Geoderma 232–234 (2014) 487–499

Contents lists available at ScienceDirect

Geoderma

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

Improving forest soil carbon models using spatial data and geostatistical approaches

Kevin Black [a,b,*], Rachel E. Creamer [c], Georgios Xenakis [d], Sally Cook [b]

[a] Forestry Division, FERS Ltd, 117 East Courtyard, Tullyvale, Cabinteely, Dublin 18, Ireland
[b] School of Environmental Sciences, University of Ulster, Coleraine Campus, Cromore Road, Coleraine, Co. Londonderry BT52 1SA, United Kingdom
[c] Teagasc, Johnstown Castle, Wexford, Ireland
[d] The University of Edinburgh, School of GeoSciences, Crew Building, The King's Buildings West Mains Road, Edinburgh EH9 3JN, United Kingdom

ARTICLE INFO

Article history:
Received 28 March 2014
Received in revised form 26 May 2014
Accepted 27 May 2014
Available online 17 June 2014

Keywords:
YASSO
Spatial auto-correlation
Soils
Forestry
Land use change

ABSTRACT

Forest soils store large amounts of carbon (C), and stock changes in this C pool may significantly increase the CO_2 concentration in the atmosphere. However, estimation of soil organic carbon (SOC) stocks and stock changes following land use transition to forestry is subject to large uncertainty. Many currently used geochemical modelling approaches, such as YASSO, are used to estimate regional changes in forest SOC stocks, but these are difficult to calibrate to reflect regional conditions because of limited availability of sufficient SOC data. In addition, most model frameworks give little consideration regarding the appropriate use of geospatial climatic and topographical data, as dependent variables in the model. As a result, many regional models may exhibit spatial autocorrelation (SAC) of residuals, which contributes to overall model error. In this paper, we develop a method for assessing SOC stock changes in Irish forests by compiling a spatial SOC database and using these data to calibrate and improve on an existing YASSO model. Careful consideration was given to the use of available climatic and digital elevation GIS data in YASSO with the aim of reducing SAC of model residuals and to more precisely predict soil- and site-specific variations in SOC stock changes following transition to forestry.

Analysis of the complied national SOC database shows that stock changes in afforested mineral soils may increase or decrease depending on previous land use and soil type. During refinement of the YASSO model, conventional statistical approaches confirmed that model performance can be improved by using climatic GIS data at the appropriate scale (resolution), together with additional use of novel topographical spatial data. The current YASSO model does not use these topographical factors as dependent variables, nor is there any consideration given to the spatial or temporal resolution of GIS datasets used. Use of GIS geo-statistical approaches to determine if SAC was reduced, as the YASSO model accuracy was improved on, produced conflicting results. We suggest that the use of Anselin Local Moran's I outlier analysis may not be suitable for this purpose because it may falsely detect spatial outliers due to the presence of neighbouring points with very high or low residual values. In contrast, semi-variogram analysis appeared to be the most useful geo-statistical measure of the spatial dependency, distribution and scale at which residual SAC occurs. Use of fine resolution (50 m) slope and topographical position index (TPI) raster datasets to predict forest SOC stocks significantly improved the final YASSO model accuracy and precision. In addition, semi-variogram analysis confirmed that the final YASSO model residuals exhibited no spatial dependency and residual error was uniformly distributed over the entire sample area, from which the SOC database was derived. However, the final YASSO model we describe requires considerable refinement using more intensive sampling studies and independent validation before it can be applied at a national level. In the future, particular emphasis should be directed to sampling forest brown earth soils, which are suggested to result in a net emission of C following transition from grassland to forest land.

1. Introduction

Land use and land use change accounts for 25% of global anthropogenic emissions of carbon dioxide (CO_2) to the atmosphere, second only to emissions from the combustion of fossil fuels (Houghton and Skole, 1990). Soils are the largest terrestrial carbon pool, nearly twice that of the atmosphere. Small changes in the soil carbon gains or losses could therefore have a large influence on atmospheric CO_2 concentration (Schlesinger, 1995). Changes in soil organic carbon (SOC) stocks are dependent on land use management, soil type or structure and geochemical processes influenced by climatic and topographical factors (Don et al., 2011; Houghton and Skole, 1990; Liski et al., 2005). The amount of SOC in a forest ecosystem is a function of the difference

* Corresponding author. Tel.: +353 1 272 2675; fax: +353 1 282 7272.
E-mail address: kevin.g.black@gmail.com (K. Black).

http://dx.doi.org/10.1016/j.geoderma.2014.05.022

between the input of carbon (C) as surface litter, belowground biomass, and root exudates to the soil profile, and the losses due to decomposition, fragmentation, erosion, deposition and leaching. In forest soils, there is a potential for a net accumulation of SOC due to large C inputs of biomass residues and root exudates into the soil profile (Liski et al., 2005). The decomposition rate of C compounds decreases with an increasing chemical complexity. Highly lignified woody necromass eventually forms a recalcitrant C pool which is not subjected to microbial decomposition. In contrast, labile C pools such as simple carbohydrates and proteins are quickly decomposed (Chapin et al., 2002). Although, these bio-geochemical processes are relatively well defined and understood, more precise estimates of SOC changes are required for the reporting of national green-house gas emissions and removals associated with of land use change and forestry to the United Nations Framework Convention on Climate Change (UNFCCC) and the Kyoto protocol. There is currently no available technique for accurately estimating and reporting SOC stock changes associated with land use change to forestry in Ireland (Duffy et al., 2012).

Afforestation of histosols (organic soils e.g. peats) can lead to a significant emission of CO_2 due to drainage (Alm et al., 2007). However, the effect of afforestation on SOC in mineral soils remains unclear. Based on a meta-analysis of 385 studies in the tropics, it is suggested that afforestation may result in an increase of SOC following conversion from crop and grasslands (Don et al., 2011). In contrast, paired plot and chronosequence approaches conducted in Ireland and New Zealand show no significant change in SOC following land use transition to forestry (Black et al., 2009a; Tate et al., 2003; Wellock et al., 2011). Clearly, the direct estimation of SOC stocks and stock changes following land use transitions is subject to large error and methodological bias (Goidts et al., 2009; Tate et al., 2003; Wellock et al., 2011). Therefore, a combination of field sampling, modelling and geo-processing approaches are increasingly being used to refine national green-house gas estimates of changes in SOC (Coleman and Jenkinson, 1996; Liski et al., 2005; Parton et al., 1987; Scott et al., 2002; Tornquist et al., 2009). These modelling frameworks, such as YASSO or CENTURY, incorporate mathematical representations of bio-geochemical and pedogenic processes, which describe changes in SOC behaviour in space and time. To achieve this, GIS data layers, such as soil maps and climate data, are often used in an attempt to provide spatially explicit estimates of SOC at the regional or national scale (Liski et al., 2005; Tornquist et al., 2009). Although it is widely recognised that validation of soil C models is important to provide robust estimates, little consideration is given to evaluation of residual spatial error introduced in the overall model error (Pringle and Lark, 2006). Conventional statistical validation of model performance by correlation of observed and predicted values assumes independence among observations. The problem is that soil survey/sample data may display spatial autocorrelation (SAC), i.e. locations close to each other exhibit more similar values than those further apart. So if model residuals exhibit SAC, then the key assumption that the residuals are independent and similarly distributed is violated (Dormann et al., 2007).

The most common factors causing SAC include a) biological processes such as forest species–soil interactions are distance related; b) nonlinear relationships between an environmental variable and the dependent variable are erroneously modelled as linear relationships and c) the model fails to account for an important environmental determinant, which is spatially structured resulting in spatial dependency in the response to that variable (Besag, 1974; Legendre et al., 2002). The spatial dependency response is important because a key feature of biological processes is that they are the combined effect of different components across different spatial scales[1] (Pringle and Lark, 2006). For example, the CENTURY and YASSO models (Liski et al., 2005; Parton et al., 1987) use temperature to explain variations in the decomposition of soil C inputs at a regional or continental scale. However, models may not consider microclimatic gradients or describe other climatic processes, such as soil moisture deficits, which may influence soil geochemical processes at a stand or localised scale. This means that the use of coarse scale spatial and temporal climatic data may not account for soil type- or site-specific variations in SOC accumulation of decomposition processes. Clearly, if these finer scale interactions have not been defined in the model to reflect the site-specific SOC spatial dependent responses to climate or topography, then there is a higher risk of SAC in the model residual.

A variety of geo-statistical techniques have been developed to identify and correct for the effects of SAC (reviewed by Dormann et al., 2007). A common approach used to treat SAC is progressive inclusion of spatial factors in the model, such as environmental covariates, thereby reducing residual SAC (Higgins et al., 1999; Warren et al., 2005). However, if the scale of measurement (spatial resolution) of the environmental variable is to coarse to describe changes in SOC over the entire range (climatic or geophysical) from which calibration data is being derived, then there is a risk in increasing residual SAC (Dormann, 2007). This poses a significant challenge since many climatic datasets are measured at a coarse scale. Therefore, interpolation techniques such as cokriging (Rawlins et al., 2009) or weighted regression (Fotheringham et al., 2002) are employed to derive spatially representative environmental covariate values. This in turn introduces inherent SAC problems (Dormann et al., 2007).

The primary aim of this paper was to develop a method for assessing SOC stock changes in Irish forests by improving on an existing YASSO model. The YASSO model was calibrated using a compiled SOC database, which was compiled using nationally available soil survey data. Various GIS downscaling techniques were applied using available coarse resolution climatic and digital elevation data to derive environmental variables, which could be incorporated into the YASSO model to more precisely predict soil- and site-specific variations in SOC stock changes following transition to forestry. It is hypothesised that improved model predictions may be obtained by either increasing the resolution of GIS data sets used, or by including additional spatial variables, to reduce SAC of model residuals. To test this hypothesis, numerous geostatistical techniques (e.g. outlier analysis, Moran's correlogram and semi-variogram models) were used to determine if the use of refined environmental variables can reduce SAC of model residuals whilst still improving on the accuracy of the YASSO model.

2. Material and methods

2.1. Study area and SOC database

The Republic of Ireland has a total land area of 71,112 km^2, of which ca. 10% is covered by forests (Duffy et al., 2012). Grasslands are the most dominant land use in Ireland (58.2% of the total area in 1990), followed by wetlands (17.2%, Duffy et al., 2012). Approximately 250 kha grasslands and wetlands have been converted to forest since 1990 (Black et al., 2009b, 2012). Major soil types in Ireland can be categorised in broad groups, namely Podzols, Brown Podzolics, Grey Brown Podzolics, Acid Brown Earths, Gleys, Brown Earths, Rendzinas, Lithosols, Alluvial soils and Histosols (Gardiner and Radford, 1980).

The national data used to calibrate models were compiled using available SOC survey data sampled from mineral soils representing major land use categories in the republic of Ireland. The mineral soil database comprises of a total of 227 sample sites (Table 1) obtained from the Soil C project (Wellock et al., 2011), the Irish national soil database of Ireland (NCD, see Xu et al., 2011), the An Foras Taluntas project (Creamer, R., unpublished data), and the CARBiFOR project (Black et al., 2009a). The definition used for mineral soils includes all soils with a carbon concentration of less than 30% (NFI, 2007). This excludes peat lands (histosols) and organo-mineral soils, such as peaty-gleys and

[1] As defined by Pringle and Lark (2006): "Components of spatial variation at a low frequency are of a 'coarse scale' relative to 'finer-scale' components at a higher frequency (i.e. shorter range)."