Environmental Modelling & Software 80 (2016) 26–32

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Confidence in soil carbon predictions undermined by the uncertainties in observations and model parameterisation

Zhongkui Luo [a,*], Enli Wang [a,**], Quanxi Shao [b], Mark K. Conyers [c], De Li Liu [c]

[a] CSIRO Agriculture, GPO Box 1666, Canberra, ACT 2601, Australia
[b] CSIRO Digital Productivity & Services, Private Bag 5, Wembley, WA 6913, Australia
[c] NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, PMB, Wagga Wagga, NSW 2650, Australia

ARTICLE INFO

Article history:
Received 17 March 2015
Received in revised form
8 December 2015
Accepted 8 February 2016
Available online 23 February 2016

Keywords:
Carbon cycle
Carbon sequestration
Measurement uncertainty
Model optimisation
Prediction uncertainty

ABSTRACT

Soil carbon (C) responds quickly and feedbacks significantly to environmental changes such as climate warming and agricultural management. Soil C modelling is the only reasonable approach available for predicting soil C dynamics under future conditions of environmental changes, and soil C models are usually constrained by the average of observations. However, model constraining is sensitive to the observed data, and the consequence of using observed averages on C predictions has rarely been studied. Using long-term soil organic C datasets from an agricultural field experiment, we constrained a process-based model using the average of observations or by taking into account the variation in observations to predict soil C dynamics. We found that uncertainties in soil C predictions were masked if ignoring the uncertainties in observations (i.e., using the average of observations to constrain model), if uncertainties in model parameterisation were not explicitly quantified. However, if uncertainties in model parameterisation had been considered, further considering uncertainties in observations had negligible effect on uncertainties in SOC predictions. The results suggest that uncertainties induced by model parameterisation are larger than that induced by observations. Precise observations representing the real spatial pattern of SOC at the studied domain, and model structure improvement and constrained space of parameters will benefit reducing uncertainties in soil C predictions. The results also highlight some areas on which future C model development and software implementations should focus to reliably infer soil C dynamics.

1. Introduction

Integrated approaches such as process-based mechanistic models are increasingly used to systematically explain, explore and predict responses of natural and human-managed systems to environmental changes (Holzworth et al., 2015; Laniak et al., 2013). Under changing environmental conditions such as global warming and land management change, soil carbon (C) dynamics are particularly important as their core role in determining the C and nutrient cycling in terrestrial ecosystems and the relevant environmental footprints. Soil C models are the only reasonable approach available for assessing soil C dynamics across spatio-temporal scales and for understanding mechanisms underpinning soil C stability under various environments (Friedlingstein et al., 2006; Thornton et al., 2007). In order to use these models to design effective management strategies, it is vital to understand the confidence level in model predictions (Bennett et al., 2013, Kelly (Letcher) et al., 2013). A number of studies have recognised that the reliability of model predictions must be carefully assessed due to uncertainties in model inputs, model parameters and structure, and scaling of model outputs etc (Clifford et al., 2014; He et al., 2014; Luo et al., 2013; Ogle et al., 2010; Post et al., 2008; Xia et al., 2013). An additional source of uncertainty comes from the spatial variability in soil C measurements that are used to constrain soil C models, and has not been explicitly quantified in current soil C modelling.

Most soil C models divide soil C into several conceptual pools with different decomposability, and simulate decomposition of each pool by first-order kinetics (Smith et al., 1997). Some of the pools cannot be directly measured and there is no agreed process to initialise these pools with bulked measurements of total soil C. As a result, derivation of the decomposition rate of those pools is constrained by local observed data of total SOC. The common practice is

* Corresponding author.
** Corresponding author.
E-mail addresses: zhongkui.luo@csiro.au, zhongkui.luo@gmail.com (Z. Luo), Enli.Wang@csiro.au (E. Wang).

http://dx.doi.org/10.1016/j.envsoft.2016.02.013
1364-8152/

to derive a single set of model parameters that enables a good match between simulated soil C and the average of measured soil C (based on different replicates) at a given time. Such constrained models are then used to predict possible soil C changes under different management practices across environments (Grace et al., 2006; Li et al., 2003; Lugato et al., 2014; Todd-Brown et al., 2014).

Such predictions are subject to unknown uncertainties. Firstly, the uncertainty in initialisation of soil C pools interacts with the uncertainty in derived decomposition rates. For example, similar model performance can be achieved either by optimising decomposition rates or by partitioning total soil C to different conceptual pools with different decomposition rates. Secondly, inaccuracy and/or variability in bulked measurements of total soil C further impact on the initialisation and parameterisation of the model. As a result, model initialisation and parameterisation become very sensitive to the data that are used for model constraining (Hararuk et al., 2014; Juston et al., 2010; Keenan et al., 2012; Xenakis et al., 2008; Ogle et al., 2010; Weng and Luo, 2011). Even at the field scale, great spatial variability exists in soil properties including soil C (Cambardella et al., 1994). For example, Zhou et al. (2008) demonstrated that soil C varied strongly spatially and its spatial autocorrelation only occured in ~2 m of distance in a grazed grassland. Todd-Brown et al. (2013) suggested that uncertainty in bulked measurements of soil C content must be integrated with errors involved in extrapolating the data from individual soil profiles to the regional scale. This spatial variability makes it difficult, if not impossible, to derive the real estimate of average soil C for model initialisation and/or parameterisation based on limited samples.

To date, however, most of studies focused on the uncertainty in model inputs in terms of limited availability of a specific data source (e.g., edaphic characteristics at higher resolution, and information on land use and management) that is needed to initialise and/or parameterise the model (Falloon et al., 2011; Luo et al., 2013). Some other studies also addressed the effect of data derived from different environments on model parameterisation thereby model outputs (Hararuk et al., 2014; Juston et al., 2010). In this study, we used a process-based biophysical model, the Agriculture Production Systems sIMulator APSIM (Holzworth et al., 2014; Keating et al., 2003), together with a long-term soil organic C observational (measured 20 times during a 25-year experiment) dataset, to quantify: 1) the potential uncertainty in soil C predictions caused by model initialisation and parameterisation, and 2) the additional uncertainty caused by the variation in soil C measurements that are used for model initialisation and parameterisation. For the latter, the variation in replicates for each of the 20 observations during a 25-year agricultural experiment were investigated when using the data to constrain the model. The results can provide insights into the collection of effective data sets for model constraining and development of next generation models.

2. Materials and methods

2.1. Study site and data source

The field experimental data collected by the Wagga Wagga Agricultural Institute of NSW Department of Primary Industries was used in this study. The experimental site was located at the Wagga Wagga, New South Wales, Australia (35.11°S, 147.37°E). It has a temperate climate with uniform rainfall distribution across the year. Mean annual temperature was 15.9 °C and mean annual rainfall was 538 mm. The soil is a chromic luvisol, and the site was maintained as an annual pasture for 19 years from 1960 to 1978, except for a crop of lupins (*Lupinus Angustifolius*) in 1975 and oats (*Avena Sativa*) in 1976. The surface 10 cm soil was a clay loam with 29% clay, 15% silt and pH 4.9 in 1979. Data collected under two treatments, continuous wheat with (100 kg N ha^{-1} yr^{-1}, N100) and without nitrogen (N0) fertilizer application, were used for this study. In both treatments, crop residues were burned, and the soil organic C content (%) in the 0–10 cm soil layer was observed 20 times for each treatment from 1979 to 2004, each time with five or up to 12 replicates for each observation (Fig. 1). Soil bulk density was also measured along with the measurement of soil C content, and they were used to calculate soil C stock (t ha^{-1}). More detailed information on experimental design, soil sampling strategy, land use, soil conditions, management and observations for this experiment can be found in Heenan et al. (2004, 1995).

2.2. The APSIM model

The Agricultural Production Systems sIMulator APSIM (Holzworth et al., 2014; Keating et al., 2003) was used to simulate the observed soil C dynamics in the two treatments (N0 vs N100). The APSIM model is a process-based bio-physical model designed to study productivity, nutrient cycling and environmental impacts of farming systems as influenced by climate variability and management interventions. The ability of APSIM to simulate soil C and soil nitrogen (N) dynamics has been verified under various cropping systems and agricultural management (Luo et al., 2011; Probert et al., 1998). The model divides soil C into six pools and simulates each pool as a first-order process with the rate constants being modified by factors involving soil temperature, moisture and nutrient availability in the soil layer, which is similar to other widely used soil C models such as Century (Parton et al., 1987) and RothC (Jenkinson, 1990). A detailed conceptual diagram of the model for simulating soil C dynamics is presented by Probert et al. (1998) and Luo et al. (2014).

The APSIM runs on a daily time-step and needs daily weather data as inputs, including radiation, maximum and minimum temperatures, and rainfall. Other required soil parameters includes soil C content, C to N ratio of the bulk soil, soil bulk density, hydraulic parameters, initial soil water and nutrient conditions (NO_3^- and NH_4^+) in each soil layer. All these data and the relevant model initialisation and parameterisation processes were adopted from two former APSIM simulation studies using the same dataset by Luo et al. (2014, 2011).

2.3. Constraining the APSIM model

The model was constrained with observed soil C data from the N0 and N100 treatments. Model parameters were derived through constraining model simulations against the observed data from both treatments. Two parameters were targeted: the potential decomposition rate of humic C pool (*rdhum*) and the amount of recalcitrant C pool in total soil C (*finert*). The two parameters directly control the turnover time and decomposability of soil C. Sensitivity analysis of eight main parameters that directly link to C decomposition has indicated that these two parameters were the two most important parameters to which soil C dynamics are most sensitive in the APSIM model (Luo et al., 2015).

Three optimisation strategies were used to optimise the model by minimising the combined rooted mean of squared errors (RMSE, i.e., the objective function for *optimise* and *DEoptim*, see below) between simulated and observed soil C under the two treatments. That is, 1) optimise *finert* only (Opt1), 2) optimise *rdhum* only (Opt2), and 3) optimise *rdhum* and *finert* synchronously (Opt3). We used the similar Bayesian approach of Yeluripati et al. (2009) to derive the posterior distributions of the two parameters (*finert* and/or *rdhum*). Both parameters were bounded within a range that is biologically and physically possible, therefore eliminating solutions in conflict with prior knowledge. For *rdhum,* we assumed that