Simulation of soil carbon dynamics in Australia with ROTH C

Juhwan Lee[1], Raphael A. Viscarra Rossel[1], Zhongkui Luo[2], and Ying Ping Wang[3]

[1]School of Molecular and Life Sciences, Curtin University, GPO Box U1987, Perth WA 6845, Australia.
[2]College of Environmental and Resource Sciences, Zhejiang University, Hangzhou, Zhejiang, China.
[3]CSIRO Oceans and Atmosphere, Private Bag 1, Aspendale, VIC 3195, Australia.

Correspondence: Raphael A. Viscarra Rossel (r.viscarra-rossel@curtin.edu.au)

Abstract. We simulated soil organic carbon (C) dynamics across Australia with the Rothamsted carbon model (ROTH C) by connecting new spatially-explicit soil measurements and data with the model. This helped us to bridge the disconnection that exists between datasets used to inform the model and the processes that it depicts. Under this framework, we compiled publicly available continental-scale datasets and pre-processed, standardised and configured them to the required spatial and temporal resolutions. We then calibrated ROTH C and run simulations to estimate the baseline soil organic C stocks and composition in the 0–0.3 m layer at 4,043 sites in cropping, modified grazing, native grazing, and natural environments across Australia. We used data on the C fractions, the particulate, humus, and resistant organic C (POC, HOC and ROC, respectively) to represent the three main C pools in the ROTH C model's structure. The model explained 97–98% of the variation in measured total organic C in soils under cropping and grazing, and 65% in soils under natural environments. We optimised the model at each site and experimented with different amounts of C inputs to simulate the potential for C accumulation under constant climate in a 100-year simulation. With an annual increase of 1 Mg C ha^{-1} in C inputs, the model simulated a potential soil C increase of 13.58 (interquartile range 12.19–15.80), 14.21 (12.38–16.03), and 15.57 (12.07–17.82) Mg C ha^{-1} under cropping, modified grazing and native grazing, and 3.52 (3.15–4.09) Mg C ha^{-1} under natural environments. Soils under native grazing were the most potentially vulnerable to C decomposition and loss, while soils under natural environments were the least vulnerable. An empirical assessment of the controls on the C change showed that climate, pH, total N, the C:N ratio, and cropping were the most important controls on POC change. Clay content and climate were dominant controls on HOC change. Consistent and explicit soil organic C simulations improve confidence in the model's estimations, contributing to the development of sustainable soil management under global change.

1 Introduction

Soil carbon (C) represents the most abundant terrestrial C pool (Batjes, 1996). It can be a significant source or sink of atmospheric CO_2 (Scharlemann et al., 2014). Sequestration of soil organic C, via the adoption of improved land management strategies, offers opportunities for improving soil and ecosystem health, sustainable food production and climate change mitigation (Lal, 2016; Paustian et al., 2019; Smith et al., 2020). However, a better understanding of soil organic C dynamics is needed to determine the size of the soil C pool accurately and to assess the potential for those opportunities.

Biogeochemical models represent our mechanistic understanding of processes such as organic C cycling in soil and can serve different purposes. They can be used to simulate soil C cycling under various combinations of soil, climate, land use and management (Conant et al., 2011), to evaluate the potential for C sequestration or loss, and to assess the impacts of environmental and human-induced change on the C cycle. In conjunction with long-term measurements, models can estimate the effects of management practices and climate change on soil C, as well as subsequent feedbacks. Therefore, the simulation of soil organic C with biogeochemical models has received much attention in the literature (Campbell and Paustian, 2015; Falloon and Smith, 2000).

The Rothamsted carbon model (ROTH C) (Jenkinson, 1990; Coleman and Jenkinson, 1996) and the CENTURY model (Parton et al., 1987) are most commonly used to simulate soil organic C dynamics in cropping, grassland and forest systems. Although developed under northern hemisphere conditions, since their inception in the 1980s, these models have been used for many different applications worldwide (Campbell and Paustian, 2015; Wang et al., 2016). They are the soil biogeochemical component in Earth systems models (Todd-Brown et al., 2013). They do not explicitly represent current theories around the mechanisms of microbial decomposition and physicochemical protection (Lehmann and Kleber, 2015), but they are still being used because they capture the general principle of soil organic C dynamics. In essence, the flow of C in the models occur through a cascading of C via several conceptual pools turning over at different rates, according to first-order kinetics and modified by climate and soil texture. Other reasons for their continued popularity might be that there is ample documentation on them; they are relatively simple and general and are therefore also well understood.

The ROTH C model has been adjusted and tested for use under Australian conditions (Janik et al., 2002; Skjemstad et al., 2004). Skjemstad et al. (2004) showed that the size of the main conceptual C pools in ROTH C, the resistant plant material, humic and inert organic matter pools, can be initialised with measurements of the particulate, humus and resistant organic C (POC, HOC and ROC, respectively) fractions. ROTH C is one of a few models that can be initialised with measured C fractions. Skjemstad et al. (2004) calibrated the decomposition rate constants under Australian conditions, and Janik et al. (2002) assessed a sensitivity of the C pools to model parameters to highlight the potential complexity in the implementation of ROTH C. Since then, researchers in Australia have used ROTH C in different research (Paul and Polglase, 2004; Hoyle et al., 2013; Lee and Viscarra Rossel, 2020). ROTH C is a sub-model of the Fully Integrated Carbon Accounting Model (FullCAM) (Richards and Evans, 2004), used in Australia's National Greenhouse Gas Inventory System. Together, they are the core of the Australian model-based Emission Reduction Fund (ERF) methodology, which allows farmers and landholders to generate extra income by storing C in their soils and thereby reducing emissions (England and Viscarra Rossel, 2018; Paustian et al., 2019). However, the soil and environmental conditions required to maintain current soil organic C stocks and composition are poorly understood. This hampers the reliable estimation of C stock and sequestration potentials at a large scale.

Simulation of soil C dynamics with biogeochemical models can be challenging. Despite the development of new models with updated representations of current understanding (Abramoff et al., 2018; Robertson et al., 2019; Wieder et al., 2014), there remains a disconnection between measurements and datasets used to inform the model and the theories represented in them (Blankinship et al., 2018; Harden et al., 2018). In practice, a lack of data restricts model parameterisation and optimisation and missing temporal datasets limit our ability to simulate and verify long-term changes in soil C stocks and composition (Smith

et al., 1997). Hence, there is also little agreement on how input datasets should be synthesised, processed and used (Manzoni and Porporato, 2009), leading to inconsistent model calibrations (Conant et al., 2011; Seidel et al., 2018) and inaccurate model estimations (Shi et al., 2018). In this context, the development of frameworks for soil organic C modelling and simulation, to synthesise and integrate measurements and datasets with models are critical (Harden et al., 2018; Ogle et al., 2010; Paustian et al., 1997; Smith et al., 2020). The frameworks should allow for their efficient updating, with new measurements, data and models, as they become available (England and Viscarra Rossel, 2018; Smith et al., 2020). They should also enable a more systematic approach for calibration and validation, making simulations more reliable and reproducible.

Here, we report on simulations of the organic C stocks in Australian soils with ROTH C using a standardised approach that syntheses and processes measurements and data for prediction at a correct scale. Our motivation for developing this research is to help answer questions around soil C that are pertinent to Australian soils and ecosystems under different different land uses and management. Our aims are to: (i) derive baseline estimates of soil organic C stock and composition by site-specifically initialising the model with measurements of POC, HOC and ROC and an optimised ratio of decomposable plant material (DPM) to resistant plant material (RPM), which represents the decomposability of incoming biomass, (ii) simulate over a 100-year period, with constant climate and a plausible range of C inputs, the potential to increase organic C stocks as well as the potential vulnerability to C loss across Australia, and (iii) to identify the soil and environmental controls of the change in soil C stocks.

2 Materials and methods

2.1 The Rothamsted carbon model (ROTH C)

ROTH C is a soil process model for the turnover of organic C in non-flooded soils (Jenkinson, 1990; Coleman and Jenkinson, 1996). The model partitions total organic C (TOC) into pools that represent decomposable plant material (DPM), resistant plant material (RPM), microbial biomass (BIO), humified organic matter (HUM), and inert organic matter (IOM) (Coleman and Jenkinson, 1996). The model simulates on a monthly time step changes in its active pools, in response to climate, soil type, land use and management. Annual C inputs from crops and manure represent different land use and management regimes. We used the ROTH C model version 26.3, which is the version that was re-calibrated for a range of Australian soils (Skjemstad et al., 2004). The decomposition rate constants for the DPM, RPM, BIO, and HUM pools are 10, 0.3, 0.15, and 0.02 year^{-1}, respectively (Skjemstad et al., 2004). Its reference state for the decomposition rate constants was reported by Jenkinson and Rayner (1977). The decomposition of each active pool is assumed to increase, following first-order kinetics, with air temperature, but reduced by soil water deficits and the presence of vegetated soil cover. Temperature effects on soil organic matter decomposition increase following a sigmoid function, while the topsoil moisture deficit reduces it by a factor of 0.2 to 1 (no moisture stress). The soil cover factor is 1.0 for bare soil and 0.6 when soil is vegetated, to slow organic matter decomposition. The main conceptual pools RPM, HUM and IOM are replaced with the measured particulate, humus and resistant organic C fractions (POC, HOC and ROC, respectively) (Skjemstad et al., 2004). The POC fraction includes any DPM available in the soil at the time of measurement. The BIO pool was initially set to zero (Sparling, 1992).

2.2 Standardised soil C simulations

We simulated C dynamics across Australia (Figure 1), using a framework that enabled us to efficiently standardise and then integrate measurements and publicly available data on soil properties and environmental controls with the ROTH C model for simulation at the appropriate scale. The approach encompasses five stages as follows (Figure 1): 1) data compilation and synthesis, 2) data pre-processing and standardisation, 3) configuration of data on management regimes, and 4) model simulation supported by consistent initialisation and verification and 5) prediction.
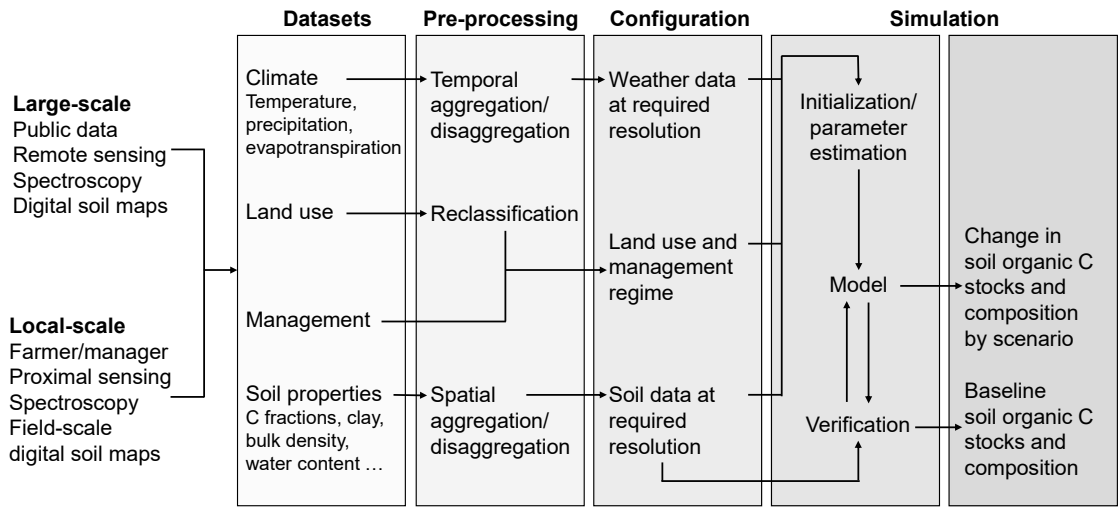
Figure 1. Soil carbon (C) simulation under a framework enables explicit standardisation and better connection between datasets and a soil process model at the appropriate scale.

2.3 Soil C simulations

2.3.1 Data compilation and synthesis

ROTH C requires site coordinates, POC, HOC, ROC, clay content, and sampling depth (in our case 0–0.3 m). The available water capacity (AWC) of the soil to a soil depth of 1 m is needed to modify evapotranspiration from pan evaporation when a plant is present and to run a crop model (see below). We selected a total of 4,431 out of 5,721 sites across Australia (Viscarra Rossel

et al., 2019) (Figure 2). The selected sites were under the dominant land use, namely cropping, grazing of modified pastures and native vegetation, and natural conservation and protected areas. Native forests and production forestry were excluded because of a lack of simulation capacity. The C fractions, clay content, and AWC were estimated with visible–near infrared spectra (Viscarra Rossel and Webster, 2012; Viscarra Rossel et al., 2015). Maximum air temperature, minimum air temperature, precipitation and pan evaporation are also required to run the model. We obtained gridded daily climate data (approximately 5-km resolution) from the SILO database of Australian climate data (SILO, 2020). We used the Australian Bureau of Agricultural and Resource Economics and Sciences land use map (ABARES, 2016) to determine detailed land cover across Australia. Agricultural activity data from Unkovich et al. (2017) provided data for croplands and modified pastures at Statistical Area Level 2 (SA2) (ABS, 2016), which are functional areas that represent socially and economically coherent communities. Current and historical events and agricultural practices, such as crop type and harvest, can be specified from 1970 to 2014. The other data required to run the model include an estimate of the decomposability of incoming biomass, soil cover, and monthly inputs of plant C and farmyard manure. These C input variables, if not measured, must be estimated at each site (see below).
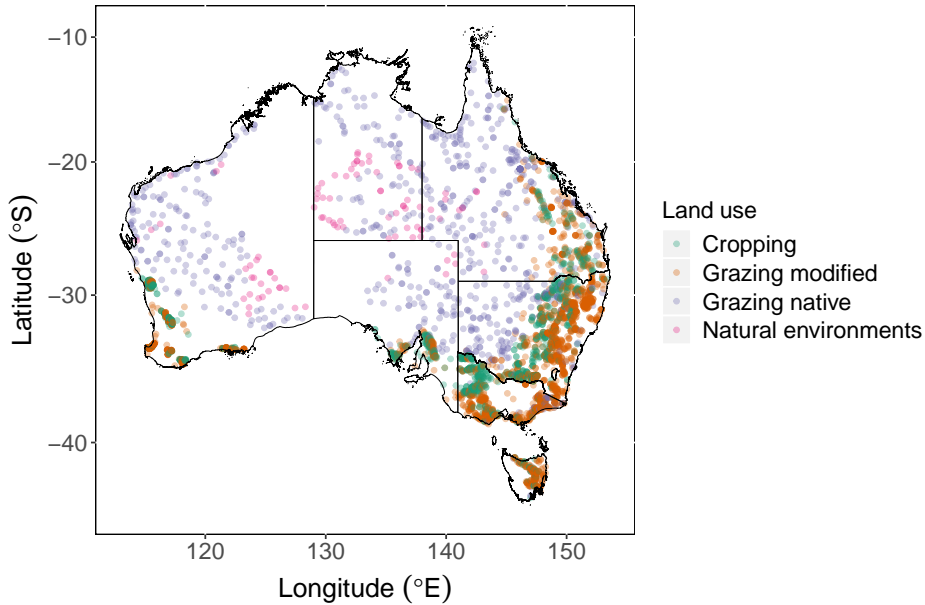
Figure 2. Location of 4,431 sites across Australia selected for this study. 1,261 sites were under cropping, 2,269 sites under grazing modified pastures, 807 sites under grazing of native vegetation, and 94 sites in natural environments, mostly under minimal use or managed resource protection in arid climates.

2.3.2 Data pre-processing and standardisation

The datasets were pre-processed and configured to provide standard and consistent values and units of measurement. Daily weather was extracted at each of the 4,431 sites for the 20 years from 1991 to 2010. The mean of the minimum and maximum

daily temperatures derived the average daily temperatures. Aggregation of the daily weather data produced monthly average temperature, precipitation, and pan evaporation.

We used the Australian land use map (ABARES, 2016) to re-classify each site into the following broad land uses: cropping, modified grazing, native (unmodified) grazing, and natural environments. We defined cropping as land under broadacre crops. Modified grazing was defined as land used for livestock grazing on improved pastures with exotic vegetation cover. Native grazing was defined as land used for grazing on native pastures. Natural environments include the areas for nature conservation, indigenous uses, and other minimal uses. We used the gridded Köppen climate classification from the Bureau of Meteorology (BOM, 2016) to identify sites under natural environments in arid climates. The area of cropping, modified grazing, native grazing, and arid natural environments occupy 292,104 km^2 (or 3.8%), 706,099 km^2 (9.2%), 3,439,468 km^2 (44.8%), and 1,507,616 km^2 (19.6%) of Australia, respectively. Data on agricultural practices at the SA2 level obtained from Unkovich et al. (2017) were used to select a crop or grass to represent typical management regimes in the sites under cropping and modified grazing.

2.3.3 Configuration of land management regimes and initial estimation of C inputs to soil

ROTH C does not calculate plant growth or the quantity of soil C inputs. Therefore, we estimated monthly plant C returns and farmyard manure added to the soil (e.g. managed or deposited by animals grazing on pasture) using the following approach. The initial estimate was made to set the starting values of the C inputs and to match the timing of C inputs to the crop or grass grown.

We assumed that crops were grown in rotations, but at sites under modified pastures, only a single grass species was considered. We used the activity data from Unkovich et al. (2017) to determine crop rotations and a representative grass species for each site during the baseline period between 1991 through 2010. For each of the periods 1990–1994, 1995–1999, 2000–2004, 2005–2009, 2010–2014, we calculated the cumulative frequency by regime. We used it to randomly select the crop or grass species (both annual and perennial) through time with a probability approach. The probability to have a certain crop was dependent on the cumulative frequency assigned to each crop type and regime. The crops grown in all years were selected and then used to determine the most dominant crop species. For the sites under native grazing, we considered a native perennial grass only.

For annual plant species, we used a crop model (Unkovich et al., 2018) that uses the amount of water available to the plant (derived from the measured AWC) to calculate a potential dry matter increment that is water-limited (WLDM) in kg ha^{-1}:

$$\text{WLDM} = ((ET \times Ts) + (DD \times Td)) \times TE$$

where ET is the evapotranspiration (mm) from pan evaporation, DD is any deep water drainage (mm) that occurs during the fallow season, Ts is a fraction of ET that goes through the transpiration, Td is a fraction of deep water drainage that goes through the transpiration, and TE is the transpiration efficiency that is the amount of biomass produced per unit of water transpired (kg mm^{-1}) of a cropping or grazing system. Daily evapotranspiration was estimated by multiplying pan evaporation with a ratio of soil water content over plant AWC and the maximum evapotranspiration by crop or grass. The maximum dry

matter production (DM_{max}) is the sum of dry matter increments over the growing season. This model then back calculates dry matter accumulation (kg ha^{-1}) over the season (DM$_{\mathrm{acc}}$):

$$\mathrm{DM_{acc}} = \frac{DM_{\mathrm{max}}}{\left(1 + e^{-\frac{Day - a \times Days_{\mathrm{max}}}{b \times Day_{\mathrm{sow}} \times Days_{\mathrm{max}}}}\right)}$$

155 where Day is the current day as the season progresses, $Days_{\mathrm{max}}$ is the number of total growing days, Day_{sow} is the day of planting, and a and b are growth coefficients specific for the plant. For a perennial system, daily growth (G) in kg ha^{-1} is calculated as:

$$\mathrm{G} = WLT \times TI \times TE$$

where WLT is the amount of water-limited transpiration (mm) that is evapotranspiration multiplied by vegetation cover, TI
160 is the temperature index function (Nix, 1981), and TE is the transpiration efficiency of a perennial system (kg mm^{-1}). The perennial plant growth is used to calculate dry matter accumulation over the season. The model estimates root biomass using a fixed root-to-shoot ratio of 0.3 (Bol).

For both modified and native pastures, we assumed grazing to occur if the grass accumulated 1.2 Mg ha^{-1} of shoot dry-matter, with no grazing effect on its growth (DPIRD, 2020). The effects of grazing on grass production and animal returns
165 were represented in a pasture growth model developed and integrated in the APSIM model (Johnson et al., 2008). Specifically, grazing animals consumed 50% of daily shoot growth, returned 50% of the consumption to the soil as dung, and shed 50% of daily root growth. When the available soil water fell to < 15% of water holding capacity, 1% and 0.5% of the shoot dry matter and the root dry matter were assumed to die daily. We assumed the C content of above-ground and below-ground residues to be 42% by mass, which is the value used in the FullCAM (Richards and Evans, 2004).
170 For the sites under natural environments, however, we did not use the plant model because we had no data on plants in this region. Instead, we assumed small but consistent C inputs from plant residues only (Wang and Barrett, 2003), which was set to be 0.049 Mg ha^{-1}/month. We assumed no soil cover because in these regions, vegetation cover is typically sparse.

2.3.4 Simulation: optimisation of C inputs to the baseline soil organic C

We initialised the stocks of POC, HOC and ROC pools using the C fractions measured at 4,431 sites. We assumed that the
175 initial soil organic C stocks were at equilibrium, and ran the model to reproduce their equilibrium condition. We based our assumption on data from the National Carbon Accounting System (NCAS) that include temporal soil organic C changes at 73 sites in Australia, recorded from 1911 to 2000 (Skjemstad and Spouncer, 2003). The DPM/RPM ratio determines the decomposability of incoming biomass. By default, the recommended DPM/RPM ratio is 1.44 for most crops and improved pastures and 0.67 for unimproved grasslands (Coleman and Jenkinson, 1996). The DPM/RPM ratio depends on the quality of
180 C in plant residues and manure. It is site-specific, differs with land-use (Post and Kwon, 2000), and is unknown for Australian native grazing or natural environments. We tested six different DPM/RPM ratios (0.67, 0.96, 1.17, 1.44, 1.78 and 2.23) to estimate baseline C inputs and to assess the sensitivity of the simulated TOC, POC and HOC to this parameter. These chosen ratios correspond to allocations of incoming plant material to DPM in the range 40–69%, and proportionally, to RPM in the

range 60–31%. For each DPM/RPM ratio, we run the simulations at each of the 4,431 sites for 100 years. Specifically, for each ratio at each location, we performed the simulations iteratively by re-initialising the POC and HOC pools with the measured C fractions and with a change in monthly input of plant residues and farmyard manure equivalent to 1/100 of their initial values. This was repeated 1000 times or until the model achieved equilibrium. We considered only monthly C inouts in the simulations. The weather data used in the simulations represents the conditions of the baseline period between 1991–2010, which were repeated over the 100-year period.

Equilibrium condition occurred when 1) both POC and HOC did not significantly change over time ($P > 0.05$), or 2) we observed an absolute change of < 0.0025 Mg C ha^{-1} in both POC and HOC. We used a time series linear model with a trend and seasonality to fit the change in POC and HOC over time. An equilibrium condition was also assumed if the direction of the trend (positive or negative) in either pool changed. This condition prevented unrealistic simulations because both POC and HOC showed the same trend in response to C inputs. Depending on the DPM/RPM, at 12 to 14 out of the 4,431 sites, the model was not able to simulate the equilibrium condition. We note that, for the sites that failed, changing C inputs only is insufficient for making both the POC and HOC pools reach equilibrium simultaneously.

We report the stocks of TOC, POC and HOC at the end of the 100-year simulation. We selected 100 years because it is the permanence period for soil C sequestration of the Australian ERF. The difference between the measured and the simulated TOC stock provided an estimate of the model deviation. We also calculated the range of monthly variation in simulated TOC stocks. For each site, we selected the DPM/RPM ratio based on the minimum deviation of TOC. Three hundred and eighty-eight sites had a model deviation and range of monthly change in TOC stock ≥ 10 Mg C ha^{-1}, so we excluded them. We based the 10 Mg C ha^{-1} threshold on the range of measured annual changes in TOC. The median TOC stock at these sites was 75.04 Mg C ha^{-1} (range 52.58–111.44 Mg C ha^{-1}), and mostly, they occurred under modified grazing (data not shown). Finally, we optimised the amount of monthly C input and the DPM/RPM ratio at 4,043 sites and used them as the baseline. We determined the dominant values of the DPM/RPM ratio for each land-use across Australia, based on their relative frequency.

2.3.5   Simulation: potential for C sequestration under changing C inputs

Using the calibrated model, we simulated potential changes in soil organic C over 100 years, in response to changes in C inputs. We selected different rates of C input to the soil by multiplying the optimised baseline with the factors 0 (no input), 0.25, 0.5, 0.75, 1.25, 1.5, 2, 4 and 6. These rates were selected to represent a wide range of C input levels that would be either physically achievable or manageable (e.g. manure addition) (Maillard and Angers, 2014). Because we already calculated the timing of C inputs and the sensitivity to the DPM/RPM ratio, we did not consider scenarios that varied the timing or quality of C inputs. We chose 100 years in order to simulate the long-term response of TOC, POC and HOC.

We calculated 11-year moving averages of the stocks of TOC, POC and HOC and the potential vulnerability of soil C to decomposition over the 100-years. The vulnerability of soil C was derived using POC/(HOC + ROC) (Viscarra Rossel et al., 2019). We calculated changes in soil organic C by changing C inputs. We reported the median stocks for the last 11 years of the simulation when it reached a new equilibrium. We also approximated the lower and upper 95% confidence intervals for the median to quantify the variation of their responses to different C inputs (Conover, 1998).

2.3.6 Empirical assessment of controls on the simulated C change

There are soil and environmental controls on organic C that are not accounted for by ROTH C. To gain a better understanding
of the controls on the change in soil organic C under changing C inputs, we modelled the change in TOC, POC and HOC
as a function of the four land-use classes and a set of environmental variables. The environmental variables included i) soil
properties, such as total nitrogen (N), total phosphorous (P), and C:N (Viscarra Rossel et al., 2015), ii) climate, iii) clay
minerals (illite, kaolinite, and smectite) (Viscarra Rossel, 2011) and iv) potassium (K), thorium (Th) and uranium (U) from
gamma radiometrics, which represent mineralogy and parent material (Minty et al., 2009). For the modelling, we used the
machine learning method CUBIST (Quinlan, 1992). Briefly, CUBIST uses a recursive partitioning of the predictor variable
space and divide the data into subsets that are more similar with respect to the predictors in the data (Quinlan, 1992). A series
of rules derived from if-then conditions define the partitions, and each condition is based on a threshold for one or more of
the predictors. When the conditions in each rule are satisfied, piecewise linear least squares regressions are used to model the
response within each partition. To build precise and stable models, we tested combinations of committees (1, 2, 5, 10, and 20)
and the number of neighbours (0, 2, 5, and 9) using 10-repeated cross-validation (Hastie et al., 2009). We used the minimum
root mean squared error (RMSE) to select the best model. We then assessed the relative importance of each variable based on
the usage of each variable in the conditions and the models.

3 Results

3.1 Effect of different quality of C inputs on soil organic C

The median stocks of TOC, POC and HOC in the 0–0.3 m soil layer, calculated across Australia are 26.01, 3.43 and 16.10 Mg C ha^{-1},
respectively (Figure 3). The measured TOC stocks under natural environments, native grazing, modified grazing and cropping
are 15.45 Mg C ha^{-1} (interquartile range 11.89–18.11 Mg C ha^{-1}), 24.61 Mg C ha^{-1} (18.96–34.17), 51.48 Mg C ha^{-1}
(39.01–74.60) and 35.38 Mg C ha^{-1} (25.39–43.55), respectively. The POC and HOC fractions consist of 11% and 68% of the
measured TOC stocks under natural environments, 11% and 67% under native grazing, 18% and 52% under modified grazing,
and 16% and 53% under cropping.

With each of the DPM/RPM ratios tested, the model simulated the measured TOC, POC and HOC stocks at equilibrium
(Supplement Table S1), but the amount of annual C input needed to maintain the soil organic C stocks was sensitive to the
varying quality of incoming plant material. The C inputs increased from 1.47 to 1.83 Mg C ha^{-1} when the DPM/RPM ratio
increased from 0.67 (low decomposability) to 2.23 (high decomposability). With those changes, the rate of C inputs into DPM
rose from 0.59 to 1.26 Mg C ha^{-1} yr^{-1}, while the rate into RPM decreased from 0.88 to 0.57 Mg C ha^{-1} yr^{-1}. The addition
of biomass C with different qualities affected the levels of POC and HOC at equilibrium (Supplement Table S1).

With an optimised DPM/RPM ratio at each location, the model was able to explain 97–98% of the measured variation in
TOC at sites under native grazing, modified grazing and cropping. RMSE values ranged from 2.45 to 3.55 Mg C ha^{-1} (Figure
3). At locations under natural environments, the model explained only 65% of the variation in TOC, but with a similar RMSE

250 of 3.22 Mg C ha^{-1}. The model could explain less of the variation in POC (55–89%) compared to TOC and HOC. Also, the model did not perform as well for POC in the soils under cropping. Across Australia, the most frequent DPM/RPM ratio was 2.23 (1,773 sites), followed by the value 0.67 (for 829 sites) and 0.96 (415 sites) (Figure 3). The simulated median and total TOC, POC and HOC stocks for each land-use class and overall were similar to the measured data (Table 1). This gave us confidence in the performance of the model in Australia.

Table 1. The simulated stocks of total, particulate and humus organic C (TOC, POC and HOC) under different land uses and in Australia. The total stocks of soil organic C were calculated from the median with the uncertainties expressed as an approximate 95% confidence interval (CI).

Land use		Median (Mg/ha)	1st quantile (Mg/ha)	3rd quantile (Mg/ha)	Total (Gt)	Lower 95% CI (Gt)	Upper 95% CI (Gt)
Cropping	TOC	40.25	30.87	47.74	1.18	1.16	1.20
(n = 1182)	POC	7.76	5.63	10.22	0.23	0.22	0.23
	HOC	21.03	15.17	25.13	0.61	0.60	0.63
Modified grazing	TOC	53.65	42.24	74.37	3.79	3.72	3.86
(n = 2008)	POC	10.66	7.70	15.13	0.75	0.73	0.77
	HOC	27.87	21.74	35.91	1.97	1.92	2.01
Native grazing	TOC	23.85	17.98	33.60	8.20	7.90	8.54
(n = 777)	POC	3.86	2.77	5.73	1.33	1.27	1.40
	HOC	14.08	10.65	19.05	4.84	4.64	4.99
Natural environment	TOC	13.00	9.61	17.05	1.96	1.65	2.30
(n = 76)	POC	2.29	1.72	3.28	0.35	0.29	0.42
	HOC	6.84	4.98	10.18	1.03	0.88	1.30
Australia*	TOC	25.45	19.37	34.94	19.52	18.63	20.52
	POC	4.46	3.23	6.44	3.43	3.24	3.64
	HOC	14.22	10.75	19.10	10.91	10.37	11.52

* The Australian-wide estimates were the area weighted averages of the medians for the four land-use classes. The area of cropping, modified grazing, native grazing, and arid natural environments occupy 3.8%, 9.2%, 44.8%, and 19.6% of Australia (total area 7673138 km^2), respectively.

3.2 Effect of changing C inputs on soil organic C

The TOC, POC and HOC stocks at equilibrium were positively related to the level of C inputs (Figure 4). Annual C inputs to the soil under natural environments, native grazing, modified grazing and cropping were 2.38, 0.77, 1.86 and 1.60 Mg C ha^{-1},
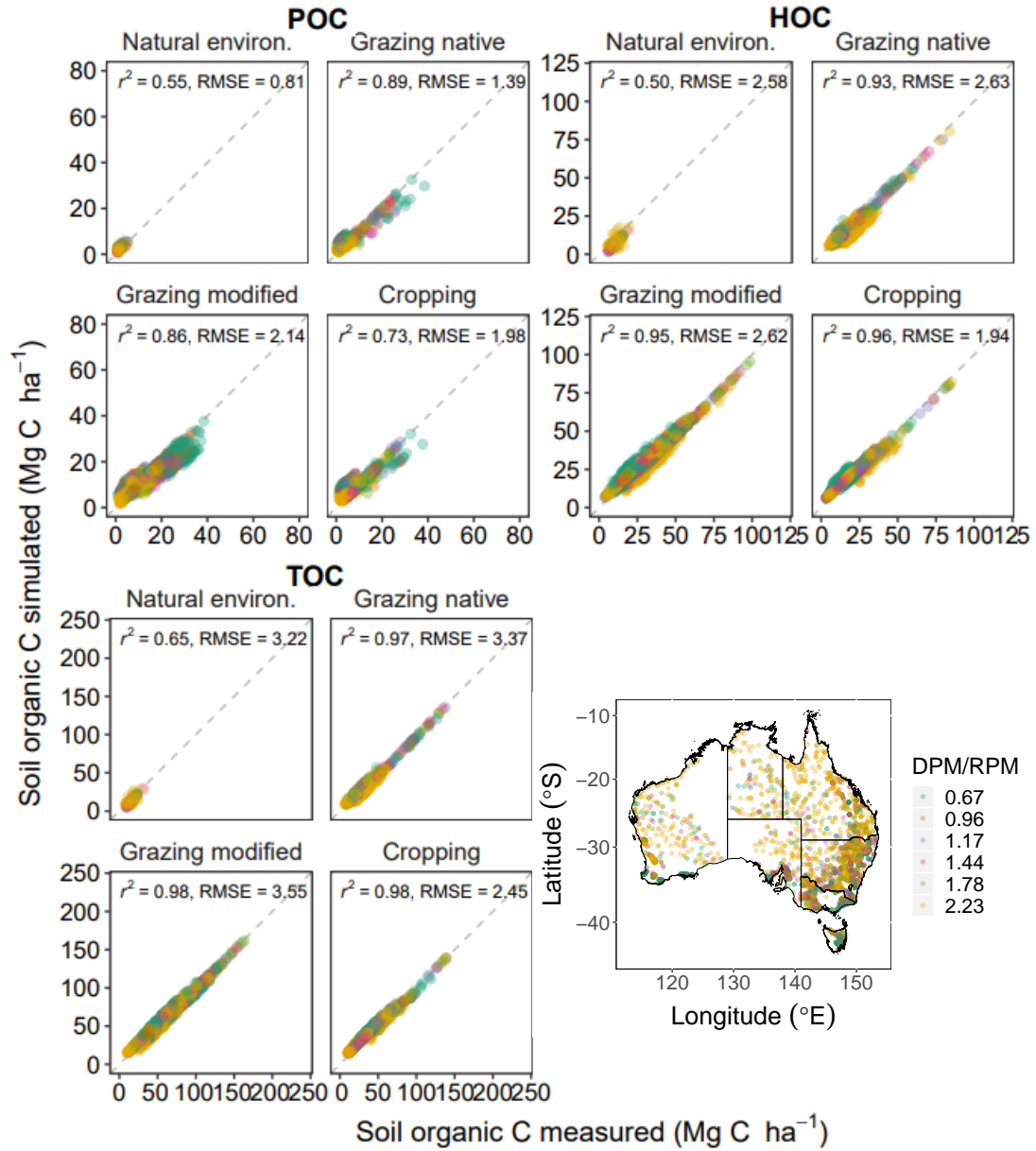
Figure 3. Simulation of equilibrium soil organic C levels after optimisation, based on its sensitivity to the changes in the allocation of incoming plant and manure C into the decomposable plant material (DPM) and resistant plant material (RPM) components (n = 4,043). The map shows the geographic distribution of DPM/RPM values, where high values correspond to faster decomposition.

respectively. Therefore, the model estimated the largest amount of C inputs required to maintain soil organic C under natural environments compared to the other land uses. The corresponding interquartile range was 1.11–3.57 Mg C ha^{-1} for natural

260 environments. In comparison, there was a wider range of C inputs for native grazing (0.57–1.13 Mg C ha^{-1}), modified grazing (1.37–3.01 Mg C ha^{-1}), and cropping (1.20–2.18 Mg C ha^{-1}) (Figure 4). For the agricultural soils, clay affected the relationship between soil organic C stocks and C inputs as soil with more clay (predominantly in eastern Australia) could hold more organic C (Figure 4). Under grazing and cropping land uses, the response of HOC and TOC to increasing C inputs appears to depend on clay content. This pattern was not evident for POC as this pool is not directly associated with clay in the model.

265 The model explained 78, 80, and 50% of the variation in TOC, HOC, and POC by increasing C input under cropping (Figure 4). The relationship was poorer under native and modified grazing (r^2 = 0.54–0.69) (Figure 4). There was a relatively weak and divergent relationship between soil organic C stocks and C inputs to the soil under natural environments (r^2 = 0.35–0.40), mostly due to differences in precipitation. We found that soil organic C was more responsive to C inputs at the sites with less annual precipitation (approximately 170 mm).

270 The TOC, POC and HOC stocks at a new equilibrium responded linearly to changing soil C inputs from the baseline (Figure 5). With an annual increase of 1 Mg C ha^{-1} in C inputs and under current climatic conditions, soils under natural environments can potentially increase TOC stocks by 3.52 Mg C ha^{-1} (interquartile range (IQR) = 3.15–4.09 Mg C ha^{-1}). In this case, the stocks of POC and HOC increased by 0.92 Mg C ha^{-1} (IQR = 0.73–1.35) and 2.48 Mg C ha^{-1} (IQR = 2.23–2.99), respectively. Soils under the other land use were more sensitive to changing C inputs. Under native grazing, TOC, POC and

275 HOC stocks changed by 15.57 Mg C ha^{-1} (IQR = 12.07–17.82), 5.49 Mg C ha^{-1} (IQR = 3.93–6.97) and 9.18 Mg C ha^{-1} (IQR = 7.46–9.96), respectively, and at the same rate as the C inputs in the simulation. Under modified grazing, TOC stocks changed by 14.21 Mg C ha^{-1} (IQR = 12.38–16.03) and POC and HOC accounted for 5.34 Mg C ha^{-1} (IQR = 4.46–6.86) and 8.12 Mg C ha^{-1} (IQR = 7.23–8.93) of the change, respectively (Figure 5). Changes in TOC, POC and HOC stocks under cropping were 13.58 Mg C ha^{-1} (IQR = 12.19–15.80), 4.69 Mg C ha^{-1} (IQR = 3.76–6.48) and 8.35 Mg C ha^{-1} (IQR =

280 7.49–8.96), respectively.

With increasing C inputs, the potential vulnerability of soil organic C to loss increased under all four land uses (Figure 5). The soil under native grazing was the most vulnerable with the increase in POC (35%) and HOC (59%), showing that the labile POC increased more proportionally than that of the other land uses. Soil under natural environments was the least vulnerable to change because of the smaller increase in POC (26%) relative to HOC (70%).

285 **3.3 The controls on the simulated soil organic C change**

Climatic variables, particularly temperature and potential evaporation, controlled the changes in TOC, POC and HOC (Figure 6). Clay content had a dominant effect on the changes in HOC because in ROTH C, clay determines the ratio of CO_2 released to HOC formed, during decomposition. Total N, the C:N ratio and pH were important controls for the changes in POC (Figure 6), and might be related to a capacity of the soil to form POC. Cropping affected the changes in POC, possibly because

290 of the crop-specific distribution of C inputs. The controls on POC were similar to those on TOC because their changes were
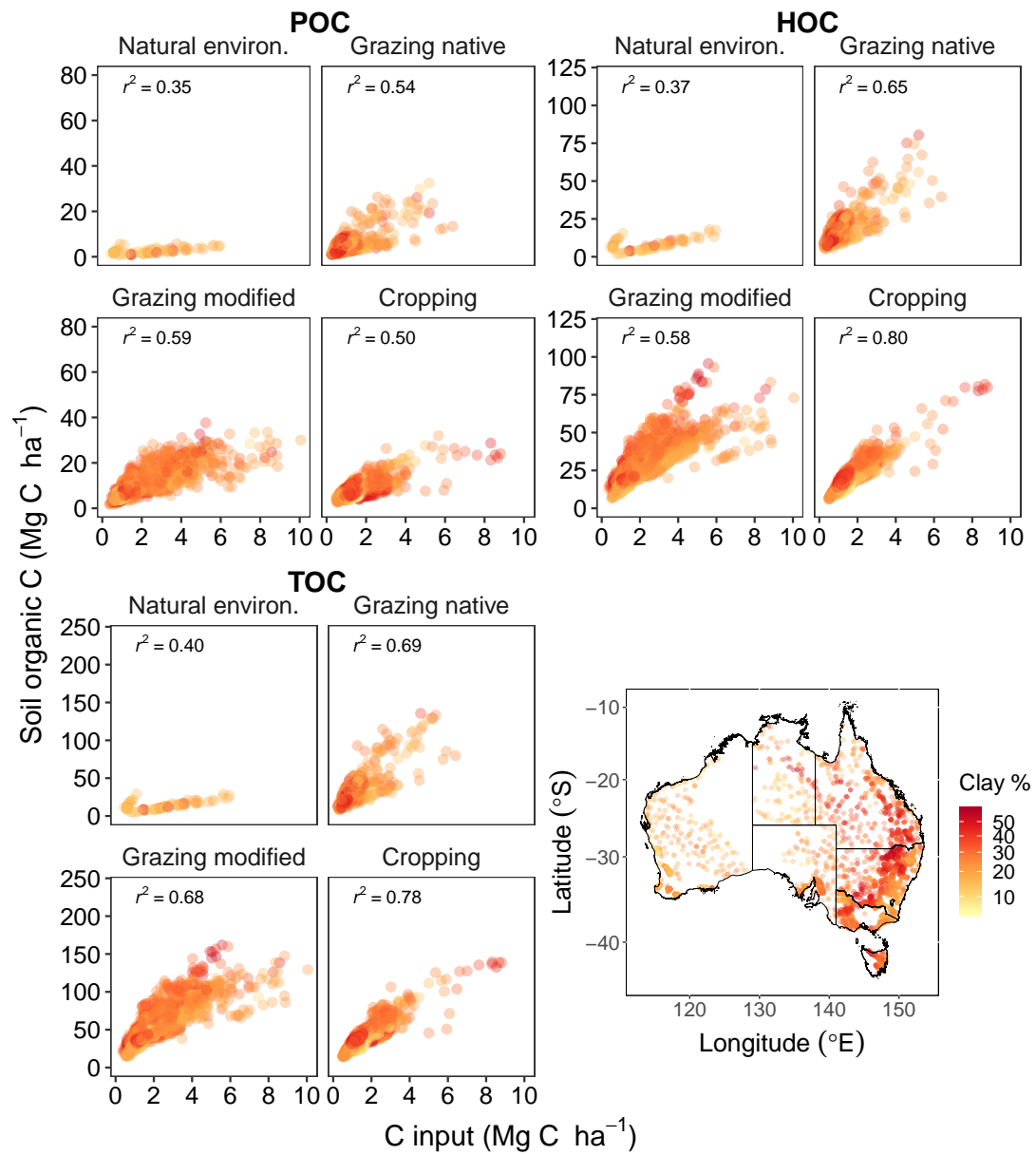
Figure 4. Changes in total, particulate and humus organic C (TOC, POC and HOC) with C inputs by land use (n = 4,043).
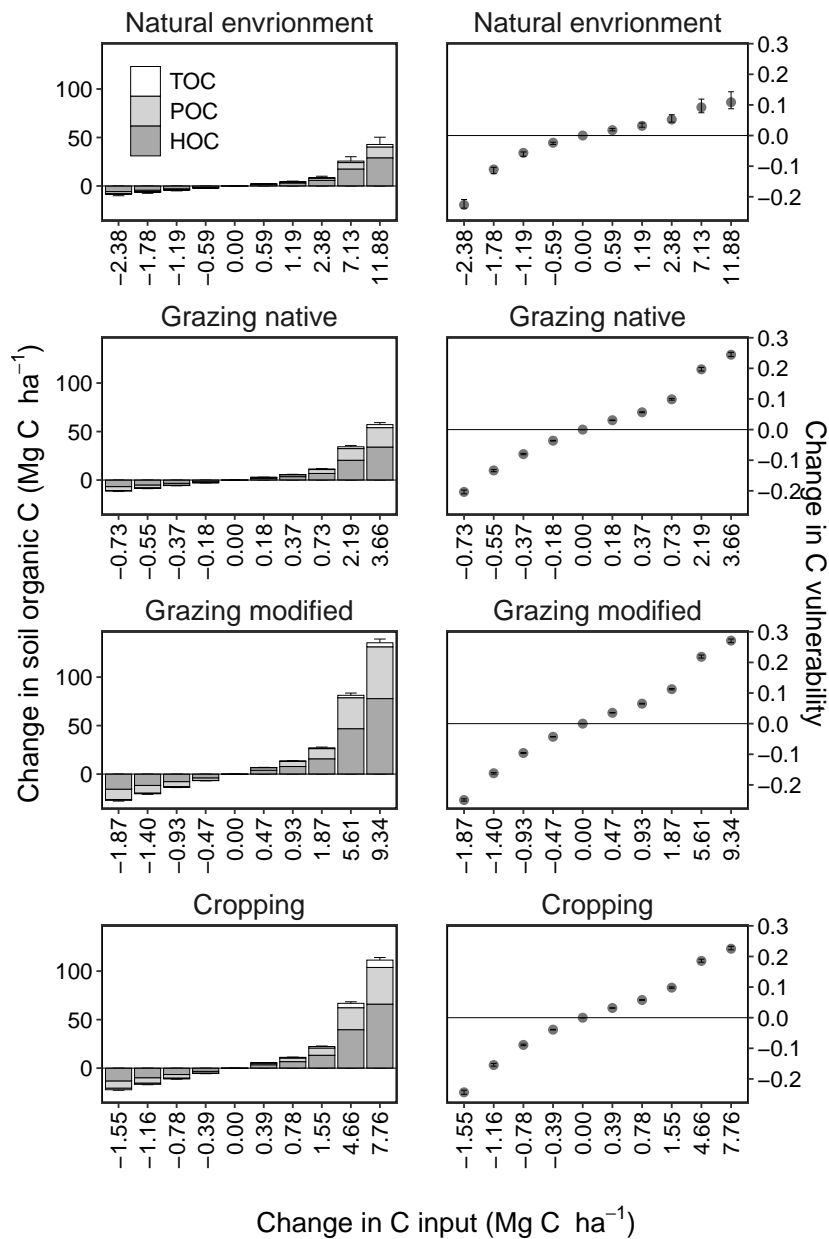
Figure 5. Simulated potential changes in total organic C (TOC) in the topsoil (0–0.3 m) following the changes in C input (n = 4,043), consisting of particulate and humus organic C (POC and HOC) (left) and corresponding changes in C vulnerability (right). At each site, baseline C input was multiplied by the factor 0, 0.25, 0.5, 0.75, 1.25, 1.5, 2, 4 and 6 to derive different C input levels. A positive change in the potential C vulnerability shows increased vulnerability to decomposition, while a negative change indicates more resistance to loss. The potential vulnerability is derived by POC/(HOC + ROC). The error bar represents the model variation within the interquartile range from the median.

proportional. The land use in natural environments affected the changes in HOC (Figure 6), suggesting that we need a greater understanding of the potential for C sequestration in low clay content soils in arid climates.
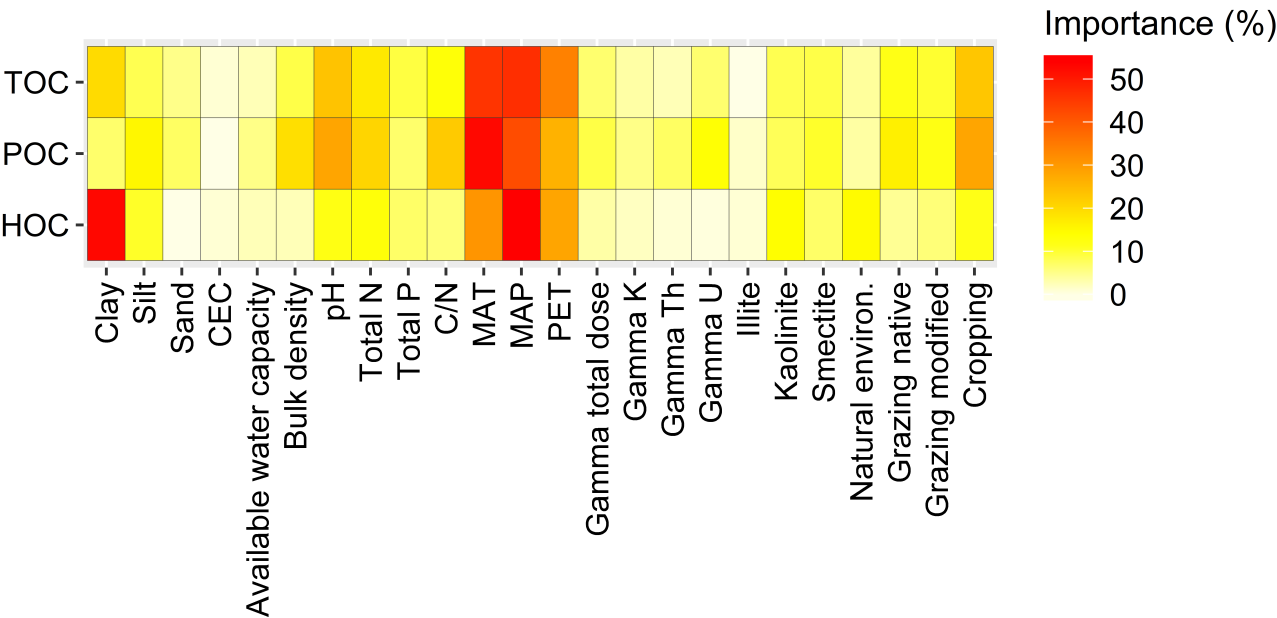
Figure 6. Importance of the environmental variables that contribute to potential changes in total, particulate and humus C organic C (TOC, POC and HOC) by changing C inputs. Climatic variables, i.e. mean annual temperature (MAT), mean annual total precipitation (MAP) and potential evapotranspiration (PET), are averaged over a period of 1991–2010. CEC is the cation exchange capacity of a soil.

4 Discussion

4.1 Current status and potential change in organic C of Australian soils

295 We used ROTH C because it requires few parameters, it initialises its main pools with measured C fractions, and it was adjusted to suit Australian conditions (Janik et al., 2002; Skjemstad et al., 2004). Further, the model is in Australia's National Greenhouse Gas Inventory System and the ERF, and so we thought it useful to comply. The climatic and soil property inputs needed to run ROTH C are readily available from publicly available datasets (see Methods) or are relatively easily measured, for instance, with proximal sensors (England and Viscarra Rossel, 2018).

The main soil C pools of ROTH C can be initialised with measured C fractions (POC, HOC, ROC), there is no need for spin-up simulations (i.e. simulations until the model reaches equilibrium), making it possible to run the model site-specifically at any location in Australia, without additional effort. Further, using measured C fractions in the model allows for correct assignment of the primary pool structure and the measurements serve as internal verification of the model. In our case, we could empirically assess how well the baseline simulations matched the model's corresponding dynamic pools, which gave us additional confidence that the model was performing well and could well represent Australian soils. Such data-driven model initialisation helps with the selection and site-specific estimation of other 'unknown' model parameters, such as the amount and quality of C inputs, which is important for a more consistent calibration of the model (Aber, 1997; Seidel et al., 2018). Our simulations successfully optimised both the amount and the quality of C inputs to maintain the current baseline soil organic C stocks.

The model performed well in soils that are under agriculture (both cropping and grazing), but the simulation under natural environments in semi-arid and arid climates need improving. Together with the relatively large C inputs, required to maintain baseline TOC (9.61 to 17.05 Mg C ha^{-1}), this poor performance suggests that the model did not represent well the complex decomposition processes described by the (hot and dry) climate and soil under natural environments. We hope to address this in subsequent research because soil C in semi-arid and arid climates might represent a crucial C sink in Australia and other similar regions of the world (Farina et al., 2013).

The simulated baseline estimate of the total TOC stock in Australia is 19.52 Gt, which is less than the 24.97 Gt estimate of Viscarra Rossel et al. (2014), as soils under land uses that contain more carbon, e.g. forests were not included in this study. Our estimates for soil under natural environments and native grazing are 1.96 Gt (with 95% confidence intervals of 1.65–2.30 Gt) and 8.02 Gt TOC (7.90–8.54 Gt), respectively. The soil under native grazing has the largest organic C stocks compared to the other land uses. The contribution of native grazing to the national soil organic C budget is considerable due to the large extent of land that it covers. This estimate was well within the confidence intervals derived by Viscarra Rossel et al. (2014), although slightly larger. Estimates of the total TOC stocks for soils under modified grazing and cropping are 3.79 Gt (with 95% confidence intervals 3.72–3.86 Gt) and 1.18 Gt (1.16–1.20 Gt), respectively. These estimates were also somewhat larger than those of Viscarra Rossel et al. (2014). Our estimates of the total POC and HOC stocks across all four land uses are 3.43 and 10.91 Gt, which are smaller than the 7.8 and 27.3 Gt, respectively, estimated derived by Viscarra Rossel et al. (2019). However, our estimates are within the range of their confidence intervals. A reason for the differences between our estimates and those of Viscarra Rossel et al. (2014, 2019) might be that our estimates from the simulations are based on a relatively sparse sample (Figure 1), while theirs are from a complete enumeration of Australia with spatial machine learning models. Nevertheless, the results from our simulations suggest that the ROTH C model can explain the soil processes under different land uses that are important for estimating the baseline total stocks of soil organic C and its composition.

There are few quantitative assessments of soil C dynamics in Australia. Primarily they are for cropping regions (Luo et al., 2014; Lam et al., 2013; Wang et al., 2016), some present local case studies with few data (Hoyle et al., 2013), and some report estimates that are uncertain because of the lack of comprehensive surveys and scarcity in data (Gifford, 2010). Here, we simulated soil organic C at 4,043 sites across Australia to estimate changes in C stocks from a range of plausible changes in C

335   inputs to the soil. With an annual increase of 1 Mg C ha^{-1} in C inputs, the model estimated the largest potential soil C increase in soil under native grazing (12.07–17.82 Mg C ha^{-1}) and modified grazing (12.38–16.03 Mg C ha^{-1}). The potential increase in soils under cropping was smaller (12.19–15.80 Mg C ha^{-1}), possibly due to the effect of soil disturbances and cultivation on decomposition. However, the difference between grazing and cropping is small as the effects of climate and soil texture on organic matter and its decomposition are likely to be similar over the large areas that these land uses occupy. Soils in natural

340   environments had the smallest potential to accumulate C (3.15–4.09 Mg C ha^{-1}), because they occur over large areas with semi-arid to arid climates characterised by low precipitation, generally below 500 mm yr^{-1} and high temperatures up to $50°$ (ABS, 2016).

We also simulated the potential vulnerability of soil organic C to loss. Results suggest that with increases in organic C, Australian soils would potentially become more vulnerable to losses. This is because of the preferential accumulation of C in

345   the more labile POC pool, which is more sensitive to decomposition with increasing temperatures. We need to better understand mechanisms of C stabilisation and management to capture the proportion of new C sequestered in the soil in the more stable pools.

We did not use net primary productivity (NPP) as a proxy for C inputs to the soil. Although large-scale estimates of NPP might be a good proxy for the C inputs in natural environments, they would be inadequate for managed systems (Haverd et al.,

350   2013). To derive estimates of NPP for managed land uses, such as croplands, one needs fine spatial resolution land cover data with crop-specific information (Li et al., 2014; Turner et al., 2006). These are not readily available continentally. Large-scale (global, continental or regional) estimates of NPP, such as those available from coarser resolution remote sensing, would not be suitable for agricultural environments, also because depending on the method used to derive NPP, the estimates would be largely uncertain (Roxburgh et al., 2005; Ciais et al., 2010). Therefore, using NPP as an estimate of C inputs for all four land

355   uses would have made our simulations more uncertain. We thought it important to maintain a consistent approach for deriving the C inputs, so we used a wide but plausible range of values to represent the C inputs across the whole of Australia. The range of C inputs that we used are representative of values that might be expected from management practices that enhance rates of primary production and C input to the soil, including manure addition (Lal, 2016; Paustian et al., 2019). Our results suggest that the baseline rate of C inputs into the active POC and HOC pools is site-specific, and managing its rate locally is needed to

360   avoid soil C loss from land-use change. Importantly, these estimates of C inputs are useful to locate soils where C capture is possible under limited availability of water resources and nutrients (Baldock et al., 2012).

The long-term changes in organic C are primarily determined by the C inputs into the soil, and the sensitivity of the change can be affected by local conditions. The results from the empirical modelling suggest that simulations might improve if we can modify the environmental effects on decomposition, separately for each of the pools. For example, clay content did not

365   importantly affect the changes in POC, but it did affect the changes in HOC. In contrast, other studies have shown that clay has a direct effect on both C inputs and the C pools in Australian soils (Krull et al., 2003; Luo et al., 2017). Of course, this might be due to the inability of the model to simulate textural controls on POC. Total N and the C:N ratio contribute more to the changes in POC than in HOC. POC appears to be also affected by pH and more under cropping. These results demonstrate the difficulty

that ROTH C has in simulating the more labile POC dynamics and the need to represent such additional environmental factors to better explain TOC change.

4.2 Simulating soil C dynamics using a standardise approach

There is a functional disconnection between measurements, data and biogeochemical models (Blankinship et al., 2018), but by simulating under a framework, like we did here, we can bridge that disconnect. A framework provides a standardised and consistent approach for organising and processing input datasets from different sources, to facilitate calibration, verification, estimation and prediction at an appropriate scale and resolution, depending on the study. The input data may originate from field or laboratory measurements, remote sensing, digital soil maps or other data from various sources. Using a standardised approach, soil C simulations can be more versatile. They can be performed on points, areas or pixels, even when few, or no site-specific data are available. In the latter case, by using fine spatial resolution information (Viscarra Rossel et al., 2014, 2015, 2019), or like we have shown here, one can use measurements together with publicly available continental-scale datasets and processes them consistently for the simulations. When site-specific data are available, then under the framework, they are processed appropriately for the local simulations, as we have shown in Lee and Viscarra Rossel (2020).

Simulating soil organic C in a standardised manner also facilitates consistent pre-processing, quality checks and explicit definition of the simulation unit. This is important because often, datasets have a different formats and resolutions, which must be standardised and harmonised before running the simulation (Batjes et al., 2020). Datasets may need to be aggregated or disaggregated over space and time, depending on the data and the need. For example, if the need is to run the simulations over a large-scale and over grids, finer resolution data e.g. soil property data, will require aggregation to match the coarser resolution of the simulation unit. Similarly, re-classification of categorical data, e.g. land-use data, may be performed, like we have done here, to set the spatial extent of the simulations. We used the model ROTH C, however, by using our approach, one could accommodate other soil C models, with only small changes to the workflow (Figure 1). This versatility is essential for extending our theoretical understanding of C cycling and its response to human-induced and environmental change at appropriate scales (Grunwald et al., 2011; Metting et al., 2001). Of course, with other multi-pool C models, it will be important to explore further the initialisation requirements and the baseline state for the simulations. The reason is that each soil C pool could be at a different state. Other models may also drive decomposition based on different assumptions, e.g. soil enzyme kinetics or microbial growth (Smith et al., 2020).

4.3 Future needs

Plant biomass production and subsequent C inputs to the soil are critical determinants of the quantity of organic matter in soil C models. The simulations that we presented estimated the potential of soil C sequestration in response to changing C inputs under the main land uses in Australia. However, we will need data on plant growth properties, seasonal biomass data, and residue and grazing management, to better represent management practices under these land uses. Without such datasets, it is difficult to verify the balance between C inputs and the stocks and composition of soil organic C under different land-use and management combinations, except for a few cropping systems (Wang et al., 2016). For the soils under native grazing, we need

new research on the specific growing conditions of plants (e.g. nutrient availability) and how they affect the amount and timing of C inputs.

The machine learning could identify some other factors that contribute to the changes in soil organic C and determine their relative importance. Although there is no direct mechanistic understanding gained from those analyses, some of those variables are important predictors of soil C change, and they might need accounting in future model development. In practice, statistical modelling can be incorporated in the simulations to help identify the balance of C flows between the soil, plant, and atmosphere at the scale of interest. However, research to combine mechanistic and statistical modelling is still at an early stage, and more research is needed to connect data with models (O'Rourke et al., 2015; Vereecken et al., 2016), in a consistent manner and across scales, for example, Viscarra Rossel et al. (2019). With new measurements and subsequently growing datasets, we expect to identify new processes and controls from statistical modelling and to further account for these in a standardised modelling approach.

5 Conclusion

The simulations of soil organic C across Australia with ROTH C were performed using a standardised approach that establishes a much-needed connection between measurements, datasets and models. It enabled consistent processing of measurements and datasets from different sources, and standardisation and configuration of the model for calibration, verification, estimation and prediction. Our results show that the site-specific initialisation of the C pools with measurements of the C fractions (POC, HOC, ROC) are essential for accurately representing baseline soil organic C stocks and composition. With a site-specific optimisation of the DPM/RPM ratio, the model could explain 97–98% of the variation in TOC under native grazing, modified grazing and cropping, respectively, and 65% under natural environments. Site-specific initialisation is also essential for estimating changes in C stocks and the potential of soils for C capture and storage. Our 100-year simulations showed that, with an annual increase of 1 Mg C ha^{-1}, and under constant climate, the potential for C sequestration, as well as the potential vulnerability to C loss, in Australian soils is smallest in soils under natural environments, it is greater under cropping and modified grazing, and is greatest in the soils under native grazing.

Code availability. Scripts used for data processing are available from the corresponding author on reasonable request.

Data availability. The spatial datasets on climate, soil and land use are publicly available from repositories cited in section 2.3.1. Other datasets are available from the corresponding author on reasonable request.

Author contributions. RVR formulated the research and with JL designed the simulations. JL performed the simulation and with RVR the data analysis. RVR and JL wrote the manuscript with input from YW and ZL.

430 *Competing interests.* The authors declare that they have no conflict of interest.

Acknowledgements. This work was supported by funding from Curtin University.

References

Estimating shoot to root ratios and annual carbon inputs in soils for cereal crops, Agriculture, Ecosystems & Environment, 63, 61–66.

Simplified simulation models based on specified minimum data sets: the CROPEVAL concept, A.A.Balkema, Rotterdam, http://hdl.handle.net/102.100.100/292180?index=1, 1981.

ABARES: Land Use of Australia 2010–11, australian Bureau of Agricultural and Resource Economics and Sciences (ABARES), Canberra, https://www.agriculture.gov.au/abares/data, 2016.

Aber, J. D.: Why don't we believe the models?, Bulletin of the Ecological Society of America, 78, 232–233, 1997.

Abramoff, R., Xu, X., Hartman, M., O'Brien, S., Feng, W., Davidson, E., Finzi, A., Moorhead, D., Schimel, J., Torn, M., and Mayes, M. A.: The Millennial model: in search of measurable pools and transformations for modeling soil carbon in the new century, Biogeochemistry, 137, 51–71, https://doi.org/10.1007/s10533-017-0409-7, 2018.

ABS: Statistical Area Level 2 (SA2), australian Bureau of Statistics (ABS), Statistical Geography, https://www.abs.gov.au/websitedbs/D3310114.nsf/home/geography, 2016.

Baldock, J. A., Wheeler, I., McKenzie, N., and McBrateny, A.: Soils and climate change: potential impacts on carbon stocks and greenhouse gas emissions, and future research for Australian agriculture, Crop & Pasture Science, 63, 269–283, https://doi.org/10.1071/cp11170, 2012.

Batjes, N. H.: Total carbon and nitrogen in the soils of the world, European Journal of Soil Science, 47, 151–163, https://doi.org/DOI 10.1111/j.1365-2389.1996.tb01386.x, 1996.

Batjes, N. H., Ribeiro, E., and van Oostrum, A.: Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019), Earth Syst. Sci. Data, 12, 299–320, https://doi.org/10.5194/essd-12-299-2020, 2020.

Blankinship, J. C., Berhe, A. A., Crow, S. E., Druhan, J. L., Heckman, K. A., Keiluweit, M., Lawrence, C. R., Marin-Spiotta, E., Plante, A. F., Rasmussen, C., Schadel, C., Schimel, J. P., Sierra, C. A., Thompson, A., Wagai, R., and Wieder, W. R.: Improving understanding of soil organic matter dynamics by triangulating theories, measurements, and models, Biogeochemistry, 140, 1–13, https://doi.org/10.1007/s10533-018-0478-2, 2018.

BOM: The Köppen classification maps, australian Bureau of Meteorology, http://www.bom.gov.au/climate/, 2016.

Campbell, E. E. and Paustian, K.: Current developments in soil organic matter modeling and the expansion of model applications: a review, Environmental Research Letters, 10, https://doi.org/10.1088/1748-9326/10/12/123004, 2015.

Ciais, P., Paris, J. D., Marland, G., Peylin, P., Piao, S. L., Levin, I., Pregger, T., Scholz, Y., Friedrich, R., Rivier, L., Houwelling, S., Schulze, E. D., and (1), M. O. T. C. S. T.: The European carbon balance. Part 1: fossil fuel emissions, Global Change Biology, 16, 1395–1408, https://doi.org/10.1111/j.1365-2486.2009.02098.x, 2010.

Coleman, K. and Jenkinson, D. S.: Roth C-26.3 - A model for the turnover of carbon in soil, pp. 237–246, Evaluation of Soil Organic Matter Models, Springer Berlin Heidelberg, Berlin, Heidelberg, 1996.

Conant, R. T., Ogle, S. M., Paul, E. A., and Paustian, K.: Measuring and monitoring soil organic carbon stocks in agricultural lands for climate mitigation, Frontiers in Ecology and the Environment, 9, 169–173, https://doi.org/10.1890/090153, 2011.

Conover, W. J.: Practical Nonparametric Statistics, 3rd Edition, John Wiley and Sons, New York, 1998.

DPIRD: Department of Primary Industries and Regional Development, Grazing annual pastures – using feed on offer (FOO) as a guide, https://www.agric.wa.gov.au/pasture-management, 2020.

England, J. R. and Viscarra Rossel, R. A.: Proximal sensing for soil carbon accounting, SOIL, 4, 101–122, https://doi.org/10.5194/soil-4-101-2018, 2018.

470   Falloon, P. D. and Smith, P.: Modelling refractory soil organic matter, Biology and Fertility of Soils, 30, 388–398, https://doi.org/10.1007/s003740050019, 2000.

Farina, R., Coleman, K., and Whitmore, A. P.: Modification of the RothC model for simulations of soil organic C dynamics in dryland regions, Geoderma, 200-201, 18–30, https://doi.org/10.1016/j.geoderma.2013.01.021, 2013.

Gifford, R. M.: Carbon sequestration in Australian grasslands: policy and technical issues, Tech. rep., 2010.

475   Grunwald, S., Thompson, J. A., and Boettinger, J. L.: Digital Soil Mapping and Modeling at Continental Scales: Finding Solutions for Global Issues, Soil Science Society of America Journal, 75, 1201–1213, https://doi.org/10.2136/sssaj2011.0025, 2011.

Harden, J. W., Hugelius, G., Ahlström, A., Blankinship, J. C., Bond-Lamberty, B., Lawrence, C. R., Loisel, J., Malhotra, A., Jackson, R. B., Ogle, S., Phillips, C., Ryals, R., Todd-Brown, K., Vargas, R., Vergara, S. E., Cotrufo, M. F., Keiluweit, M., Heckman, K. A., Crow, S. E., Silver, W. L., DeLonge, M., and Nave, L. E.: Networking our science to characterize the state, vulnerabilities, and management

480   opportunities of soil organic matter, Global Change Biology, 24, e705–e718, https://doi.org/10.1111/gcb.13896, 2018.

Hastie, T., Tibshirani, R., and Friedman, J.: Model Assessment and Selection, in: The Elements of Statistical Learning Data Mining, Inference, and Prediction, Second Edition, chap. 7, pp. 219–260, Springer-Verlag New York, 2009.

Haverd, V., Raupach, M. R., Briggs, P. R., Canadell, J. G., Isaac, P., Pickett-Heaps, C., Roxburgh, S. H., van Gorsel, E., Rossel, R. A. V., and Wang, Z.: Multiple observation types reduce uncertainty in Australia's terrestrial carbon and water cycles, Biogeosciences, 10, 2011–2040,

485   https://doi.org/10.5194/bg-10-2011-2013, 2013.

Hoyle, F. C., D'Antuono, M., Overheu, T., and Murphy, D. V.: Capacity for increasing soil organic carbon stocks in dryland agricultural systems, Soil Research, 51, 657–667, https://doi.org/10.1071/Sr12373, 2013.

Janik, L., Spouncer, L., Correll, R., and Skjemstad, J.: Sensitivity Analysis of the ROTH-C Soil Carbon Model (Ver. 26.3 Excel©), National Carbon Accounting System Technical Report No. 30, Australian Greenhouse Office, Canberra, ACT, 2002.

490   Jenkinson, D. S.: The turnover of organic carbon and nitrogen in soil, Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 329, 361–368, https://doi.org/10.1098/rstb.1990.0177, 1990.

Jenkinson, D. S. and Rayner, J. H.: The turnover of soil organic matter in some of the Rothamsted classical experiments, Soil Science, 123, 298–305, https://doi.org/doi.org/10.1097/00010694-197705000-00005, 1977.

Johnson, I. R., Chapman, D. F., Snow, V. O., Eckard, R. J., Parsons, A. J., Lambert, M. G., and Cullen, B. R.: DairyMod and EcoMod:

495   biophysical pasture-simulation models for Australia and New Zealand, Australian Journal of Experimental Agriculture, 48, 621–631, https://doi.org/10.1071/EA07133, 2008.

Krull, E. S., Baldock, J. A., and Skjemstad, J. O.: Importance of mechanisms and processes of the stabilisation of soil organic matter for modelling carbon turnover, Functional Plant Biology, 30, 207–222, <GotoISI>://WOS:000181046800007, 2003.

Lal, R.: Beyond COP21: Potential and challenges of the "4 per Thousand" initiative, Journal of Soil and Water Conservation, 71, 20a–25a,

500   https://doi.org/10.2489/jswc.71.1.20A, 2016.

Lam, S. K., Chen, D., Mosier, A. R., and Roush, R.: The potential for carbon sequestration in Australian agricultural soils is technically and economically limited, Scientific Reports, 3, 2179, https://doi.org/10.1038/srep02179, 2013.

Lee, J. and Viscarra Rossel, R. A.: Soil carbon simulation confounded by different pool initialisation, Nutrient Cycling in Agroecosystems, 116, 245–255, https://doi.org/10.1007/s10705-019-10041-0, 2020.

505   Lehmann, J. and Kleber, M.: The contentious nature of soil organic matter, Nature, 528, 60–68, https://doi.org/10.1038/nature16069, 2015.

Li, Z., Liu, S., Tan, Z., Bliss, N. B., Young, C. J., West, T. O., and Ogle, S. M.: Comparing cropland net primary production estimates from inventory, a satellite-based model, and a process-based model in the Midwest of the United States, Ecological Modelling, 277, 1–12, https://doi.org/10.1016/j.ecolmodel.2014.01.012, 2014.

510   Luo, Z. K., Wang, E. L., Baldock, J., and Xing, H. T.: Potential soil organic carbon stock and its uncertainty under various cropping systems in Australian cropland, Soil Research, 52, 463–475, https://doi.org/10.1071/Sr13294, 2014.

Luo, Z. K., Feng, W. T., Luo, Y. Q., Baldock, J., and Wang, E. L.: Soil organic carbon dynamics jointly controlled by climate, carbon inputs, soil properties and soil carbon fractions, Global Change Biology, 23, 4430–4439, https://doi.org/10.1111/gcb.13767, 2017.

Maillard, E. and Angers, D. A.: Animal manure application and soil organic carbon stocks: a meta-analysis, Global Change Biology, 20, 666–679, https://doi.org/https://doi.org/10.1111/gcb.12438, 2014.

515   Manzoni, S. and Porporato, A.: Soil carbon and nitrogen mineralization: Theory and models across scales, Soil Biology and Biochemistry, 41, 1355–1379, https://doi.org/http://dx.doi.org/10.1016/j.soilbio.2009.02.031, 2009.

Metting, F. B., Smith, J. L., Amthor, J. S., and Izaurralde, R. C.: Science needs and new technology for increasing soil carbon sequestration, Climatic Change, 51, 11–34, https://doi.org/10.1023/a:1017509224801, 2001.

Minty, B., Franklin, R., Milligan, P., Richardson, M., and Wilford, J.: The radiometric map of Australia, Exploration Geophysics, 40, 325–333, https://doi.org/10.1071/Eg09025, 2009.

520   Ogle, S. M., Breidt, F. J., Easter, M., Williams, S., Killian, K., and Paustian, K.: Scale and uncertainty in modeled soil organic carbon stock changes for US croplands using a process-based model, 16, 810–822, https://doi.org/10.1111/j.1365-2486.2009.01951.x, 2010.

O'Rourke, S. M., Angers, D. A., Holden, N. M., and McBratney, A. B.: Soil organic carbon across scales, Global Change Biology, 21, 3561–3574, https://doi.org/doi:10.1111/gcb.12959, 2015.

525   Parton, W. J., Schimel, D. S., Cole, C. V., and Ojima, D. S.: Analysis of factors controlling soil organic matter levels in Great Plains grasslands, Soil Science Society of America Journal, 51, 1173–1179, https://doi.org/10.2136/sssaj1987.03615995005100050015x, 1987.

Paul, K. I. and Polglase, P. J.: Calibration of the RothC model to turnover of soil carbon under eucalypts and pines, Australian Journal of Soil Research, 42, 883–895, https://doi.org/10.1071/Sr04025, 2004.

Paustian, K., Levine, E., Post, W. M., and Ryzhova, I. M.: The use of models to integrate information and understanding of soil C at the
530   regional scale, Geoderma, 79, 227–260, https://doi.org/10.1016/S0016-7061(97)00043-8, 1997.

Paustian, K., Larson, E., Kent, J., Marx, E., and Swan, A.: Soil C sequestration as a biological negative emission strategy, Frontiers in Climate, 1, https://doi.org/10.3389/fclim.2019.00008, 2019.

Post, W. M. and Kwon, K. C.: Soil carbon sequestration and land-use change: Processes and potential, Global Change Biology, 6, 317–327, 2000.

535   Quinlan, J. R.: Learning with continuous classes. Proceedings of the 5th Australian Joint Conference On Artificial Intelligence, pp. 343–348, 1992.

Richards, G. P. and Evans, D. M. W.: Development of a carbon accounting model (FullCAM Vers. 1.0) for the Australian continent, Australian Forestry, 67, 277–283, https://doi.org/10.1080/00049158.2004.10674947, 2004.

Robertson, A. D., Paustian, K., Ogle, S., Wallenstein, M. D., Lugato, E., and Cotrufo, M. F.: Unifying soil organic matter formation and
540   persistence frameworks: the MEMS model, Biogeosciences, 16, 1225–1248, https://doi.org/10.5194/bg-16-1225-2019, 2019.

Roxburgh, S. H., Berry, S. L., Buckley, T. N., Barnes, B., and Roderick, M. L.: What is NPP? Inconsistent accounting of respiratory fluxes in the definition of net primary production, Functional Ecology, 19, 378–382, https://doi.org/10.1111/j.1365-2435.2005.00983.x, 2005.

Scharlemann, J. P. W., Tanner, E. V. J., Hiederer, R., and Kapos, V.: Global soil carbon: understanding and managing the largest terrestrial carbon pool, Carbon Management, 5, 81–91, https://doi.org/10.4155/Cmt.13.77, 2014.

545 Seidel, S. J., Palosuo, T., Thorburn, P., and Wallach, D.: Towards improved calibration of crop models – Where are we now and where should we go?, European Journal of Agronomy, 94, 25–35, https://doi.org/https://doi.org/10.1016/j.eja.2018.01.006, 2018.

Shi, Z., Crowell, S., Luo, Y. Q., and Moore, B.: Model structures amplify uncertainty in predicted soil carbon responses to climate change, Nature Communications, 9, 11, https://doi.org/10.1038/s41467-018-04526-9, 2018.

SILO: Australian Climate Database, the Queensland Department of Environment and Science, https://legacy.longpaddock.qld.gov.au/silo, 550 2020.

Skjemstad, J. and Spouncer, L.: Integrated Soils Modelling for the National Carbon Accounting System, National Carbon Accounting System Technical Report No. 36, Tech. rep., The Australian Greenhouse Office, 2003.

Skjemstad, J. O., Spouncer, L. R., Cowie, B., and Swift, R. S.: Calibration of the Rothamsted organic carbon turnover model (ROTH C ver. 26.3), using measurable soil organic carbon pools, Australian Journal of Soil Research, 42, 79–88, https://doi.org/10.1071/sr03013, 2004.

555 Smith, P., Smith, J., Powlson, D., McGill, W., Arah, J., Chertov, O., Coleman, K., Franko, U., Frolking, S., Jenkinson, D., Jensen, L., Kelly, R., Klein-Gunnewiek, H., Komarov, A., Li, C., Molina, J., Mueller, T., Parton, W., Thornley, J., and Whitmore, A.: A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments, Geoderma, 81, 153–225, https://doi.org/10.1016/S0016-7061(97)00087-6, 1997.

Smith, P., Soussana, J.-F., Angers, D., Schipper, L., Chenu, C., Rasse, D. P., Batjes, N. H., van Egmond, F., McNeill, S., Kuhnert, M., 560 Arias-Navarro, C., Olesen, J. E., Chirinda, N., Fornara, D., Wollenberg, E., Álvaro Fuentes, J., Sanz-Cobena, A., and Klumpp, K.: How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal, Global Change Biology, 26, 219–241, https://doi.org/10.1111/gcb.14815, 2020.

Sparling, G.: Ratio of microbial biomass carbon to soil organic carbon as a sensitive indicator of changes in soil organic matter, Soil Research, 30, 195–207, https://doi.org/https://doi.org/10.1071/SR9920195, 1992.

565 Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G., and Allison, S. D.: Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations, Biogeosciences, 10, 1717–1736, https://doi.org/10.5194/bg-10-1717-2013, 2013.

Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Running, S. W., Zhao, M. S., Costa, M. H., Kirschbaum, A. A., Ham, J. M., Saleska, S. R., and Ahl, D. E.: Evaluation of MODIS NPP and GPP products across multiple biomes, Remote Sensing of Environment, 102, 570 282–292, https://doi.org/10.1016/j.rse.2006.02.017, 2006.

Unkovich, M., Baldock, J. A., and Farquharson, R.: Development of comprehensive time-series datasets of crop and pasture type and management, for the Australian continent, relevant to the National Greenhouse Gas Inventory carbon accounting procedures, Tech. rep., CSIRO, Australia, 2017.

Unkovich, M., Baldock, J., and Farquharson, R.: Field measurements of bare soil evaporation and crop transpiration, and 575 transpiration efficiency, for rainfed grain crops in Australia – A review, Agricultural Water Management, 205, 72–80, https://doi.org/10.1016/j.agwat.2018.04.016, 2018.

Vereecken, H., Schnepf, A., Hopmans, J. W., Javaux, M., Or, D., Roose, D. O. T., Vanderborght, J., Young, M. H., Amelung, W., Aitkenhead, M., Allison, S. D., Assouline, S., Baveye, P., Berli, M., Bruggemann, N., Finke, P., Flury, M., Gaiser, T., Govers, G., Ghezzehei, T., Hallett, P., Franssen, H. J. H., Heppell, J., Horn, R., Huisman, J. A., Jacques, D., Jonard, F., Kollet, S., Lafolie, F., Lamorski, 580 K., Leitner, D., McBratney, A., Minasny, B., Montzka, C., Nowak, W., Pachepsky, Y., Padarian, J., Romano, N., Roth, K., Roth-

fuss, Y., Rowe, E. C., Schwen, A., Simunek, J., Tiktak, A., Van Dam, J., van der Zee, S. E. A. T. M., Vogel, H. J., Vrugt, J. A., Wohling, T., and Young, I. M.: Modeling soil processes: Review, key challenges, and new perspectives, Vadose Zone Journal, 15, 1–57, https://doi.org/10.2136/vzj2015.09.0131, 2016.

Viscarra Rossel, R. A.: Fine-resolution multiscale mapping of clay minerals in Australian soils measured with near infrared spectra, Journal of Geophysical Research-Earth Surface, 116, https://doi.org/10.1029/2011jf001977, 2011.

Viscarra Rossel, R. A. and Webster, R.: Predicting soil properties from the Australian soil visible-near infrared spectroscopic database, European Journal of Soil Science, 63, 848–860, https://doi.org/10.1111/j.1365-2389.2012.01495.x, 2012.

Viscarra Rossel, R. A., Webster, R., Bui, E. N., and Baldock, J. A.: Baseline map of organic carbon in Australian soil to support national carbon accounting and monitoring under climate change, Global Change Biology, 20, 2953–2970, https://doi.org/10.1111/gcb.12569, 2014.

Viscarra Rossel, R. A., Chen, C., Grundy, M. J., Searle, R., Clifford, D., and Campbell, P. H.: The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project, Soil Research, 53, 845–864, https://doi.org/10.1071/SR14366, 2015.

Viscarra Rossel, R. A., Lee, J., Behrens, T., Luo, Z., Baldock, J., and Richards, A.: Continental-scale soil carbon composition and vulnerability modulated by regional environmental controls, Nature Geoscience, 12, 547–552, https://doi.org/10.1038/s41561-019-0373-z, 2019.

Wang, G. C., Luo, Z. K., Han, P. F., Chen, H. S., and Xu, J. J.: Critical carbon input to maintain current soil organic carbon stocks in global wheat systems, Scientific Reports, 6, https://doi.org/10.1038/srep19327, 2016.

Wang, Y. P. and Barrett, D. J.: Estimating regional terrestrial carbon fluxes for the Australian continent using a multiple-constraint approach I. Using remotely sensed data and ecological observations of net primary production, Tellus B: Chemical and Physical Meteorology, 55, 270–289, https://doi.org/10.3402/tellusb.v55i2.16706, 2003.

Wieder, W. R., Grandy, A. S., Kallenbach, C. M., and Bonan, G. B.: Integrating microbial physiology and physio-chemical principles in soils with the MIcrobial-MIneral Carbon Stabilization (MIMICS) model, Biogeosciences, 11, 3899–3917, https://doi.org/10.5194/bg-11-3899-2014, 2014.