Mapping soil profile depth, bulk density and carbon stock in Scotland using remote sensing and spatial covariates

Mapping soil depth and carbon stock in Scotland

AITKENHEAD, M.J., COULL, M.C.

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

matt.aitkenhead@hutton.ac.uk

+44 (0)1224 395257

Abstract

The spatial distribution of soil organic carbon is an important factor in land management decision making, climate change mitigation and landscape planning. In Scotland, where approximately one-quarter of the soils are peat, this information has usually been obtained using field survey and mapping, with digital soil mapping only carried out recently. Here a method is presented that integrates legacy survey data, recent monitoring work for peatland restoration surveys, spatial covariates such as topography and climate, and remote sensing data. The aim of this work was to provide estimates of the depth, bulk density and carbon concentration of Scotland's soils in order to allow more effective carbon stock mapping. A neural network model was used to integrate the existing data, and this was then used to generate a map of soil property estimates for carbon stock mapping at 100 metre resolution over Scotland. Accuracy assessment indicated that the depth mapping to the bottom of the organic layer was achieved with an r^2 of 0.67, while carbon proportion and bulk density were estimated with an r^2 of 0.63 and 0.79, respectively. Modelling of these three properties allowed estimation of soil carbon in mineral and organic soils in Scotland to a depth of one metre (3498 megatons) and overall (3688 megatons).

Keywords

Soil carbon; climate change; remote sensing; digital soil mapping; neural network

Highlights

- Scotland's soil organic carbon was mapped using a digital soil mapping approach.
- This provides a high-resolution map available for scientists, regulatory bodies and policymakers.
- The method largely agreed with previous work but improved the spatial resolution of the mapping.
- Significant soil carbon stocks are held in both organic (peat) and non-peat soils.

1. Introduction

Soil, particularly peat, contains significant quantities of carbon and presents both opportunities (mitigation through increased carbon storage) and risks (oxidation and GHG release) in relation to climate change. Peat also provides ecosystem services beyond carbon storage, including water storage and filtration, and biodiversity support. The management, protection and restoration of soil is therefore of importance for several environmental and policy reasons. To achieve appropriate soil carbon management, it is necessary to know where the carbon-rich soils are, and how much carbon they hold (and at what depth). In Scotland for example, approximately one-quarter of the country's surface area is classed as peat soil, but the spatial distribution and depth of this peat is unknown within existing soil map units.

An understanding of peat depth is also important to know in order to determine peat GHG emission rates and other properties/functions. Artz et al. (2006) showed that as depth increases, the microbial cycling of carbon changes in activity, relating to the level of humification in the peat. Dixon et al. (2017) showed that in deeper peats in North America, depth influences vegetation response to evaporative stress. Depth within peat also affects fungal community structure and plant functional group effects (Lamit et al., 2017).

Peat is defined in various soil classification systems as soil with an organic topsoil deeper than some defined depth, which varies according to the classification system used. Soil organic horizon depth data can come from several different sources, including traditional depth surveys, surveys carried out to assess the effects of peat restoration efforts, and paleoecological surveys (e.g. Ratcliffe & Payne, 2016). In addition to direct field survey and sampling, other approaches exist. The use of remote sensing data for delineating peatland areas is an important component of inventorying peatland carbon stocks (Nicoletti et al., 2003; Sheng et al., 2004) and is mapped better when used in combination with other covariates and with machine learning approaches (Minasny et al., 2019). The use of remote sensing and digital soil mapping approaches for mapping peat presence/absence specifically in Scotland has been previously demonstrated (Aitkenhead, 2017; Poggio et al., 2019). However, these works did not provide information on depth or carbon stock per unit area.

Peatland restoration activities in Scotland have, as a requirement for government funding, carried out grid-based depth and site condition surveys across over 200 peat bogs in order to provide evidence that peat not only exists at these locations, but that restoration work would be appropriate. As part of satisfying funding requirements, peat depth information across bogs must be provided. While estimating peat depth using Ground Penetrating Radar (GPR) can be carried out in the field and may be more accurate than manual probing (Parry et al., 2014), for reasons of cost and other practicalities, the manual approach remains the standard for peat depth survey at multiple points across an area of interest. Because of this, rod-based peat survey information has been carried out across many peat bogs in Scotland.

Manual soil survey work can provide baseline data for improved mapping of soils. Peat depth observations can be used to develop maps of estimated depth using spatial statistics and maps of other factors such as topography that influence the formation and development of peat (e.g. Rudiyanto Setiawan et al., 2014). Topographic data combined with statistical approaches and peat depth survey across a study area can be used to map estimated peat depth (Holden & Connolly, 2011; Parry et al., 2012). Buffam et al. (2010) also used digital elevation and slope information to estimate mean depths of small peat basins in Wisconsin in the United States.

Several approaches to the digital mapping of soil properties have been demonstrated, including regression functions (e.g. ten Caten et al., 2012), decision trees (Illes et al., 2011) and fuzzy classification (Odgers et al., 2011a, 2011b). One of the most flexible approaches to digitally mapping soils is the use of neural networks (e.g. Behrens et al., 2006), which are particularly effective in relating known parameters to unknowns of interest (McBratney et al., 2003). Zhang et al. (2017) gives a good review of different DSM (digital soil mapping) modelling methods and highlights recent advances in a number of approaches. They also identify neural networks as a strong approach in this field with many successes.

Wadoux (2019) used the LUCAS dataset with a neural network approach to produce soil maps with associated uncertainty estimates, another important factor that was highlighted by Zhang et al. (2017). Aitkenhead & Coull (2019) found an ANN approach for mapping soil classes in Scotland and found that even where the model assigned the wrong class, there was greater probability of it assigning a functionally similar class than one that was totally different. This indicates that relationships that exist between soil formation factors and soil properties can be captured within this modelling approach. Wadoux et al. (2019) and Padarian et al. (2019) used local covariates with a neural network approach to estimate soil properties at multiple depths, which is necessary in order to estimate soil carbon concentrations down the profile.

Many factors have been demonstrated as proxy indicators of soil properties, in accordance with the seminal hypotheses of Dokuchaev and Jenny (Florinsky, 2012), and have been used in DSM. These include vegetation (e.g. Sharma et al., 2006; Ballabio et al., 2012), topography (Sharma et al., 2006; Bodaghabadi et al., 2011; ten Caten, 2011), geology (Bui & Moran, 2001; Chagas et al., 2011) and climate (Fujii et al., 2013).

Digital soil mapping is very dependent on the accuracy and distribution of field data, and all applications from the field scale (e.g. Quenum et al., 2012) to the continental scale (e.g. Grunwald et al., 2011) attempt to maximise the effectiveness with which the available field data is used. The spatial distribution of field data is important not only for providing enough coverage of soil and relevant environmental properties, but also to allow statistical validity and robustness to any model developed.

The approach demonstrated here for mapping depth, carbon concentration and carbon stock in Scottish soils uses neural networks to integrate multiple spatial covariates and uses multiple data sources including the Scottish Soil Database (Brown et al., 1987; Lilly et al., 2004) and local peat depth surveys to provide training data. It is also intended to improve on the information provided by Aitkenhead & Coull (2016) in which soil carbon stocks were mapped to a depth of 1 metre, by estimating the depth and carbon stock of the full soil profile across the country.

2. Methods

2.1 Peatland Action survey data

Scottish Natural Heritage (SNH), as the Scottish Government agency responsible for environmental conservation and habitat protection, provides funding to landowners for restoration of degraded peatland in Scotland. As part of the funding application process, landowners must provide spatial information about the peatland to be restored, including a peat depth survey carried out in a grid across the site. This data is used by SNH as part of the assessment for awarding restoration funding.

Peat depth survey data for over 200 sites (the work is ongoing with additional sites added frequently) across Scotland was collated by SNH and made available for this peat depth mapping work. As of August 2018, there were over 10000 depth values. Landowners were required to provide this data in a specific format and using standard measurement techniques (e.g. 100 metre grid, use of marked rods). Further information about the assessment protocols and requirements are given at www.nature.scot/peatlandaction.

2.2 National Soil Inventory of Scotland data

The National Soils Inventory of Scotland (NSIS) datasets contain soil and site data taken from 10 km and 20 km sampling grids across Scotland (Lilly et al., 2010), with samples taken at different depths from multiple horizons in each profile. Soil types included in the dataset were peats/histosols, gleys, podzols, immature soils (alluvial, lithosols, rankers) and brown earths/cambisols. Sample analysis included organic carbon content and bulk density, along with depth down the profile. Maximum profile depth data used for these datasets was 250 cm.

2.3 Additional Scottish soil data

The Scottish Soil Database, of which the NSIS data is part, contains information from many additional soil survey campaigns going back as far as the 1940s. Each data point also contains the coordinates, in the UK Ordnance Survey grid reference system, at which the sample was taken. These coordinates were used to determine environmental parameters at each location, from several different datasets (see below).

The Scottish Soil Database was explored for sample data that provided depth, organic carbon or bulk density information. Criteria for selection were that the analytical method used was included and that the location information was considered accurate within 100 metres of that given. Combined with the Peatland Action and NSIS data, this gave a dataset with 10141 specific values of depth to bottom of organic soil material, 1527 values for bulk density (311 from peat) and 27833 values of organic carbon (with large/coarse fragments removed prior to sampling). In Section 2.5, further description is given of how this data was split into subsets for training in a manner designed to avoid overfitting from using neighboring sample points.

Figure 1 shows box plots of the distribution in values for the three variables of interest, while Figure 2 shows the distribution of survey points across Scotland that were used in this work, separated by dataset used.

Figure 1. Box plots of profile depth (A), bulk density (B) and organic carbon (C) for the datasets used.

Figure 2. Distribution of survey points used in this work: (A) NSIS data, (B) other Scottish Soil Database data, (C) Peatland Restoration data.

2.4 Spatial covariates

The following spatial datasets were used for mapping the three soil properties, both in generating training data for the model and for mapping soil organic carbon once the model was trained:

- Ordnance Survey 50 m resolution Panorama DEM (Digital Elevation Model).
- Land Cover Map 2007 25 m resolution (LCM2007) (Morton et al., 2011).
- Land Cover of Scotland 1:25 000 scale (LCS88) (Macaulay Land Use Research Institute, 1993).
- Soil Map of Scotland at 1:250 000 scale, providing information on the percentage presence of Major Soil Group (12 classes) within soil mapping units.
- Monthly mean temperature and rainfall data derived from UK Meteorological Office data. This was taken from 1460 Meteorological Office Stations from 1941 to 1970 interpolated to a 100 m resolution across Scotland (Matthews *et al.* 1994; Lilly & Matthews, 1994), using a combination of stepwise multiple linear regression followed by residual kriging.

- Geological class at 1:250 000, derived from parent material information on soil mapping units in the Scotland Soil Map.
- Landsat 8 data (30-60 metre resolution, depending on spectral band), used to generate a 98% cloud-free coverage of Scotland, captured during June and July 2017. This data was downloaded for free from the USGS data transfer service. Individual band values were used rather than indices, to allow the neural network model to identify suitable combinations of bands.

From the DEM, a total of 7 further topographic spatial datasets were generated. These included slope, overall curvature (second derivative of the DEM), profile curvature (in the direction of maximum slope) and plan curvature (perpendicular to the direction of maximum slope), aspect, aspect from North and aspect from East. These last two are the minimum angle between the actual aspect and North and East, respectively, and provide values for aspect that do not have a large numerical discontinuity between values slightly east and slightly west of North (i.e. to avoid 0°being the same as 360°).

The reason for using two land cover maps was that while the LCS88 is considered extremely accurate for land cover in 1988, it is now over thirty years old and land cover will have changed since it was produced. The LCM2007 dataset, while being more recent, is not considered as accurate for Scotland, particularly for grassland, heath and peat land cover types. For the LCM2007 and LCS88 datasets, a reduced categorisation was generated with only 10 classes, which allowed a separate map to be generated for each class type for the two land cover maps. The broad categorisation of land cover that was used was selected to allow both LCM2007 and LCS88 maps to be translated easily and consistently, and included the following categories: arable, improved grassland, rough grassland, heath, peat, bare ground, water, montane, coniferous forest and deciduous forest. Classes were selected largely based on definitions in the LCM2007 dataset, and in most cases the definition of the corresponding LCS88 class being assigned had an identical or near-identical definition. For heath and peat classes, some adjustment of the mapping to match the two systems was required. Also from the soil map information, we produced a map of parent material that contained 19 parent material types.

Each of the above spatial datasets was resampled to 100 metre resolution. Where a spatial dataset had a coarser spatial resolution than 100 metres (e.g. temperature and rainfall monthly means), it was subsampled and smoothed linearly between the existing points. Where a dataset had finer resolution than 100 metres (e.g. the land cover maps), the nearest cell to each 100 metres location was selected and used to represent that grid cell.

2.5 Neural network modelling

Backpropagation neural network (NN) models (Bishop, 1995) were developed to estimate soil properties based on input parameters. This network design had two hidden layers with ten nodes each and used a gradient descent value of 0.05. The number of hidden layers was set to two as we have found through previous experience that this provides better accuracy than having one layer and that three layers do not provide much improvement but adds to the computational cost. The activation of each node was calculated using a sigmoid function.

The number of input nodes X equalled the number of input parameters that exist in the training dataset, and the number of output nodes Y equalled the number of output parameters. All input parameters were normalised to lie within the range [0, 1], while the output parameters were normalised to lie within the range [0.25, 0.75]. This was done to avoid the need for extremely large node activation values due to the sigmoid function. Training steps were set at 100,000 after trial and error to find the best validation accuracy.

Model calibration was carried out using a k-fold cross-validation approach, with the full dataset split into 13 approximately equally sized subsets at random. Ten of these subsets were used for model testing in the standard 10-fold cross-validation approach (nine subsets used for training and one for testing, with this process repeated ten times). The remaining three subsets were used for validation using an ensemble of the ten trained models, with the mean of each output across ten models used to provide estimated values.

As the Peatland Action peat depth data came from clustered sampling areas, a further preprocessing step for the profile depth estimate was carried out when data points were assigned to one of the 13 subsets. When one data point was randomly assigned to a subset, all the data points from the same restoration site were included in that subset. This was done to ensure that model performance was not biased by testing models using data points that were spatially correlated with the training data. This will not remove all of the spatial correlation that exists within the dataset but as the sites are distributed across the whole of Scotland, it was assumed that it would minimize the effects of spatial correlation.

Subset clustering restrictions were also placed on the organic carbon and bulk density datasets to ensure that all the values from one soil profile were kept within the same subset. Additionally, data from profiles that were within 1000 metres of one another were also put into the same subset. This was done for the same reason as the clustering of the peat depth data, to avoid biasing the models. Analysis of the spatial autocorrelation of the three modelled variables gave the following semivariograms: depth (nugget = 13.3, range = 1600 m, sill = 54.6); bulk density (nugget = 0.12, range = 1500 m, sill = 0.39), organic carbon (nugget = 3.2, range = 900 m, sill = 10.8). For each variable, the semivariance at 1000 m was greater than 80% of the sill value determined, and so it was assumed that clustering values within 1000 m of one another was valid.

Prior to model training, all input and output variables were normalized to the range [0, 1] and then transformed using a power function selected to minimize the variance in histogram bucket size. This was done in order to reduce skewing in the population distribution of variables and used ten equally sized histogram buckets for each variable. Post-training, validation data estimates were transformed in reverse to achieve a similar population distribution and range of values as the training data. The effect of this was to train the models using data that was closer to being normally distributed, while still providing modelled estimates whose distribution reflected that of the properties studied. No analysis was carried out on the effects of transforming and back-transforming the data in this way, in terms of outliers or distribution tails.

Three NN models were developed, for (1) organic layer thickness, (2) carbon concentration and (3) bulk density. For the carbon content and bulk density models, depth was included as an input in order to model variation of these properties down the profile. The reasons for having separate models for carbon content and bulk density was that not all sample points had both variables measured, and that we found in practice that combining the two outputs within one NN model produced lower accuracy. Efforts to maintain correlation between these two properties (i.e. lower bulk density for higher carbon content) were not made, and as is shown in the Results, this did lead in some geographical locations to localized issues of high bulk density and high carbon content estimates.

2.6 Analysis of results

For each output variable, statistical evaluation of model performance was carried out using r^2, RMSE (Root Mean Square Error), MAE (Mean Absolute Error) and RPIQ (Ratio of Performance to Interquartile distance). RPIQ (Minasny & McBratney, 2013) was used as a useful additional metric to replace RPD, which is correlated with r^2 for large datasets and so therefore less useful.

2.7 Mapping and interpretation

The trained ensemble models were used to produce maps of bulk density, carbon concentration and peat depth. Values were calculated by taking the mean of all ten models that were individually trained as described above. Depth of organic layer was calculated as extending to where the carbon proportion became lower than a value of 20%. Definitions of various organic horizons differ in the WRB soil classification system between 12% and 20% carbon concentration, and this relatively high value was chosen as it was judged to match the Scottish soil classification system definition of organic soil most closely.

Estimates of carbon content and bulk density were made in 5 cm increments down the profile, to a maximum depth of 10 metres. At each increment, carbon proportion (the proportion of the soil that is carbon by weight) and soil bulk density estimates were used to calculate carbon stock density in kg per square metre for that 5 cm layer. These values were summed to provide estimates of carbon stock per square metre for both the organic layer and the full profile. This also enabled peat presence/absence to be modelled by using the threshold of 50 cm organic layer thickness which is used to discriminate peat from other soils using the Scottish soil classification system.

Model estimates were produced using the neural network models by using the spatial covariate datasets previously described at 100 metre resolution.

3. Results

3.1 Spatial covariates

Values for spatial covariates at the sample locations covered all or close to the full range found in the Scottish landscape. Table 1 lists these and their min/max values. Optimized power function transformation of the input variables as described above showed that many of the variables had skewed distributions (e.g. elevation which was skewed towards lower values). The normalization process produced variable distributions that were much more evenly distributed than pre-normalization, with a lower standard deviation in range populations.

Table 1. Spatial covariate value ranges at sample locations.

[*] Monthly mean

3.2 Model evaluation

Table 2 shows the statistical evaluation of the three NN models. Bulk density was modelled most accurately, both in terms of r^2 and RPIQ (for which it is common to take values above 2 as 'good' although this does not have a mathematical or statistical basis). Values for all three output variables had non-normal distributions, with depth skewed towards smaller values and both carbon and bulk density having bimodal distributions (lots of low and high values, but fewer mid-range values). Bulk density was also correlated with depth, which may explain the higher accuracy of this model (possibly because of the factor of load weight causing increased compression with depth).

Table 2. Validation statistics for depth, carbon content and bulk density models.

It is important to note that while the RMSE value for organic profile depth estimation is high overall, the RMSE for depths less than 200 cm is much smaller at only 18.8 cm. Of the 10141 depth data points used, 86% had depth shallower than 200 cm. This means that estimates of carbon stock deeper than 200 cm depth are less reliable, but also that they have less impact on the total C (carbon) stock estimate.

In calculating carbon stock for a location and depth, it is difficult to estimate C stock error as carbon and bulk density are to some extent correlated to one another. Figure 3 gives a diagrammatic representation of the error rates for the carbon and bulk density models, showing where carbon stock error rates are likely to be low, medium or high for different ranges of carbon and bulk density. The error rate range given is as a percentage of the estimated value in each case and is derived from an evaluation of the error distributions in the model outputs.

These error distributions are not the same across all values of the modelled values, possibly due to the number of examples within each value range, or to factors not considered in the modelling that influence soils of different types.

Figure 3. Representation of likely error ranges and estimated likelihoods of occurrence for different carbon proportion and bulk density combinations.

3.3 Mapping

Maps of peat depth and total soil profile depth are shown in Figure 4. The legend in each case is not linear but used the ‘standard deviation’ distribution in ESRI ArcMap 10.1. Deeper soil profiles in the Flow Country (northern mainland) and Lewis (northern Outer Hebrides) are associated with deeper peat depth. Soils in western parts of mainland Scotland however have predominantly shallower full profile depths than in the east of the country, despite having deeper peat profiles. This indicates that in the west of Scotland, soil organic horizons tend to make up a greater proportion of each soil profile.

Eastern Scotland is predominated by flatter topography and a drier climate with mineral, agricultural soils derived from Brown Earths (Cambisols) and Gleys. Meanwhile, western Scotland has a higher proportion of steep slopes and high elevations with a wetter climate, meaning shallower soils with more organic matter as stated above.

Figure 4. Maps of peat depth (A) and total profile depth (B) for Scotland.

Figure 5 shows maps of carbon stock calculated in the surface 5 cm, and in the depth range 45-50 cm. These were calculated for each depth range by multiplying the bulk density (g cm^{-3}) and carbon concentration (percentage divided by 100 or g g^{-1}) at that depth and location to obtain a carbon density in units of g cm^{-3}. This was then converted, using a 5 cm thickness layer (500 m^3 ha^{-1}) in units of kg ha^{-1}.

These two maps for different depths show dramatic differences in the distribution of low and high carbon stock density. At the surface layer, east coast and Central Belt areas have areas with much higher carbon stocks despite these soils being predominantly mineral because they have a high topsoil bulk density (values around 1.5 g cm^{-3} are common).

This means that even for a relatively low topsoil organic matter content for Scotland (5%), carbon stock values for these soils tend to be greater than the relatively low-density organic soils near the surface. Additionally, for a large proportion of the year, arable soils contain a large amount of crop roots in the top few centimeters, which is included in the organic carbon stock estimate.

At 45-50 cm depth however, the soils with the greatest stock density are the peaty podzols and peats of western Scotland, while the arable soils of the east coast have very little carbon at this depth. Deep peats in the Outer Hebrides (the island chain off the north-west coast) and the far north of Scotland have intermediate carbon stock density at this depth range, as their bulk density is still lower than the organomineral peaty podzols. The very high values represented in Figure 5 are outliers; most of the values even in high C stock areas are less than 50,000 kg ha^{-1} per 5 cm layer.

Figure 5. Maps of carbon stock per hectare at (A) 0-5 cm and (B) 45-50 cm depth.

Figure 6 shows the maps of carbon stock per hectare at 95-100 cm and 195-200 cm. These maps (particularly for the deeper layer) show large areas in white where the profile is modelled as not reaching to this depth. Where there is peat however, the larger C stock values per hectare are approximately the same as for shallower layers. At 100 cm, the greatest contribution to soil

C stocks is from the Outer Hebrides, the Flow Country (northern mainland) and the Shetland Isles (far north). There are some isolated pockets of deep peat in other areas, but these are relatively small. At 200 cm, the greatest contribution is from the Flow Country and the western islands – there is soil carbon scattered around the country in other places at this depth, but the stock density is relatively low.

Figure 6. Maps of carbon stock per hectare at (A) 95-100 cm and (B) 195-200 cm depth.

Figure 7 shows total carbon stock to 100 cm depth for peats and for all soils. Large areas on map (A) are white because there is no peat, and these correspond in map (B) to large areas of low carbon stock. However, there are significant increases in the north and west of the country where carbon stock has fewer gaps when all soils are included (Fig. 6B). At 1 metre depth for example, the Isle of Skye shows up as having a large proportion by area of high peat carbon stock in Fig. 6A in comparison to nearby mainland areas, while for all soils (Fig 6B) this area of high carbon stock also extends across Lochaber (far west mainland) and the larger islands just west of Lochaber.

Figure 7. Maps of total carbon per hectare to 100 cm in (A) peat and (B) all soils.

In Figure 8, maps of total peat and total soil carbon per hectare are shown. These include the full profile depth as modelled. While the distribution of peat soils in Figure 8 is the same as in Figure 7, the maximum and distribution of values is noticeably different. Deeper peats in the Flow Country, Outer Hebrides and south-west of Scotland have added more carbon than soils in Lochaber and Mull, which are modelled as generally being between one and two meters in depth.

The maximum carbon stock density per hectare also increases greatly from Figure 7 to Figure 8, due to small areas of very deep peat. The Flow Country and northern Outer Hebrides (Isle of Lewis) also have much more carbon stock values when looking at the full profile. It is important to note that the small difference between maximum peat carbon stock and total soil maximum carbon stock are due to the presence of some soil carbon below the peat profile depth.

Figure 8. Maps of (A) total peat carbon and (B) total soil carbon per hectare.

3.4 Carbon stock variation with depth

In Figure 9, the values of carbon stock per 5 cm layer across the whole of Scotland are given, for peat, non-peat (mineral soils) and all soils. In the Scottish soil terminology, soils are organic, organomineral (having an organic topsoil that is less than 50 cm thick) or mineral – here all non-peat soils are referred to as mineral. These show that at depths below 40 cm, non-peat soils contribute more carbon to the total soil C stock than peats, but that this contribution from non-peat soils drops rapidly with depth and is negligible below approximately 80 cm. As the technical definition of peats in Scotland is that they have an organic layer thicker than 50 cm, this partially explains this reduction in contribution from non-peat soils at around this depth – if the organic layer is deeper than 50 cm that soil is defined as peat and is not included in the 'non-peat' contribution.

The contribution from peats continues to increase up to 60 cm depth, after which it also falls rapidly. This is because while most peats are deeper than 60 cm, the proportion within each depth range decreases with depth. It tails off more gradually than for non-peat soils, but after 150-200 cm is also considered negligible. This agrees with the information shown in Figure 6, where soil C stocks at 200 cm are from many very small areas of deep peat.

Figure 9. Graph at 5 cm depth increment showing total soil carbon, peat soil carbon and non-peat soil carbon in 5 cm layers to 200 cm depth.

The total carbon in peat soils is calculated as 1889 MT (1719 MT to 100 cm and 1883 MT to 200 cm), and that in non-peat soil is calculated as 1799 MT (1779 to 100 cm and 1797 MT to 200 cm). This gives a total soil carbon estimate for Scotland of 3688 MT (3498 MT and 3680 MT to 100 and 200 cm respectively). The total area estimated as peat is 23958 km^2, which is 29.1% of Scotland's surface area (compared to the estimate of 22.4% by Chapman et al., 2009).

4. Discussion

The profile depth, carbon concentration and bulk density models developed were considered sufficiently accurate to allow estimates of carbon stock with depth, and to produce maps of this stock at 100 metre resolution across Scotland. Combining survey datasets from different sources made it possible to do this through a neural network modelling approach.

The peatland restoration survey data is important as it provides depth information for many peatland locations. The existing Scottish Soils Database has several hundred data points for peat but not enough to provide a representative sampling of Scotland's peat soils for modelling purposes. One issue with the restoration survey data however is that it may be biased towards sites that require restoration; if true, these sites are more likely to be degraded. While peat degradation does not always imply carbon loss, it can if the degradation has led to erosion and therefore loss of carbon and therefore depth. It is possible therefore that deep peat values have been underestimated. Certainly, there is anecdotal evidence and a small number of physical observations in some datasets for deep peat depths greater than 10 meters in Scotland.

The three datasets used have different spatial distributions, numbers of points and depth ranges. Because of this, there may be some impact on model accuracy for different soil types or areas of Scotland. The NSIS data is distributed evenly in a spatial grid (Figure 2A) and so therefore can be assumed to provide a good representation of the range of values seen for soil carbon and bulk density, but there are very few profiles in this dataset that go deeper than 1 metre and in practice, some profiles that were deeper than 1 metre were not explored below this depth. However, while this potentially biases any model, the reality of Scotland's soils is that the vast majority of them are shallower than 1 metre (71% are recorded in the NSIS 10-km grid dataset as having rock within 1 metre of the surface) and so this is not likely to be causing much of a problem. Profiles deeper than 1 metre are nearly always peat (over 99% of those in the NSIS dataset and all the Peatland Restoration data) so the data is considered to provide a reasonable distribution of values for this soil type and the modelling carried out.

The other Scottish Soil Database profile data used in this work is concentrated in lowland, commonly agricultural areas (Figure 2B). It therefore provides more data on soils of high bulk density and low carbon content, on lower altitude and in warmer, drier climates than the mean for Scotland. This potential bias is assumed to have been at least partially corrected for in the dataset normalisation process described in the methodology but may still have biased the environmental data used in the model to be more accurate for these types of conditions. The Peatland Restoration data may have partially countered this by providing a bias towards cool, wet climatic areas and flat slopes at higher elevations. Arguably therefore, the environmental conditions least represented within the model training is steeper slopes, as only the grid-based NSIS sampling system will have captured the representative slope distribution for Scotland. It is therefore possible that the model is less accurate on steep slopes than flat ones.

The models developed are considered to have given good estimates of bulk density and moderately good estimates of profile depth and soil carbon concentration. A visual evaluation of the mapped results was also carried out by overlaying colour-coded maps of C and bulk density produced by the model and spending time scrolling through this zoomed map. This (admittedly subjective) analysis showed some areas on the west coast of Scotland where bulk density and C proportion did not appear to match; for a small number of examined locations both bulk density and C were high in the top 50 cm. As soil C and bulk density are inversely correlated, this was not realistic and as mentioned in the Methodology, is likely due to having separate models for C and bulk density. However, for most locations and depths examined, the values of C and bulk density made sense in relation to one another.

The total area of peat estimated (23958 km^2) is higher than previous estimate of 19000 km^2 (Chapman et al., 2009). A lot of this difference appears to be due to soils in western Scotland that are estimated as having organic profile depth just over 50 cm deep, and thus being classed as peats. In existing maps of Scottish soils, this area is dominated by peaty gleys and peaty podzols rather than peats, and further work should be carried out to evaluate this discrepancy between legacy mapping and the current work.

An important piece of information that has not been fully addressed here is the uncertainty associated with the estimates of soil carbon stock at each location. There is also uncertainty around the area of peat, linked to estimates of depth which may be producing false negatives or false positives. This uncertainty can be quantified using multiple runs of the same model (e.g. Poggio & Gimona, 2014), runs of multiple model approaches (Poggio et al., 2019; Poggio et al., 2018) or statistical evaluation of the sample data used for the mapping (e.g. Odgers et al., 2015). However, this quantification is highly dependent on the approach used to determine the uncertainty (and the actual definition of the uncertainty itself), and the author does not feel that this area has been tackled sufficiently in the Digital Soil Mapping domain.

The estimates of total soil carbon to 1 metre depth are close to existing estimates (2954 MT for Aitkenhead & Coull (2016) vs. 3498 in this work). Chapman et al. (2009) estimated total peat carbon at 1610 MT compared to this estimate of 1889 MT. Estimates in the literature for total soil carbon stocks to 1 metre include 2187 MT (Bradley et al., 2005), 2055 MT (Chapman et al., 2013) and ~3000 MT (Campbell et al., 2012). The Scottish Government states on the Scotland's Soils website that "Scotland's soils contain more than 3000 million tonnes of carbon", with this information coming from the State of Scotland's Soils Report in 2011 (Dobbie et al., 2011).

The greater contribution of soil carbon stocks near the surface from non-peat soils indicates that while peat restoration and conservation is vital, vulnerable mineral soils should also be monitored and protected. Carbon closer to the surface is inherently more vulnerable to a range of natural processes and human activities, and so is more likely to be lost. North-west and south-west Scotland have large areas of carbon-rich soils that are not technically peat due to having an organic topsoil less than 50 cm deep, and these soils hold a significant amount of Scotland's carbon stocks.

While this work is potentially useful in providing improved resolution mapping of Scotland's soil carbon stocks, there are ways in which the work could be improved. First and foremost is that of spatial resolution; even a 100 metre grid cell size will miss a lot of small peat 'pockets' that should be factored into local land management decision-making. The overall model accuracy should be improved, particularly that of profile (and organic profile) depth. While errors in estimation of depth of deep peat (>2 metres) are arguably less important as these areas do not contribute significantly to Scotland's soil carbon stock, the estimation of organic profile depth for shallower peats should be improved.

Building on this work and acknowledging that greater accuracy is important, it is also vital to recognize that of even greater urgency is the need to make effective use of this kind of data. Legislation and advice for land managers should be improved to ensure that soil carbon stocks are restored and protected, and that land managers are aware of the multiple benefits accrued from large soil carbon stocks.

Acknowledgements

The author would like to acknowledge Dr Rebekka Artz, Dr Zizis Gagkas from the James Hutton Institute for their help in developing this paper, Dr Jackie Potts from BioSS (Biomathematics & Statistics Scotland) for help with geostatistical interpretation and Dr Patricia Bruneau at Scottish Natural Heritage for advice relating to Peatland Action data.

This work was supported by RESAS (the Scottish Government's Rural and Environment Science and Analytical Services Division). RESAS had no role in the study design, data analysis or reporting of this work.

References

1. Aitkenhead, M.J. (2017). Mapping peat in Scotland with remote sensing and site characteristics. European Journal of Soil Science, 68(1), 23-38.
2. Aitkenhead, M.J. & Coull, M.C. (2019). Digital mapping of soil ecosystem services in Scotland using neural networks and relationship modelling-Part 1: Mapping of soil classes. Soil Use and Management 35(2), 205-216. DI 10.1111/sum.12492
3. Aitkenhead, M.J. & Coull, M.C. (2016). Mapping soil carbon stocks across Scotland using a neural network model. Geoderma, 262, 187-198.
4. Artz, R.R.E., Chapman, S.J. & Campbell, C.D. (2006). Substrate utilisation profiles of microbial communities in peat are depth dependent and correlate with whole soil FTIR profiles. Soil Biology & Biochemistry, 38(9), 2958-2962.
5. Ballabio, C., Fava, F. & Rosenmund, A. (2012). A plant ecology approach to digital soil mapping, improving the prediction of soil organic carbon content in alpine grasslands. Geoderma, 187, 102-116.

6. Beamish, D. (2014). Peat Mapping Associations of Airborne Radiometric Survey Data. Remote Sensing, 6(1), 521-539.

7. Behrens, T. & Scholten, T. (2006). Digital soil mapping in Germany - a review. Journal of Plant Nutrition and Soil Science, 169(3), 434-443.

8. Bishop, C.M. (1995). Neural networks for pattern recognition. 498pp. Clarendon Press, Oxford.

9. Bodaghabadi, M.B., Salehi, M.H., Martinez-Casasnovas, J.A., Mohammadi, J., Toomanian, N. & Borujeni, I.E. (2011). Using Canonical Correspondence Analysis (CCA) to identify the most important DEM attributes for digital soil mapping applications. Catena, 86(1), 66-74.

10. Bradley, R.I., Milne, R., Bell, J., Lilly, A., Jordan, C. & Higgins, A. (2005). A soil carbon and land use database for the United Kingdom. Soil Use and Management, 21(4), 363-369.

11. Buffam, I., Carpenter, S.R., Yeck, W., Hanson, P.C. & Turner, M.G. (2010). Filling holes in regional carbon budgets: Predicting peat depth in a north temperate lake district. Journal of Geophysical Research-Biogeosciences, 115, AR G01005.

12. Bui, E.N., Moran, C.J.A.F. (2001). Disaggregation of polygons of surficial geology and soil maps using spatial modelling and legacy data. Geoderma 103(1-2), 79-94. DI 10.1016/S0016-7061(01)00070-2

13. Campbell, C.D., Lilly, A., Towers, W., Chapman, S.J., Werritty, A. & Hanley, N. (2012). Land use and a low-carbon society. Earth and Environmental Science Transactions of The Royal Society of Edinburgh, 103(2), 165-173.

14. Chapman, S., Bell, J., Donnelly, D. & Lilly, A. (2009). Carbon stocks in Scottish peatlands. Soil Use and Management, 25, 105-112.

15. Chapman, S.J., Bell, J.S., Campbell, C.D., Hudson, G., Lilly, A., Nolan, A.J., … Towers, W. (2013). Comparison of soil carbon stocks in Scottish soils between 1978 and 2009. European Journal of Soil Science, 64(4), 455-465.

16. Chagas, C.S., de Carvalho Jr, W. & Bhering, S.B. (2011). Integration of Quickbird data and terrain attributes for digital soil mapping by artificial neural networks. Revista Brasileira de Ciencia do Solo, 35(3), 693-704.

17. Dixon, S.J., Kettridge, N., Moore, P.A., Devito, K.J., Tilak, A.S., Petrone, R.M., … Waddington, J.M. (2017). Peat depth as a control on moss water availability under evaporative stress. Hydrological Processes, 31(23), 4107-4121.

18. Dobbie, K.E., Bruneau, P.M.C., Towers, W. (eds.), 2011. The State of Scotland's Soil. Natural Scotland, www.sepa.org/land/land_publications.aspx.

19. Florinsky, I.V. (2012). The Dokuchaev hypothesis as a basis for predictive digital soil mapping (on the 125th anniversary of its publication). Eurasian Soil Science, 45(4), 445-451.

20. Fujii, K., Morioka, K., Hangs, R., Funakawa, S., Kosaki, T. & Anderson, D.W. (2013). Importance of climate and parent material on soil formation in Saskatchewan, Canada as revealed by soil solution studies. Pedologist, 57(1), 27-44.

21. Grunwald, S., Thompson, J.A. & Boettinger, J.L. (2011). Digital Soil Mapping and Modeling at Continental Scales: Finding Solutions for Global Issues. Soil Science Society of America Journal, 75(4), 1201-1213.

22. Holden, N.M. & Connolly, J. (2011). Estimating the carbon stock of a blanket peat region using a peat depth inference model. Catena, 86(2), 75-85.

23. Illes, G., Kovacs, G. & Heil, B. (2011). Comparing and evaluating digital soil mapping methods in a Hungarian forest reserve. Canadian Journal of Soil Science, 91(4), 615-626.

24. Keaney, A., McKinley, J., Graham, C., Robinson, M. & Ruffell, A. (2013). Spatial statistics to estimate peat thickness using airborne radiometric data. Spatial Statistics, 5, 3-24.

25. Kempen, B. (2011). Updating soil information with digital soil mapping, 218pp.

26. Lamit, L.J., Romanowicz, K.J., Potvin, L.R., Rivers, A.R., Singh, K., Lennon, J.T., … Lilleskov, E.A. (2017). Patterns and drivers of fungal community depth stratification in Sphagnum peat. FEMS Microbiology Ecology, 93(7), AR fix082.

27. Lilly, A., Bell, J.S., Hudson, G., Nolan, A.J. & Towers, W. (Compilers) (2010). National Soil Inventory of Scotland 1 (NSIS_1): site location, sampling and profile description protocols. (1978-1988). Technical Bulletin, The Macaulay Land Use Research Institute.

28. Lilly, A. & Matthews, K.B. (1994). A soil wetness class map for Scotland: new assessments of soil and climate data for land evaluation. Geoforum, 25(3), 371-379.

29. Lilly, A., Towers, W., Malcolm, A. & Paterson, E. (2004). Report on a workshop on the development of a Scottish Soils Knowledge and Information Base (SSKIB). Macaulay Land Use Research Institute Report, 35pp.

30. Macaulay Land Use Research Institute (1993). The Land Cover of Scotland 1988 Final Report, Macaulay Land Use Research Institute.

31. Matthews, K.B., MacDonald, A., Aspinall, R.J., Hudson, G., Law, A.N.R. & Paterson, E. (1994). Climatic soil moisture deficit – Climate and soil data integration in a GIS. Climatic Change, 28, 273-287.

32. McBratney, A.B., Santos, M.L.M. & Minasny, B. (2003). On digital soil mapping. Geoderma, 117(1-2), 3-52.

33. Minasny, B. & McBratney, A.B. (2013). Why you don't need to use RPD. Pedometron 33, 14-15.

34. Minasny, B., Berglund, O., Connolly, J., Hedley, C., de Vries, F., Gimona, A. et al. (2019). Digital mapping of peatlands – a critical review. Earth-Science Reviews 196. DI 10.1016/j.earscirev.2019.05.014.

35. Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G. & Simpson, I.C. (2011). Final report for LCM2007 – the new UK land cover map. CS Technical Report No 11/07 NERC/Centre for Ecology & Hydrology 108pp. (CEH project number: C03259).

36. Nicoletti, V., Silvestri, S., Rizzetto, F., Tosi, L., Putti, M. & Teatini, P. (2003). Use of remote sensing for the delineation of surface peat deposits south of the Venice Lagoon (Italy). IGARSS 2003: IEEE International Geoscience and Remote Sensing Symposium, Jul 21-25, 2003, Toulouse, France.

37. Odgers, N.P., McBratney, A.B. & Minasny, B. (2011). Bottom-up digital soil mapping. I. Soil layer classes. Geoderma, 163(1-2), 38-44.

38. Odgers, N.P., McBratney, A.B. & Minasny, B. (2011). Bottom-up digital soil mapping. II. Soil series classes. Geoderma, 163(1-2), 30-37.

39. Odgers, N.P., McBratney, A.B. & Minasny, B. (2015). Digital soil property mapping and uncertainty estimation using soil class probability rasters. Geoderma 237, 190-198. https://doi.org/10.1016/j.geoderma.2014.09.009

40. Padarian, J., Minasny, B., McBratney, A.B. (2019). Using deep learning for digital soil mapping. SOIL 5(1), 79-89. DI 10.5194/soil-5-79-2019

41. Parry, L.E., Charman, D.J. & Noades, J.P.W. (2012). A method for modelling peat depth in blanket peatlands. Soil Use and Management, 28(4), 614-624.

42. Parry, L.E., West, L.J., Holden, J. & Chapman, P.J. (2014). Evaluating approaches for estimating peat depth. Journal of Geophysical Research-Biogeosciences, 119(4), 567-576.

43. Poggio, L., & Gimona, A. (2014). National scale 3D modelling of soil organic carbon stocks with uncertainty propagation – an example from Scotland. Geoderma 232, 284-299. https://doi.org/10.1016/j.geoderma.2014.05.004.

44. Poggio, L., Gimona, A., Gagkas, Z. & Lilly, A. (2018). 3D digital soil mapping for Scottish soils using remote sensing., Internal Report for RESAS.

45. Poggio, L., Lassauce, A., Gimona, A. (2019). Modelling the extent of northern peat soil and its uncertainty with Sentinel: Scotland as example of highly cloudy region. Geoderma 346, 63-74. DI 10.1016/j.geoderma.2019.03.017.

46. Quenum, M., Nolin, M.C. & Bernier, M. (2012). Digital mapping of soil phosphorus maximum sorption capacity. Canadian Journal of Soil Science, 92(5), 733-750.

47. Ratcliffe, J. & Payne, R.J. (2016). Palaeoecological studies as a source of peat depth data: A discussion and data compilation for Scotland. Mires and Peat, 18, AR UNSP 13.

48. Rudiyanto Setiawan, B.I., Arief, C., Saptomo, S.K., Gunawan, A. & Kuswarman Sungkono Indriyanto, H. (2014). Estimating distribution of carbon stock in tropical peatland using a combination of an Empirical Peat Depth Model and GIS. In Eds. Setiawan, Y., Lubis, M.I., Prasetyo, L.B., Siregar, I.Z., Effendi, H., 1st International Symposium on LAPAN-IPB Satellite (LISAT) for food security and Environmental Monitoring, Nov 25-26, 2014. Bogor, Indonesia.

49. Sharma, S.S., Mohanty, B.P., Zhu, J. (2006). Including topography and vegetation attributes for developing pedotransfer functions. Soil Science Society of America Journal 70(5), 1430-1440. DI 10.2136/sssaj2005.0087

50. Sheng, Y.W., Smith, L.C., MacDonald, G.M., Kremenetski, K.V., Frey, K.E., Velichko, A.A., … Dubinin, P. (2004). A high-resolution GIS-based inventory of the west Siberian peat carbon pool. Global Biogeochemical Cycles, 18(3), AR GB3004.

51. ten Caten, A., Diniz Dalmolin, R.S., Mendonca-Santos, M.L. & Giasson, E. (2012). Digital soil mapping: characteristics of the Brazilian approach. Ciencia Rural, 42(11), 1989-1997.

52. ten Caten, A., Diniz Dalmolin, R.S., Pedron, F.A. & Mendonca-Santos, M.L. (2011). Principal components as predictor variables in digital mapping of soil classes. Ciencia Rural, 41(7), 1170-1176.

53. Wadoux, A.M.J.C. (2019). Using deep learning for multivariate mapping of soil with quantified uncertainty. Geoderma 351, 59-70. DI 10.1016/j.geoderma.2019.05.012

54. Wadoux, A.M.J.C., Padarian, J., Minasny, B. (2019). Multi-source data integration for soil mapping using deep learning. SOIL 5(1), 107-119. DI 10.5194/soil-5-107-2019

55. Werban, U., Behrens, T., Cassiani, G. & Dietrich, P. (2010). iSOIL: An EU Project to Integrate Geophysics, Digital Soil Mapping, and Soil Science. In Viscarra-Rossel, R.A., McBratney, A.B., Minasny, B, Proximal Soil Sensing: Progress in Soil Science, 103-110.

56. Zhang, G., Liu, F. & Song, X. (2017). Recent progress and future prospect of digital soil mapping: a review. Journal of Integrative Agriculture 16(12), 2871-2885. DI 10.1016/S2095-311(17)61762-3.

Table Captions

Table 1. Spatial covariate value ranges at sample locations.

Table 2. Validation statistics for depth, carbon content and bulk density models.

Tables

Covariate	Minimum	Maximum	Covariate	Minimum	Maximum
Elevation /m	0	1237	Soil type (%)	0	100
Curvature	-1.62	1.50	Temp (°C)[*]	-2.2	16.5
Slope /°	0	40.42	Rainfall (mm)[*]	32	477
Aspect /°	0	359.68	Geology	All types	All types
Land cover	All types	All types			

Variable	Min	Max	r^2	RMSE	MAE	RPIQ
Depth /cm	0	1000	0.67	58.3	42.5	1.69
Carbon /%	0.01	70.84	0.63	8.1	5.4	2.13
Bulk density /g cm^{-3})	0.05	1.80	0.79	0.17	0.11	4.62

Figure captions

Figure 1. Box plots of profile depth (A), bulk density (B) and organic carbon (C) for the datasets used.

Figure 2. Distribution of survey points used in this work: (A) NSIS data, (B) other Scottish Soil Database data, (C) Peatland Restoration data.

Figure 3. Representation of likely error ranges and estimated likelihoods of occurrence for different carbon proportion and bulk density combinations.

Figure 4. Maps of peat depth (A) and total profile depth (B) for Scotland.

Figure 5. Maps of carbon stock per hectare at (A) 0-5 cm and (B) 45-50 cm depth.

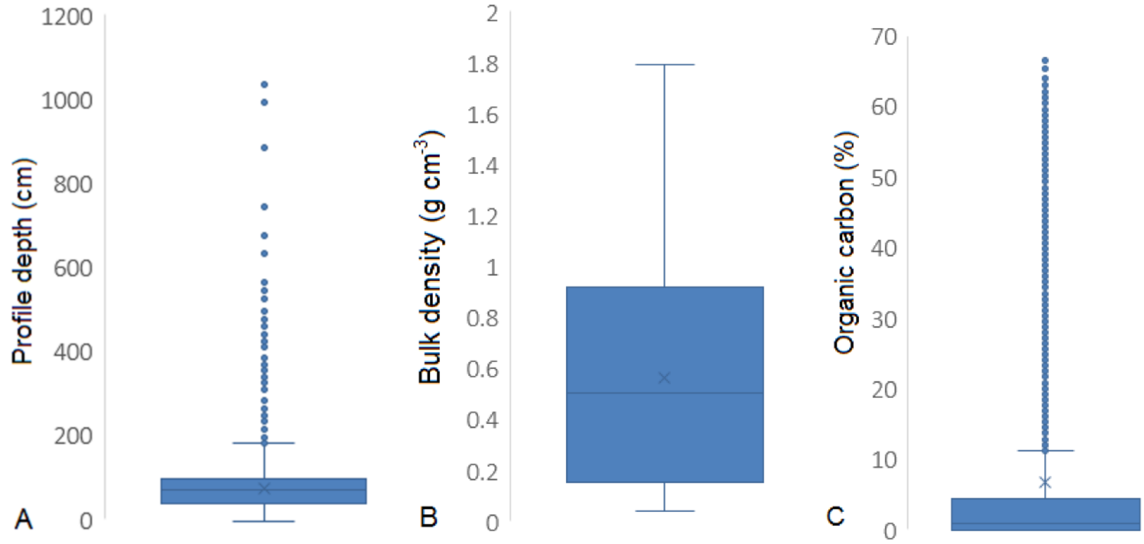
Figure 6. Maps of carbon stock per hectare at (A) 95-100 cm and (B) 195-200 cm depth.

Figure 7. Maps of total carbon per hectare to 100 cm in (A) peat and (B) all soils.

Figure 8. Maps of (A) total peat carbon and (B) total soil carbon per hectare.

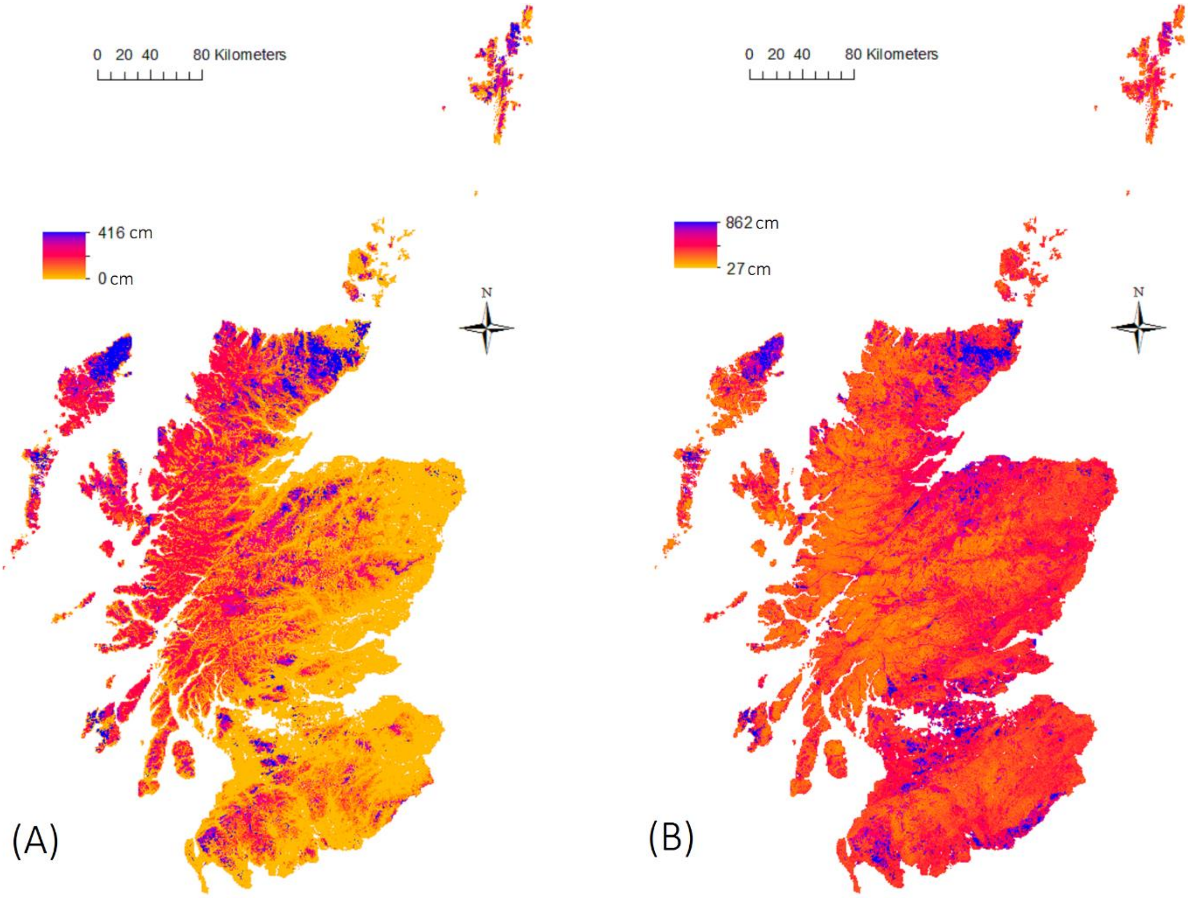
Figure 9. Graph at 5 cm depth increment showing total soil carbon, peat soil carbon and non-peat soil carbon in 5 cm layers to 200 cm depth.

Figures

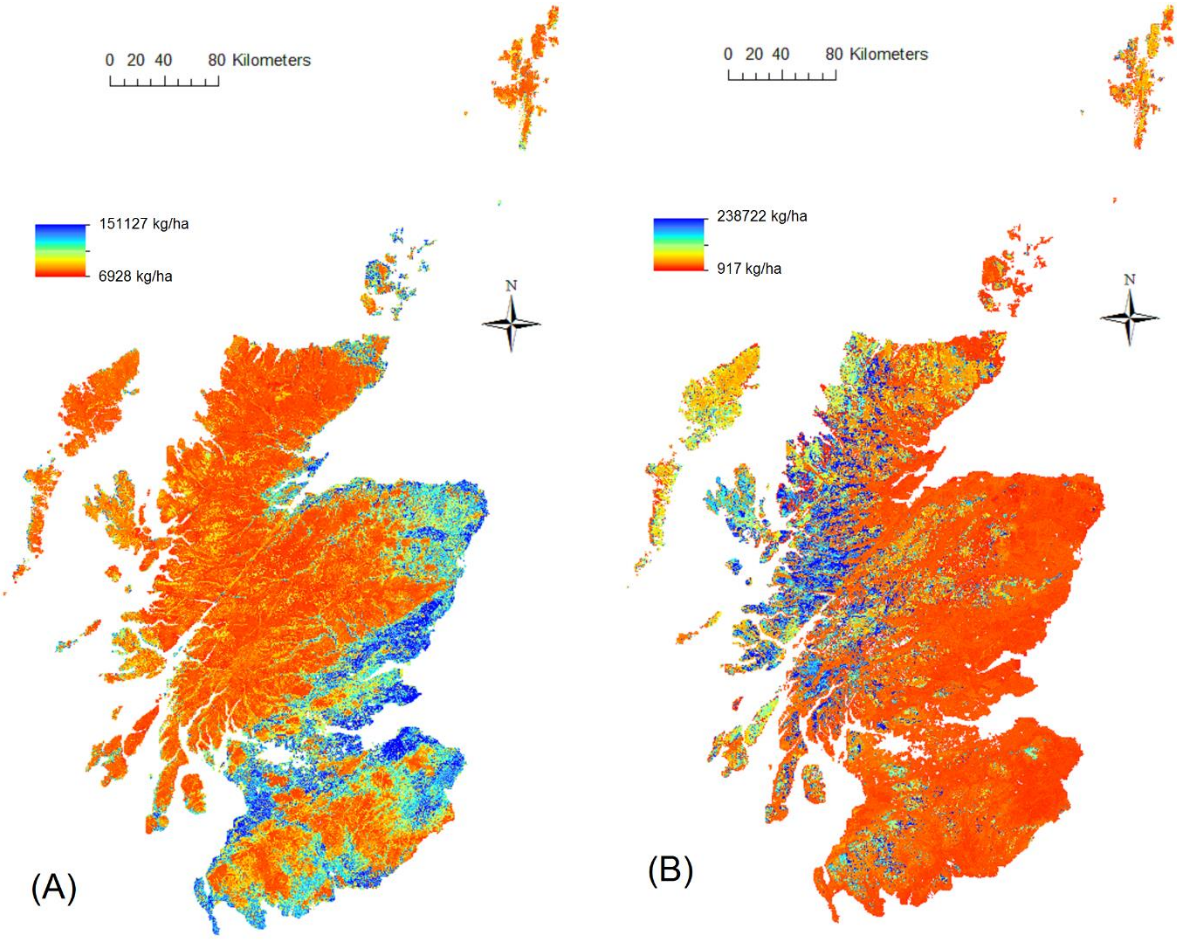

Profile depth (cm)
1200
1000
800
600
400
200
0
A
Bulk density (g cm-3)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
B
Organic carbon (%)
70
60
50
40
30
20
10
0
C

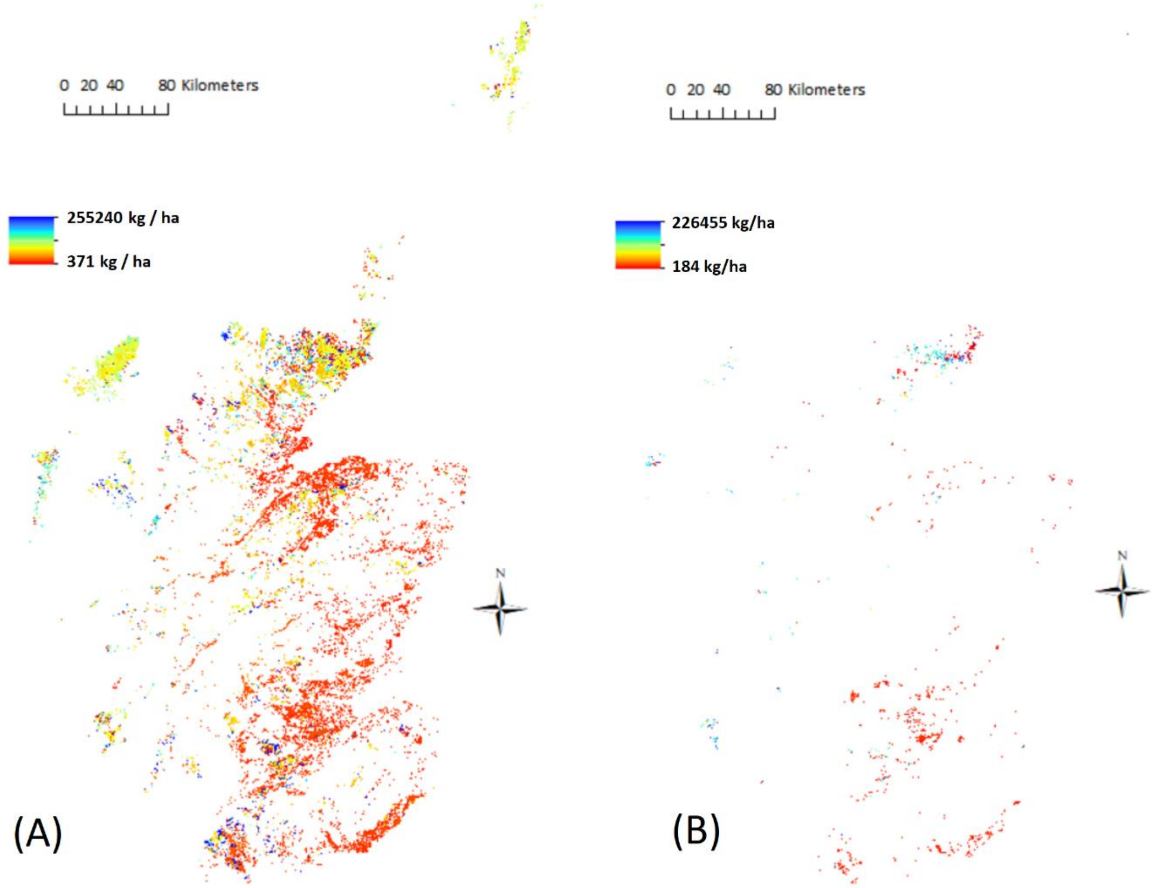

A
B
C
0 25 50 75 100
Kilometers

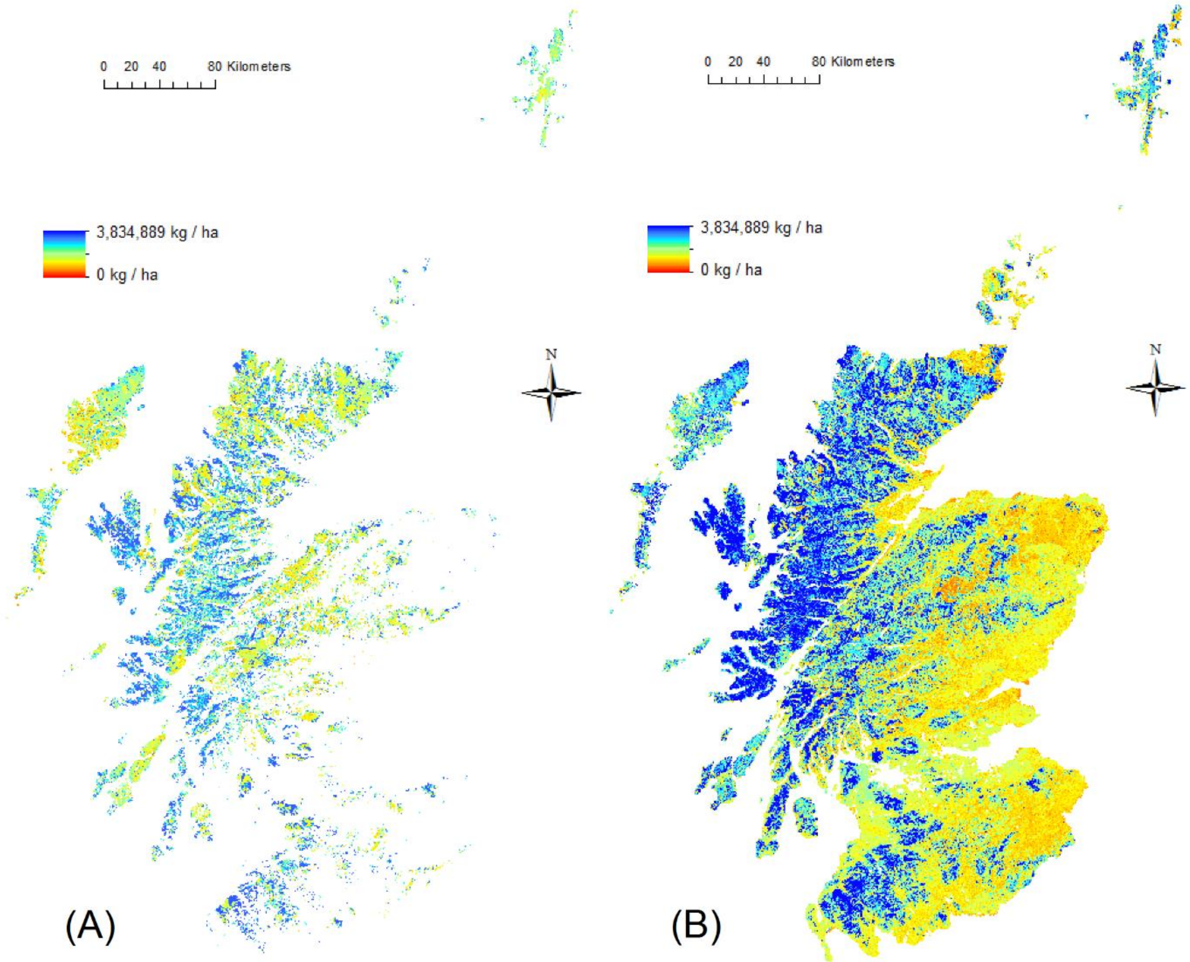
Low C Low BD Low error <1%	Low C Medium BD Low error 1-2%	Low C High BD Moderate error 40-60%
Medium C Low BD High error 1-2%	Medium C Medium BD High error 5-10%	Medium C High BD High error 1-2%
High C Low BD Moderate error 20-30%	High C Medium BD Moderate error 5-10%	High C High BD High error <1%

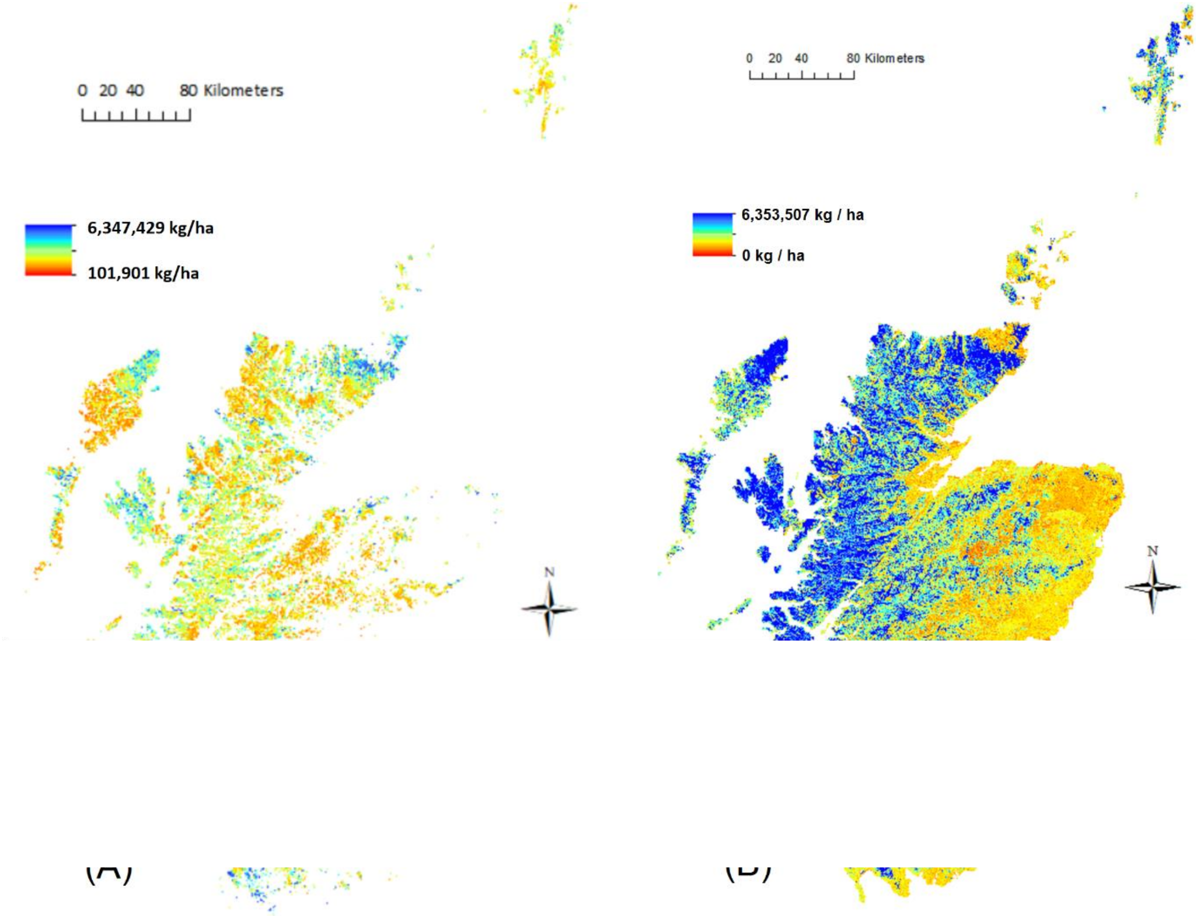

0 20 40 80 Kilometers
416 cm
0 cm
N
(A)
0 20 40 80 Kilometers
862 cm
27 cm
N
(B)

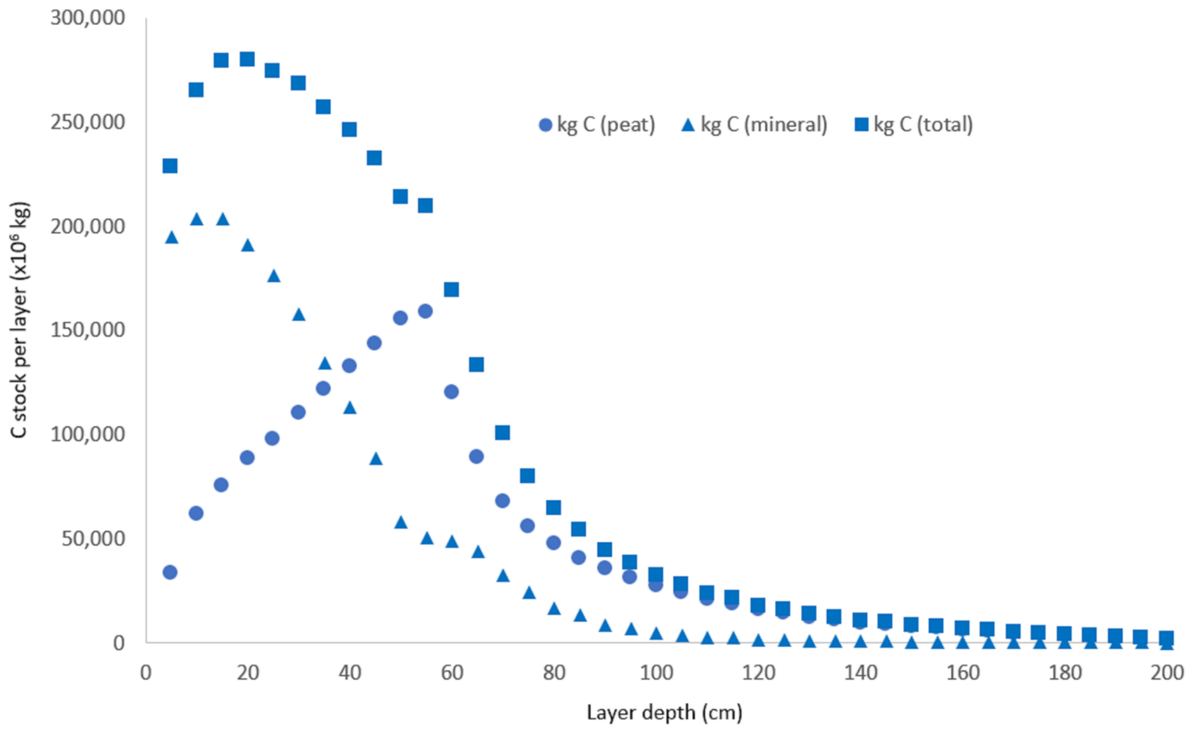
698


0 20 40 80 Kilometers
151127 kg/ha
6928 kg/ha
N
(A)
0 20 40 80 Kilometers
238722 kg/ha
917 kg/ha
N
(B)


0 20 40 80 Kilometers
255240 kg / ha
371 kg / ha
(A)
0 20 40 80 Kilometers
226455 kg/ha
184 kg/ha
N
(B)


0 20 40 80 Kilometers
3,834,889 kg / ha
0 kg / ha
N
(A)
0 20 40 80 Kilometers
3,834,889 kg / ha
0 kg / ha
N
(B)


0 20 40 80 Kilometers
6,347,429 kg/ha
101,901 kg/ha
N
0 20 40 80 Kilometers
6,353,507 kg / ha
0 kg / ha
N


kg C (peat)
kg C (mineral)
kg C (total)
C stock per layer (x10⁶ kg)
Layer depth (cm)
300,000
250,000
200,000
150,000
100,000
50,000
0
0
20
40
60
80
100
120
140
160
180
200