Geoderma 262 (2016) 187–198

Contents lists available at ScienceDirect

Geoderma

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

Mapping soil carbon stocks across Scotland using a neural network model

M.J. Aitkenhead *, M.C. Coull

The James Hutton Institute, Craigiebuckler, Aberdeen, AB15 8QH Scotland, UK

ARTICLE INFO

Article history:
Received 27 April 2015
Received in revised form 21 August 2015
Accepted 23 August 2015
Available online 4 September 2015

Keywords:
Soil
Mapping
Carbon
Scotland
Neural network
Digital Soil Mapping

ABSTRACT

A neural network model was trained to predict soil organic matter content, bulk density and soil organic matter density at different soil profile depths across Scotland. These predictions were then used to predict soil organic carbon content. The data used to train the model was developed from the National Soil Inventory of Scotland (NSIS) datasets, along with spatial datasets for topographic and climatic variables, and for geology, soil type and land cover. The trained network was tested and found to explain 79.8% of the variance in organic matter content, 77.9% of the variance in bulk density and 57.3% of the variance in profile depth. Various statistical measures were used to evaluate the predictive ability of the model, showing that it was suitable for predicting the carbon stocks of soils. The neural network model was used to make predictions from the surface to 1 m in 1 cm intervals, at 100 m spatial resolution, across Scotland. This allowed us to make a prediction of the distribution, spatially and at depth, of carbon stocks to 1 m across Scotland and to make estimates of the total carbon stock of Scottish soils (2954 Tg) and the amount stored in different soil types across the country. We found that our estimate of the amount of carbon stored in Scottish soils was in agreement with previous estimates. Mineral and organo-mineral soils are predicted to hold a large amount of carbon in the upper portion, and in terms of carbon stock are almost as important as peat soils. At increased depth, a much smaller proportion of the total Scottish soil carbon stock is held in soils not classified as organic. We provide information about the distribution of carbon stocks with depth and soil type and under different land use/land cover types. Finally, we discuss the relevance of this information in relation to efforts to store carbon within Scottish soils in the medium to long term.

1. Introduction

Mapping of soil carbon stocks across the UK is necessary for a number of reasons, including providing baseline data for monitoring the effects of climate change, carbon accounting, and informing land management decisions. Actual or potential variation in soil organic carbon stocks is now a factor considered in climate change negotiations, implying a need to accurately map the spatial distribution of soil organic carbon. Examples of monitoring frameworks incorporating soil carbon stocks include France (Martin et al., 2011), Ireland (Xu et al., 2011) and China (Wiesmeier et al., 2011; Liu et al., 2012). For each of these examples, the spatial distribution of soil organic carbon (SOC) stocks was modelled as a function of environmental and soil variables.

In addition to providing carbon stock assessment for political, regulatory and conservation purposes, any map of soil carbon that includes information about variation with depth will be useful for providing baseline data and parameterisation for soil profile models. These models could then be used to make more accurate predictions about the effects of land management and climate change over time. Hong et al. (2010) for example, produced maps of various soil water characteristics for South Korea, based on soil carbon profile predictions. Recent work on modelling soil carbon dynamics, such as that of Nieto et al. (2010) using RothC, Aitkenhead et al. (2011) using MOSES, or Oelbermann and Voroney (2011) using Century, is reliant on direct measurements of the distribution of soil carbon in the profile.

Several different approaches have been used for national-scale soil carbon stock mapping, including interpolation between sample points (Bradley et al., 2005), characterisation by map unit (Batjes, 2010); estimation from point data (Chapman et al., 2013) and raster-based soil process modelling (Smith et al., 2007; Smith et al., 2010a,b). In regions where the landscape is heterogeneous, it can be difficult to make accurate predictions of soil carbon stocks without (A) information about the relationships between landscape character and soil carbon content, and (B) high-resolution spatial datasets of the relevant environmental parameters, such as vegetation and topography. Development of models of soil organic carbon for heterogeneous landscapes requires sufficiently detailed field survey data. Examples of such models include the use of pedotransfer functions (Han et al., 2010) and kriging using existing survey point data (Dlugoss et al., 2010; Ungar et al., 2010). The work by Poggio and Gimona (2014) on modelling soil organic carbon stocks in Scotland integrated several approaches, including General Additive Modelling (GAM), and kriging of residuals to account for local variation. This approach is highly flexible, being able to incorporate a wide variety of spatial covariates, and also provides a method for determining levels of uncertainty at each location. The work presented here uses a different

* Corresponding author.
E-mail address: matt.aitkenhead@hutton.ac.uk (M.J. Aitkenhead).

http://dx.doi.org/10.1016/j.geoderma.2015.08.034
0016-7061/

though still flexible approach (neural networks) and is used to generate maps with a different resolution (100 m instead of 1 km) and also to generate estimates with a finer profile depth resolution (1 cm instead of 5 cm); however, it is not as robust in terms of uncertainty assessment as the method of Poggio and Gimona (2014).

In addition to the measurement of soil organic carbon content at specific points, accurate national-scale assessment of SOC requires an ability to relate soil organic content to proxy data. This allows more detailed assessments to be carried out at reduced cost, or where soil assessments have not been detailed enough (Xu et al., 2011). Several approaches have been taken to modelling SOC content, ranging from statistical regression (Bauer et al., 2006; Li et al., 2010), through the development of pedotransfer functions that include multiple parameters at each point of interest (e.g. Chapman et al., 2013), to the use of approaches based on artificial intelligence (AI) and using environmental parameters (Sarmadian et al., 2009; Allahyaripour and Fazli, 2011). The use of approaches taken from AI to model environmental characteristics often includes the application of neural networks to develop predictive models from large, noisy and complex datasets. For soils, this kind of approach has been very successful in producing more accurate models of soil character than existing pedotransfer functions (e.g. Allahyaripour and Fazli, 2011; Haghverdi et al., 2012).

Historically, as most SOC stock estimates require information about the proportion of the soil that is organic matter, it has been important to have an accurate estimate of the bulk density (BD) of the soil. Methods of estimating this include relationships between organic matter content and BD (e.g. Prevost, 2004; Perie and Ouimet, 2008; Ruehlmann and Koerschens, 2009). It is important however to be able to predict BD from parameters other than organic matter content alone, as otherwise there will be a tendency to rely twice on the accuracy of the OM content assessment, leading to potential error propagation (Chapman et al., 2013). Heuscher et al. (2005) and Crowe et al. (2006) demonstrated that a number of parameters such as moisture content, texture and depth in the profile could also be used as predictors of bulk density. Texture in particular has been shown as effective in predicting bulk density in mineral soils (Keller and Hakansson, 2010; Brahim et al., 2012). Approaches that incorporate additional environmental parameters such as parent material and vegetation type have also been shown to be effective (Jalabert et al., 2010; Sakin, 2012). Having said this, any pedotransfer function or model predicting bulk density for topsoils must also incorporate some information about organic matter content, as variation in this single soil characteristic has a stronger effect on bulk density than any other.

Values given for LOI are often assumed to indicate the concentration of organic matter within the soil. For mineral soils with low organic matter content, or those soils containing carbonates or with high clay content, this assumption is incorrect (Salehi et al., 2011). At low organic matter content values, it has been shown that a significant fraction of the mass loss during LOI analysis is in fact due to water loss from certain minerals (Hoogsteen et al., 2015), meaning that using LOI to estimate soil organic matter or carbon content will lead to an overestimate if the common assumption is made that a simple ratio can be used to derive carbon content from LOI.

The estimation of soil organic carbon has historically been achieved by dividing the soil organic matter content by a value of 1.724 (the van Bemmelen factor). The use of this factor over the last 180 years is described in a useful, informative and entertaining review by Pribyl (2010), who demonstrates that a value closer to 2 is more accurate, although it is also argued that using a single conversion factor is not generally appropriate. There is evidence that the conversion factor should be (A) related to depth, with larger values nearer the surface where more labile organic matter contains a lower proportion of carbon, and (B) related to organic carbon content itself, with more organic soils containing a higher proportion of recaltricent organic matter with a relatively high proportion of carbon (and therefore having a lower conversion factor).

Depth within the profile is often shown to be an important predictor of bulk density (e.g. Suuster et al., 2011). The variation of bulk density, and therefore carbon content, with depth can be modelled effectively using parameterised functions of depth (e.g. Minasny et al., 2006). The effects of overburden pressure on bulk density at depth through compaction vary between soil types (Stutter et al., 2009), and between soils under different land cover types (Bachmann and Hartge, 2006). In order to successfully predict the variation of bulk density with depth across a country such as Scotland, which has a very complex landscape with a number of different soil types, any predictive model of bulk density variation with depth should either include a number of different factors, or be calibrated for site conditions (e.g. Hollis et al., 2012). Either way, specific information about each site is required. The model we demonstrate here includes topography, geology, soil and vegetation information to predict bulk density, soil organic matter content and soil carbon density with depth, in a single unified approach.

Here we demonstrate an approach to mapping soil organic matter, bulk density and soil organic carbon (SOC) content across Scotland to develop a pedotransfer-like model that is applied at a relatively fine spatial scale (100 m). This reduces two of the largest contributing factors to soil organic carbon estimate uncertainty, namely variation with soil density and variation at small scales. The modelling approach relies on neural networks, the use of which for modelling soil characteristics has been shown to be successful (Baker and Ellison, 2008; Borgesen et al., 2008).

2. Methods

2.1. Soil data

The data used in this work are contained within the Scottish Soils Database (SSD), one of the most detailed and systematic collections of national soil data in Europe. Formation of the SSD was initiated at the Macaulay Institute in the late 1970s, with several studies and surveys resulting in a completed database in 1987 (Brown et al., 1987). This database contained soil profile descriptions and chemical and physical analyses for several thousand sample points. One of the most significant components of the SSD is information from the National Soils Inventory of Scotland (NSIS) dataset, which is derived from an objective sampling of Scottish soils. Soil and site conditions of 721 locations throughout Scotland were sampled on a 10 km grid across the entire country, aligned with the National Grid of Great Britain (Lilly et al., 2010). Samples were taken at multiple depths from soil pits and analysed to determine their physical and chemical properties.

2.2. Soil data preparation

Not all of the data points in the Scottish Soils Database contain full records of every parameter. For this work, we selected those points that had the following information:

- Dry bulk density average
- Soil organic carbon (SOC) (only a subset of the data used had this information)
- LOI (loss on ignition) at 900 °C (values at 450 °C were also available and would have been preferred as the lower temperature gives a clearer measurement of organic matter content, but there were not as many measurements available at this temperature)
- Sample depth (i.e. the depth from the surface at which the sample was taken, this was taken as the mean value of the depths given for the top and bottom of the sampling depth).

As we are interested in the quantities of organic carbon stored in the soil but do not have direct information about organic carbon content for each of the samples used, it was necessary therefore to derive OC