunities. For. Ecol. Manag. 2000, 138, 335– 356.
- 4. Stockmann, U.; Padarian, J.; McBratney, A.; Minasny, B.; de Brogniez, D.; Montanarella, L.; Hong, S.Y.; Rawlins, B.G.; Field, D.J. Global soil organic carbon assessment. Glob. Food Secur. 2015,

6, 9–16.

- Berryman, E.; Hatten, J.; Page-Dumroese, D.S.; Heckman, K.A.; D'Amore, D.V.; Puttere, J.; SanClements, M.; Connolly, S.J.; Perry, C.H.H.; Domke, G.M. Soil carbon. In Forest and Rangeland Soils of the United States under Changing Conditions; Springer: Cham, Switzerland, 2020; pp. 9–31.
- 6. Koch, A.; McBratney, A.; Adams, M.; Field, D.; Hill, R.; Crawford, J.; Minasny, B.; Lal, R.; Abbott, L.; O'Donnell, A.; et al. Soil security: Solving the global soil crisis. Glob. Policy 2013, 4, 434–441.
- 7. Oldfield, E.E.; Wood, S.A.; Bradford, M.A. Direct effects of soil organic matter on productivity mirror those observed with organic amendments. Plant Soil 2018, 423, 363–373.
- Bossio, D.; Cook-Patton, S.; Ellis, P.; Fargione, J.; Sanderman, J.; Smith, P.; Wood, S.; Zomer, R.; Von Unger, M.; Emmer, I.; et al. The role of soil carbon in natural climate solutions. Nat. Sustain. 2020, 3, 391–398.
- 9. Minasny, B.; Malone, B.P.; McBratney, A.B.; Angers, D.A.; Arrouays, D.; Chambers, A.; Chaplot, V.; Chen, Z.S.; Cheng, K.; Das, B.S.; et al. Soil carbon 4 per mille. Geoderma 2017, 292, 59–86.
- 10. Andrews, S.S.; Karlen, D.L.; Cambardella, C.A. The soil management assessment framework: A quantitative soil quality evaluation method. Soil Sci. Soc. Am. J. 2004, 68, 1945–1962.
- 11. Sanderman, J.; Hengl, T.; Fiske, G.J. Soil carbon debt of 12,000 years of human land use. Proc. Natl. Acad. Sci. USA 2017, 114, 9575–9580.
- 12. Jenkinson, D.; Adams, D.; Wild, A. Model estimates of CO2 emissions from soil in response to global warming. Nature 1991, 351, 304–306.
- 13. Hicks Pries, C.E.; Castanha, C.; Porras, R.; Torn, M. The whole-soil carbon flux in response to warming. Science 2017, 355, 1420–1423.
- Smith, P.; Martino, D.; Cai, Z.; Gwary, D.; Janzen, H.; Kumar, P.; McCarl, B.; Ogle, S.; O'Mara, F.; Rice, C.; et al. Greenhouse gas mitigation in agriculture. Philos. Trans. R Soc. Biol. Sci. 2008, 363, 789–813.
- Smith, P.; Adams, J.; Beerling, D.J.; Beringer, T.; Calvin, K.V.; Fuss, S.; Griscom, B.; Hagemann, N.; Kammann, C.; Kraxner, F.; et al. Land-management options for greenhouse gas removal and their impacts on ecosystem services and the sustainable development goals. Annu. Rev. Environ. Resour. 2019, 44, 255–286.
- Todd-Brown, K.E.; Randerson, J.T.; Post, W.M.; Hoffman, F.M.; Tarnocai, C.; Schuur, E.A.; Allison, S.D. Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. Biogeosciences 2013, 10, 1717–1736.
- 17. Scharlemann, J.P.; Tanner, E.V.; Hiederer, R.; Kapos, V. Global soil carbon: Understanding and managing the largest terrestrial carbon pool. Carbon Manag. 2014, 5, 81–91.

- 18. Batjes, N. Harmonized soil profile data for applications at global and continental scales: Updates to the WISE database. Soil Use Manag. 2009, 25, 124–127.
- 19. Batjes, N.H. Total carbon and nitrogen in the soils of the world. Eur. J. Soil Sci. 1996, 47, 151– 163.
- 20. Meentemeyer, V.; Gardner, J.; Box, E.O. World patterns and amounts of detrital soil carbon. Earth Surf. Process. Landforms 1985, 10, 557–567.
- 21. Jiménez-Lao, R.; Aguilar, F.J.; Nemmaoui, A.; Aguilar, M.A. Remote Sensing of Agricultural Greenhouses and Plastic-Mulched Farmland: An Analysis of Worldwide Research. Remote Sens. 2020, 12, 2649.
- 22. Li, J.; Pei, Y.; Zhao, S.; Xiao, R.; Sang, X.; Zhang, C. A review of remote sensing for environmental monitoring in China. Remote Sens. 2020, 12, 1130.
- 23. Overpeck, J.T.; Meehl, G.A.; Bony, S.; Easterling, D.R. Climate data challenges in the 21st century. Science 2011, 331, 700–702.
- 24. Pavlovic, M.; Ilic, S.; Antonic, N.; Culibrk, D. Monitoring the Impact of Large Transport Infrastructure on Land Use and Environment Using Deep Learning and Satellite Imagery. Remote Sens. 2022, 14, 2494.
- 25. Kattenborn, T.; Leitloff, J.; Schiefer, F.; Hinz, S. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS J. Photogramm. Remote Sens. 2021, 173, 24–49.
- 26. Toth, C.; Jóźków, G. Remote sensing platforms and sensors: A survey. ISPRS J. Photogramm. Remote Sens. 2016, 115, 22–36.
- 27. Gomez, C.; Rossel, R.A.V.; McBratney, A.B. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. Geoderma 2008, 146, 403–411.
- 28. Wang, B.; Waters, C.; Orgill, S.; Gray, J.; Cowie, A.; Clark, A.; Li Liu, D. High resolution mapping of soil organic carbon stocks using remote sensing variables in the semi-arid rangelands of eastern Australia. Sci. Total Environ. 2018, 630, 367–378.
- 29. Breiman, L. Random forests. Mach. Learn. 2001, 45, 5–32.
- 30. Elith, J.; Leathwick, J.R.; Hastie, T. A working guide to boosted regression trees. J. Anim. Ecol. 2008, 77, 802–813.
- 31. Cortes, C.; Vapnik, V. Support-vector networks. Mach. Learn. 1995, 20, 273–297.
- Castaldi, F.; Palombo, A.; Santini, F.; Pascucci, S.; Pignatti, S.; Casa, R. Evaluation of the potential of the current and forthcoming multispectral and hyperspectral imagers to estimate soil texture and organic carbon. Remote Sens. Environ. 2016, 179, 54–65.

- 33. Steinberg, A.; Chabrillat, S.; Stevens, A.; Segl, K.; Foerster, S. Prediction of common surface soil properties based on Vis-NIR airborne and simulated EnMAP imaging spectroscopy data: Prediction accuracy and influence of spatial resolution. Remote Sens. 2016, 8, 613.
- 34. Goodfellow, I.; Bengio, Y.; Courville, A. Deep Learning; MIT Press: Cambridge, MA, USA, 2016.
- 35. Yuan, Q.; Shen, H.; Li, T.; Li, Z.; Li, S.; Jiang, Y.; Xu, H.; Tan, W.; Yang, Q.; Wang, J.; et al. Deep learning in environmental remote sensing: Achievements and challenges. Remote Sens. Environ. 2020, 241, 111716.
- Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
- Rakhlin, A.; Davydow, A.; Nikolenko, S. Land cover classification from satellite imagery with U-Net and lovász-softmax loss. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–23 June 2018; pp. 262–266.
- Yang, L.; Cai, Y.; Zhang, L.; Guo, M.; Li, A.; Zhou, C. A deep learning method to predict soil organic carbon content at a regional scale using satellite-based phenology variables. Int. J. Appl. Earth Obs. Geoinf. 2021, 102, 102428.
- 39. Emadi, M.; Taghizadeh-Mehrjardi, R.; Cherati, A.; Danesh, M.; Mosavi, A.; Scholten, T. Predicting and mapping of soil organic carbon using machine learning algorithms in Northern Iran. Remote Sens. 2020, 12, 2234.
- 40. Castaldi, F.; Koparan, M.H.; Wetterlind, J.; Žydelis, R.; Vinci, I.; Savaş, A.Ö.; Kıvrak, C.; Tunçay, T.; Volungevičius, J.; Obber, S.; et al. Assessing the capability of Sentinel-2 time-series to estimate soil organic carbon and clay content at local scale in croplands. ISPRS J. Photogramm. Remote Sens. 2023, 199, 40–60.

Retrieved from https://www.encyclopedia.pub/entry/history/show/125103