Geoderma 290 (2017) 91–99

Contents lists available at ScienceDirect

Geoderma

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

Digital soil mapping of soil carbon at the farm scale: A spatial downscaling approach in consideration of measured and uncertain data

Brendan P. Malone *, Quentin Styc, Budiman Minasny, Alex B. McBratney

Sydney Institute of Agriculture, C81 Biomedical Building, The University of Sydney, New South Wales 2006, Australia

ARTICLE INFO

Article history:
Received 24 July 2016
Received in revised form 10 November 2016
Accepted 9 December 2016
Available online 2 January 2017

Keywords:
Digital soil mapping
Spatial downscaling
Soil carbon auditing
Soil carbon
Quantile regression forest
Uncertainty

ABSTRACT

In this paper a spatial downscaling method is explored for generating appropriate farm scale digital soil maps. The digital soil map product to be downscaled is an Australian national extent soil carbon map (100 m grid resolution). Taking into account the associated prediction uncertainties of this map, we used a simulation approach based on Gaussian random fields to generate plausible mapping realisations that were in turn downscaled to 10 m resolution for a farm in North-western NSW, Australia. We were able to derive both a downscaled map of soil carbon and associated prediction variance with this approach. Building further upon this development, we then incorporated a bias correction step into the spatial downscaling procedure which permits the inclusion of field observations as a way to moderate the downscaling results to better reflect actual conditions on the ground. Based on an independent validation dataset, it was found that incorporating field observations increase the concordance correlation coefficient to 0.8 from 0.2. This relatively lower correlation achieved using spatial downscaling alone was due to the national scale mapping for the study area being positively biased in the area of interest. It was found that downscaling that incorporates observational data was marginally better if not comparable to using a point-based digital soil mapping approach. The advantage of spatial downscaling is that it can be implemented in situations of data scarcity. This will be ideal for on farm soil monitoring in situations where detailed soil mapping is initially not available. For example, soil carbon auditing schemes requiring prior soil information for implementation of design-based soil sampling could potentially be universally applied with such a spatial downscaling approach.

1. Introduction

Contextualizing the sampling theory established by De Gruijter et al. (2015), follow up work by De Gruijter et al. (2016) proposed an efficient and optimizable soil sampling protocol for the unbiased estimation of on farm soil carbon stocks. Their interest was the estimation of whole farm soil carbon stocks with sufficient statistical confidence. Such information is necessary for the establishment and ongoing monitoring of soil carbon. Unbiased estimation of on-farm carbon stocks is also necessary in the broader context of carbon inventory and participation of agriculture sector in the carbon economy (Antle et al., 2003). Coincidentally, in addition to the environmental benefits associated with sequestration of carbon into soils, there are foreseeable economic benefits for farming communities too (Stockmann et al., 2013).

A design-based soil sampling approach, the *ospats* algorithm from De Gruijter et al. (2015) enables one to use prior information by way of existing soil carbon mapping (and associated prediction variances) to derive an optimal number and spatial configuration of strata, and ultimately an optimal number of samples to collect from a farm. The focus of this particular research is in regards to the prior information that is required by *ospats* – that is, the mapping of soil carbon and associated uncertainties. For universality of application, *ospats* needs relevant farm scale digital soil map of carbon stock and associated prediction uncertainties. With the exception of some farms, most agricultural landholdings will not likely have an established digital spatial information system. It is proposed in this research that such information may be obtained more-or-less globally by exploiting the availability of global and/or national digital soil mapping products.

Throughout the world there has been an upsurge of digital soil mapping projects (Minasny and McBratney, 2016). This is due mainly to enabling technologies in quantitative methodologies and geographic information systems, in addition to a global need of relevant spatial soil information systems to address critical environment issues of which soil is manifold. The vanguard of such projects has been the GlobalSoilMap project (Sanchez et al., 2009; Arrouays et al., 2014), which set as the ambitious goal to use digital soil mapping to map key soil attributes at 100 m spatial resolution and specified depth intervals

* Corresponding author.
E-mail addresses: brendan.malone@sydney.edu.au (B.P. Malone), quentin.styc@gmail.com (Q. Styc), budiman.minasny@sydney.edu.au (B. Minasny), alex.mcbratney@sydney.edu.au (A.B. McBratney).

http://dx.doi.org/10.1016/j.geoderma.2016.12.008

to 2 m (0–5 cm, 5–15 cm, 15–30 cm, 30–60 cm, 60–100 cm, and 100–200 cm) across the entire ice-free land surface of the world. Working in parallel with that project or inspired from some of its methodological approaches, similar very large extent digital soil mapping projects have also resulted throughout the world.

Such large spatial extent digital soil mapping products are invaluable for aiding the decision-making process at the spatial scales they were intended for. However, they are not particularly relevant for considering issues at the farm scale. In other words, the observed spatial variability at the farm scale is not sufficiently captured in nationally or regionally calibrated models and the resultant digital soil maps. This can often be simply a matter of grid cell resolution being too coarse for meaningful on farm analyses.

In order to create digital soil maps relevant to the farm scale in the absence of sufficient data, one possible option to consider is spatial downscaling. This means the spatial disaggregation of the national mapping using a statistical model and a library of environmental covariates that will help in defining the spatial variability of the target variable. The implicit assumption here is that the covariate information is strongly related to the target variable, which is being derived at the fine scaled resolution. This general statistical downscaling approach is embodied is the *dissever* algorithm (Malone et al., 2012) which itself is a generalization of the linear downscaling algorithm proposed by Liu and Pu (2008). Poggio and Gimona (2015) used a modified *dissever* approach which considers a correction step when downscaling climate model outputs.

The *dissever* algorithm was originally parameterized for inclusion of the uncertainties of the map to be downscaled. The underlying model within *dissever* is a weighted generalized additive model (Hastie and Tibshirani, 1990). In addition to allowing one to investigate non-linear relationship between target variable and covariates, in this model, the uncertainties of the input data (the coarse map) are used as weighting factors (inverse weighted) in the nonlinear fitting function (Malone et al., 2012). Despite this generalization, incorporating the prediction uncertainties did not ultimately mean they were propagated through to the downscaled outputs directly. In fact, the only measure of uncertainty associated with downscaling from Malone et al. (2012) was that associated with the deviation from mass balance between coarse scale mapping and associated downscaled mapping. The quantification of uncertainties is necessary for obvious reasons of assessing the reliability of mapping. Importantly for soil carbon stock auditing and using *ospats* specifically, the magnitude of the prediction uncertainties determines the spatial configuration of the sampling strata and optimal sample number.

Subsequently, this research is focused on the delivery of relevant farm scale digital soil mapping of carbon via a spatial downscaling approach. A key question of this research is how to efficiently take into account the prediction uncertainties of the national mapping so that they are in turn propagated through to the downscaled mapping in addition to the uncertainty estimated from downscaling. Exploring further the work of Poggio and Gimona (2015) a second research question is, to what effect does incorporating field observations into the downscaling process while simultaneously taking into consideration the uncertainties of the national mapping? For the first question we may hypothesize that by explicitly taking into account the prediction uncertainties within the modeling process, we may expect an associated prediction variance with the downscaled mapping, which would be an ideal outcome especially for universal implementation of *ospats*. Regarding the second question, from results obtained by Poggio and Gimona (2015) we would expect the bias corrected downscaled mapping will reflect the more present-day spatial pattern of soil carbon variation, and have relatively lower uncertainty than the mapping created from downscaling alone.

2. Material and methods

2.1. Methodological overview

First we describe the study area that is under investigation, and then outline the various data that has been collected and subsequently used. We then detail a simple approach for generating plausible realisations from national scale mapping conditioned to the spatial and statistical properties of these mapping. This is followed by description of a spatial downscaling approach of the simulated national scale mapping outputs for generating relevant farm scale predictions and associated uncertainty. The spatial downscaling entails approaches pertaining to with and without correction based on the usage of point observation data. For comparisons, we compare downscaled products with those derived from a point-based digital soil mapping approach. Validation of all outputs in this investigation is performed using an independent data set from the area under investigation.

2.2. Study area and data acquisition

The farm under investigation in this study is the University of Sydney owned and managed E.J. Holtsbaum Research Station, "Nowley" (31.35°S 150.11°E). Situated in the highly agriculturally productive Liverpool Plains region in north west NSW (Fig. 1), Nowley (approximately 2300 ha) is run as a mixed farming enterprise centered around crops of wheat, barley and canola in winter, sorghum and sunflower in summer, and a cattle herd of breeders, replacement heifers and bulls. Nowley has a combination of fertile basaltic soils together with more challenging to manage soil types that are poorly drained and with considerably high amounts of subsoil sodium. A more comprehensive description of the region and Nowley farm can be found at (Stockmann et al., 2016).

At the present time, national mapping of soil carbon (that is publically available and downloadable) pertains to total soil carbon concentration. This data is available via the repository of the Australian Soil and Landscape Grid (http://www.clw.csiro.au/aclep/soilandlandscapegrid/). Technical information regarding the Australian Soil and Landscape grid can be found in Grundy et al. (2015). This national scale digital soil mapping is based on the GlobalSoilMap specification resolved to a 100 m grid cell resolution, and is available as layers corresponding to depth intervals of: 0–5 cm, 5–15 cm, 15–30 cm, 30–60 cm, and 60–100 cm. Using a boundary extent of the Nowley Farm we clipped the national mapping which included the lower and upper bounds of a 90% prediction interval, and the predicted values as generated by *scorpan* modeling. In this study, because soil sampling was based on a 0–7.5 cm depth interval, we extracted only the mapping corresponding to 0–5 cm and 5–15 cm. With the extracted mapping we set about deriving mapping that corresponded to 0–7.5 cm which is based on collected topsoil samples. This was facilitated using the mass preserving spline depth function described in Bishop et al. (1999) for each 100 m grid cell. This procedure was done for the predictions and the associated lower and upper prediction intervals. After this, using the 90% lower and upper prediction limits we estimated the variance and standard deviation of the predictions the same way as in Malone et al. (2014).

Data collected from Nowley included a number of environmental layers related to topography and gamma radiation. The topographic data were collected (and later mapped) by Tranter (2005) by ground survey using an all-terrain vehicle with attached Ashtech Real-Time Kinetic Global Positioning System (RTK GPS). Driven across the farm in 20 m swaths, positional data with coupled elevation data were recorded every 3 s. After the completion of the survey, a series of post-processing steps were performed followed by interpolation to create a map of elevation. Interpolation was performed via local ordinary kriging onto a regularly spaced 10 m grid across the whole farm. Further processing of the 10 m elevation model entailed calculation of a number of terrain derivatives. In this particular study the following terrain derivatives were used: slope gradient, terrain wetness index, and multi-resolution valley bottom flatness index. High resolution gamma radiometric mapping was collected via aerial survey. Australia has a continental coverage of remotely-sensed radiometric data at 100 m grid resolution, but the raw data that contributes to this mapping is composed of a patchwork of aerial surveys that range in information content (Minty et al., 2009). In the area where this study is based, the information content