A nationally scalable approach to simulating soil organic carbon in agricultural landscapes

K.M. Coelli\textsuperscript{a,c}, S.B. Karunaratne\textsuperscript{b}, J.A. Baldock\textsuperscript{c}, S. Ugbaje\textsuperscript{a,d}, A.J.V. Buzacotta\textsuperscript{a}, P. Filippi\textsuperscript{a}, S. Cattle\textsuperscript{a} and T.F.A. Bishop\textsuperscript{a,d}

\textsuperscript{a} School of Life and Environmental Sciences, The University of Sydney, Sydney, New South Wales, \textsuperscript{b} CSIRO Agriculture and Food, Black Mountain, Australian Capital Territory, \textsuperscript{c} CSIRO Agriculture and Food, Adelaide, South Australia, \textsuperscript{d} Sydney Informatics Hub, The University of Sydney, Sydney, New South Wales, CSIRO Data61, Brisbane, Queensland

Email: kate.coelli@sydney.edu.au

Abstract: The world's soil provides a large potential sink to sequester carbon, but in order to incentivise increases in soil organic carbon (SOC) and for countries to use increases in SOC to meet their international greenhouse gas emission commitments, we need to accurately estimate changes in SOC. Statistical and process-based models are often used to estimate these changes, and Australia's greenhouse gas reporting framework uses Full Carbon Accounting Model (FullCAM). The dynamics of soil carbon within the FullCAM modelling framework is quantified using the Rothamstead Carbon Model (RothC). The model's ability to estimate change of SOC with time depends on (a) accurate estimation of the key drivers of model inputs and (b) the fate of those drivers on dynamics of SOC through model parameter calibration. For instance, SOC estimates could be improved with more accurate estimation of the plant residues entering the soil. Currently, the reporting framework uses crop type information based on broad spatial supports such as Statistical Area 2 (SA2) regions to derive plant residue inputs for RothC. This is a much coarser spatial resolution than the true spatial variability of plant residues.

In this work we present an application of RothC with plant residue inputs derived from freely available remotely sensed evapotranspiration (ET) data, NDVI and a land use specific scaling parameter. This is demonstrated using two case studies: Muttama Creek - a 1025 km\textsuperscript{2} catchment and Hillston: a 2650 km\textsuperscript{2} district, both in New South Wales (NSW). The Muttama Creek catchment is predominately dryland cropping and grazing land uses, while Hillston is a semi-arid irrigated cotton-growing district.

The model is initialised using SOC fractions, predicted from the spectral library built under an Australian-wide soil sampling project (SCaRP - Baldock et al., 2013a) and has an improved water balance model which uses remotely sensed ET (Wimalathunge and Bishop, 2019) to represent plant water use. This is an improvement on the water balance model built into RothC which uses evaporation to estimate plant water use.

Each catchment consists of two temporal surveys. While Hillston demonstrates a longer timeframe (2002-2015) to observe changes in SOC between surveys, the timeframe of the Muttama Creek (2013-2019) study is similar to the time between repeat sampling rounds of a soil carbon project, under the Emissions Reduction Fund (ERF - the Australian Government’s carbon crediting program). In soil carbon projects under the ERF, change in SOC from the project must be reported to the Government at least every five years, for 25 years (according to a strict set of reporting guidelines). To receive carbon credits after their first report, there must be an observable change in SOC between sampling rounds.

This study shows potential to scale ET and NDVI information according to land use to estimate plant biomass carbon inputs (C inputs) to use as an input to the RothC model. Across the two catchments used to demonstrate this approach, 77 sites, sampled at two time points were used to calibrate the model according to the four land use classes tested in this study. When the simulated SOC results were compared to the observed values for the second sampling time, the cropping systems (Lin’s concordance correlation coefficient (LCCC) = 0.61) performed the best, followed by native grazing systems (LCCC = 0.46), modified pasture systems (LCCC = 0.33) and irrigated cropping systems (LCCC = 0.11).

Keywords: Next generation soil carbon models, process-based model, RothC, soil carbon model, integration of data streams
1. INTRODUCTION

Understanding the magnitude of changes in soil organic carbon (SOC) is important for several reasons. In agriculture, as soils with low SOC are less productive (McBratney et al., 2014) farmers are interested in understanding their SOC levels and how they change, to better manage their soils and crops. In Australian domestic climate change policy, Government incentivise carbon sequestration through the Emissions Reduction Fund (ERF), where land managers can be paid for increasing their SOC levels in agricultural systems (Australian Government, 2020b). With accurate and cost-effective measurement techniques, this can contribute to climate change and drought mitigation as well providing social, environmental, and economic benefits (van Dijk et al., 2013) including an additional and diversified income stream to farmers in a changing climate.

Although two ERF methods already exist for soil carbon sequestration, the Australian Government has identified soil carbon as one of the five ERF method development priorities, where a new method will be developed to reduce the cost of soil carbon measurements (Australian Government, 2020a). Under this new Government initiative, a variety of technologies will be considered to reduce the cost of soil carbon measurement where SOC models will pay a pivotal role.

In international climate change policy, many countries across the world have agreed to reduce their net greenhouse gas emissions, including through increasing their SOC levels (Paris Agreement, 4 per mille, Kyoto Protocol). To demonstrate compliance with these agreements, countries, including Australia have carbon inventories where sources and sinks of carbon emissions are tracked. Therefore, SOC monitoring, reporting and verification (MRV) is a key aspect of SOC science. Australia uses a range of data and modelling approaches to track carbon fluxes across all sectors of the economy (Australian Government, 2020c). For the land sector, the FullCAM model is used to estimate carbon fluxes (Australian Government, 2020c) which incorporates the RothC model to estimate the soil carbon sequestration component (Paul and Roxburgh, 2017).

RothC uses simple, readily available inputs to model a monthly time-step change in topsoil SOC over a period of years to centuries (Coleman and Jenkinson, 1996). To do this, C inputs enter the soil and turnover and movement between four active carbon pools is simulated. There is also an additional inert organic matter (IOM) pool which remains stable over time (Figure 1). The SOC pools defined in the RothC model are conceptual.

Both accurate estimation of key drivers of model inputs and knowledge of their impact on the dynamics of SOC obtained through model parameter calibration are needed for the model to accurately estimate change of SOC through time. Some key model inputs include initial fractions of SOC (Nemo et al., 2017) as well as monthly estimates of organic C inputs entering the system (Keel et al., 2017; Ludwig et al., 2010). While Skjemstad et al. (2004) (and later Zimmermann et al. (2007)) have shown that measurable SOC fractions can be used to replace the conceptual pools in RothC to initialise the model, the most common way to initialise RothC is to run the model in equilibrium mode (spin-up) to derive initial starting pools of soil carbon (Cagnarini et al., 2019; Coleman and Jenkinson, 1996; Smith et al., 2005). In addition to directly influencing the final total

Figure 1. Structure of Roth-C model - adapted from Coleman and Jenkinson (1996) and replacement of conceptual pools with measurable fractions to initialise (Baldock et al., 2013b; Skjemstad et al., 2004) the model in the landscape.
SOC simulated by RothC, an accurate reflection of C inputs could also improve model initialisation. A study by Nemo et al. (2017), demonstrated that while the spin-up method does not always produce realistic results, adjusting C inputs during spin-up could produce better results.

In addition, in agricultural landscapes where land uses vary in space and time (e.g. pastures integrated into cropping rotations), we cannot assume uniform C inputs in space and time. The Australian greenhouse gas reporting framework uses crop type information based on broad Statistical Area 2 (SA2) regions to derive plant residue inputs for RothC. This is a much coarser spatial resolution than the true spatial variability of plant residues. One solution is to make use of freely available, remotely sensed information such as ET and NDVI.

The objectives of this work are to develop a nationally scalable approach to deriving C inputs so that RothC can be used to simulate SOC changes in agricultural landscapes where on farm data is limited.

2. MATERIALS AND METHODS

2.1. Soil sample collection and analysis

For this study, data was collected in two catchments across New South Wales (NSW). Timing, catchment name and land use classification of the sample locations are outlined in Table 1. The soil samples from Muttama Creek were analysed for SOC via dry combustion approach using a LECO C-144 carbon analyser, where any samples containing inorganic carbon were treated with H$_2$SO$_3$ to remove any inorganic carbon, following the process outlined in (Baldock et al., 2013a). Soil samples that were not analysed using the aforementioned methods were predicted using mid infrared spectroscopy (MIR) and partial least squares regression (PLSR) and the spectral library according to the approach outlined in Baldock et al. (2013a). The SOC for samples from Hillston were measured using a combination of dry combustion (2002) and wet oxidation (2015) techniques (Filippi et al., 2018).

While we acknowledge the potential difference in SOC for the 2015 Hillston samples which were analysed using wet oxidation, rather than dry combustion, we did not alter these measurements using a conversion factor (Skjemstad et al., 2000). This was for a few reasons, including the questionable usefulness of universal corrections (Lettens et al., 2007), and that these samples were measured using the modified Walkley & Black method, where samples are heated to minimize incomplete oxidation (Schumacher, 2002).

The measurable fractions of SOC (particulate organic carbon (POC), resistant organic carbon (ROC) and humic organic carbon (HOC)) were predicted using the process outlined by Baldock et al. (2013a). SOC and fractions were converted to stocks (t/ha) using bulk density (BD) derived from a global pedo-transfer function (Tranter et al., 2007). For the sand value required in the pedo-transfer function developed by Tranter et al. (2007), measured using particle size analysis was used if available, otherwise it was predicted using MIR and PLSR.

Table 1. Sampling rounds and number of temporally paired sites by land use type

<table>
<thead>
<tr>
<th>Catchment</th>
<th>Baseline Sampling (t0) Year</th>
<th>Subsequent Sampling (t1) Year</th>
<th>Cropping</th>
<th>Modified pasture</th>
<th>Grazing native veg</th>
<th>Irrigated cropping</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hillston</td>
<td>2002</td>
<td>2015</td>
<td>4</td>
<td>3</td>
<td>14</td>
<td>36</td>
<td>57</td>
</tr>
<tr>
<td>Muttama Creek</td>
<td>2013</td>
<td>2019</td>
<td>11</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>20</td>
</tr>
</tbody>
</table>

2.2. Model overview and description

The Rothamsted Carbon model (RothC) is a multi-pool process-based model used to model the turnover of organic carbon. It uses a monthly timestep and readily available inputs to predict carbon at a given timepoint (t$_0$) based on carbon stocks in another timepoint (t$_0$) and carbon turnover processes, on a years to centuries timescale (Coleman and Jenkinson, 1996). In this model, carbon in the system is cycled through five conceptual carbon pools over time (Figure 1). In this study, we simulate SOC change at the sites outlined in Table 1. The model is calibrated according to land use type to derive a land use specific scaling parameter for the calculation of C inputs for the model.

2.3. Model initialisation

The first step in the modelling process is to initialise the model, or obtain starting values of these five carbon pools. In this work, the model was initialised by replacing three of the five conceptual pools with measured values of carbon fractions (Figure 1). The remaining two pools (BIO and DPM) were set as zero as they will quickly stabilize at low values (Skjemstad et al., 2004). This approach to model initialisation has been
commonly used since it was proposed by Skjemstad et al. (2004) (Zimmermann et al., 2007, Baldock et al., 2013b, Herbst et al., 2018).

2.4. Model inputs and adjustments

In addition to the information required for model initialisation, RothC requires an estimate of plant residue C entering the soil system, as well as climate and soil information to determine the rate of organic carbon decomposition at each location. Some of this information is used to calculate soil moisture, which is a key driver of decomposition. While the RothC model has an integrated water balance model to account for soil moisture, this model is simplistic in the way it represents plant water use. Wimalathunge and Bishop (2019) have developed a water balance model which uses remotely sensed ET to represent plant water use. In this study, this improved water balance model has been used in place of the default water balance model in RothC.

2.5. Calculating plant biomass carbon inputs for RothC

A measure of the carbon entering the soil system from plant residue C (C inputs (t/ha)) is arguably the most important input value into the RothC model. In the context of modelling SOC change using RothC, we are interested in the plant biomass remaining on site and later convert into plant residue C. One way to represent this in a cropping system is;

\[ C \text{ inputs} = 0.44 \times T \times WUE \times \left( 1 - \frac{HI}{HI} \right) \times 0.001, \]  

where 0.44 is a factor to convert total biomass to biomass carbon (assuming 44% of plant material is carbon), T is transpiration (mm), WUE is water use efficiency (kg yield/mm transpiration) and HI is harvest index which is the ratio of yield to total above ground biomass. 0.001 is a unit conversion from kg to tonnes. If crop type is known, WUE and harvest index could be estimated from values in the literature.

Transpiration is a useful measure of plant water use and total plant biomass growth. As total ET is more readily available from remotely sensed products, in this work we calculate transpiration from a MODIS-derived ET product (MOD16A2 – aggregated to monthly total). We use a normalised difference vegetation index (NDVI-derived from Landsat images and aggregated to monthly mean values) to partition transpiration from evaporation and a parameter, a, to scale the NDVI and ET(mm) relationship between land uses;

\[ C \text{ inputs} = 0.44 \times a \times NDVI \times ET \times WUE \times \frac{1 - HI}{HI} \times 0.001. \]  

(2)

Since the approach we present in this study outlines an option to use in the absence of crop type information, we need to estimate parameters for WUE and HI. We have represented this using a combined parameter, b, which is a fitted parameter reflecting the NDVI and ET relationship, WUE and HI;

\[ C \text{ inputs} = 0.44 \times b \times NDVI \times ET. \]  

(3)

Although we are unable to estimate WUE and HI specific to crop type, the spatially and temporally dynamic nature of the remotely sensed products (NDVI and ET) should allow some of this variability to be captured in the estimate of plant biomass remaining on site. The fitted parameter also gives more flexibility to estimating plant biomass in systems other than cropping systems. For example, because harvest index is not required as a specific input, the proportion of biomass removed from a grazing system by animals grazing, could be captured in the combined land use specific parameter term, b.

To derive the b parameter for each land use, the RothC model was run using the ‘optim’ function in base R with RMSE as the objective function. During the optimisation process, the decomposition rate constants (Table 2) for each of the pools were also allowed to vary. In RothC, each of the active soil carbon pools (decomposable plant material (DPM), resistant plant material (RPM), microbial biomass (BIO) & humified organic matter (HUM) - Figure 1) decompose according to first order kinetics (Smith et al., 2020), where decomposition rate is influenced by temperature, soil moisture, soil cover and a pool specific decomposition rate constant (Coleman and Jenkinson, 1996). Even though Skjemstad et al. (2004) proposed halving the RPM decomposition rate constant (Table 2) for use in some circumstances to account for Australian conditions, the rate constants were still developed under European conditions at long term field experiments (Jenkinson et al., 1987; Jenkinson et al., 1992), and are unlikely to reflect decomposition in Australian agricultural systems. The approach outlined here allows remotely sensed plant information, and the fate of those C inputs entering to soils to be calibrated simultaneously rather than considering the soil system separately from the C inputs.
3. RESULTS AND DISCUSSION

In this study we tested the potential of using satellite-derived products to estimate C inputs and calibrated the model parameters according to land use type across two catchments. While the model performance could be improved, this study demonstrates the potential of calibrating SOC models across various spatial scales with freely available satellite-derived information and land use data.

![Graph showing observed SOC vs simulated SOC for different land uses](image)

**Figure 2.** Observed SOC at a subsequent sampling round vs simulated SOC between a baseline sampling round and a subsequent sampling round at a) cropping sites, b) grazing native vegetation sites c) modified pasture sites and, d) irrigated cropping sites across both Muttama and Hillston catchments

Of the four land use types tested in this study, cropping performed the best, with the highest Lin’s concordance correlation coefficient (LCCC) and the lowest root mean square error (RMSE) (Figure 2a). The grazing systems did not do very well at predicting SOC (Figure 2b and c). It is possible that accounting for the stocking rates in grazing systems would help to improve these predictions. The way a pasture system is managed, such as stocking rates and rotation systems, could influence the quantity of C inputs entering the soil carbon cycle in several ways. This is demonstrated by the introduction of the concept of safe “Livestock carrying capacities” over half a century ago – a concept which is used to estimate the appropriate stocking rate to maintain a pasture resource (e.g. Condon, 1968; Johnston et al., 1996).

As stocking rates increase, traffic over the pasture is increased and plant growth could become limited due to increased soil compaction and direct damage to the pasture. However, a balance must be reached, as studies have shown that grazing can increase pasture growth, carbon allocation and soil carbon (e.g. Macdonald et al., 2008; Wilson et al., 2018). Secondly, in grazing systems, another source of carbon entering the system is manure. Variations in this between sites would also be captured if the stocking rate was reflected in the model.

The results for irrigated cropping land use (Figure 2d), indicate that the approach outlined in this study may not be applicable in these systems. The predictions of SOC were the worst of all four land uses tested in this study, likely to be a result of flood irrigation. The flood irrigation at these sites means that for part of each year they were inundated by water. As RothC models the turnover of SOC in non-waterlogged topsoils (Coleman and Jenkinson, 1996), it assumes aerobic conditions and as such, it is possible the true decomposition process that occurred at those sites was not accurately reflected by the model. In addition, there was no site-specific information available on the irrigated water volume, which, if available, could have been converted to mm water and added to site specific rainfall. This could have improved the model simulation quality.
As outlined in Table 2, the decomposition rate constants were higher than, or equal to the default RothC decomposition rate constants in all pools and land uses except two. Larger decomposition rate constants indicate faster decomposition meaning that the model is, in general, attempting to quickly add carbon from plant material to the SOC pool. The land use specific scaling parameter, $b$, varies substantially between land use types. As this parameter scales transpiration, by land use type, to give C inputs, a higher $b$ value indicates a higher quantity of C-input entering the soil, per mm of transpiration, given the same NDVI and ET values.

<table>
<thead>
<tr>
<th>LAND USE</th>
<th>b</th>
<th>$K_{DPM}$</th>
<th>$K_{RPM}$</th>
<th>$K_{BIO}$</th>
<th>$K_{HUM}$</th>
<th>DPM:RPM_RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>NA</td>
<td>10</td>
<td>0.15*</td>
<td>0.66</td>
<td>0.02</td>
<td>1.44</td>
</tr>
<tr>
<td>Cropping</td>
<td>27.5</td>
<td>12</td>
<td>3</td>
<td>0.7</td>
<td>0.034</td>
<td>3.67</td>
</tr>
<tr>
<td>Modified pastures</td>
<td>27</td>
<td>10</td>
<td>3</td>
<td>0.66</td>
<td>0.016</td>
<td>2.67</td>
</tr>
<tr>
<td>Grazing native vegetation</td>
<td>49.5</td>
<td>10</td>
<td>3</td>
<td>0.66</td>
<td>0.023</td>
<td>2.94</td>
</tr>
<tr>
<td>Irrigated cropping (Hillston)</td>
<td>1</td>
<td>10</td>
<td>0.1</td>
<td>0.66</td>
<td>0.025</td>
<td>1.44</td>
</tr>
</tbody>
</table>

*This decomposition rate constant has been altered account for Australian conditions (Skjemstad et al., 2004)

4. CONCLUSION

While this work would clearly benefit from being calibrated on a larger dataset, it shows our progress towards developing a nationally scalable approach to simulating SOC in agricultural landscapes using RothC and freely available, readily accessible geospatial data. To improve the models, we need more data with at least two temporal measurements for each spatial observation (temporally paired) to calibrate RothC. As the value of long-term soil monitoring studies becomes more obvious to research groups, land managers and policy makers alike, we are likely to see an increase in temporal SOC datasets. Our approach would allow any temporally paired sites collected across different time periods to be used to improve the model.

A key limitation of this work is that we have not performed independent model validation, due in part to the small sample size. We are in process of performing lab work for a third catchment which has data at two time points, meaning future work could include testing the models (one for each land use type) by leave-one-catchment-out cross validation to check that land use based $b$ values and decomposition rate constants are applicable across different areas, and hopefully at a national scale. Furthermore, as the approach is applicable anywhere in Australia it could be tested at any location where there are repeated measures of soil carbon.

While this case study demonstrates the potential of this approach to estimating plant C biomass in systems that remain under a given land use type for a long period of time, another limitation of this approach is the temporal resolution of the land use maps. Often available at a snapshot in time, and updated every few years, these products may not adequately capture systems that rotate between pasture and cropping systems over time. As this approach is flexible, however, if additional local information is known, the $b$ parameter can be tailored to reflect this. Finally, unless irrigation quantities are available to manually add into monthly rainfall quantities, we do not recommend this approach is applied in cropping systems, as the water balance model means the soil will always be in a dry condition, while in reality it could be waterlogged. In summary, we have integrated an easily scalable plant model component into the RothC, driven by remotely sensed data streams, for more practical application across wide array of soil carbon projects.

ACKNOWLEDGMENTS

This project is part of a PhD undertaken by the first author and funded by the Sydney Institute of Agriculture (SIA), University of Sydney and CSIRO Data61. The authors acknowledge the technical assistance provided by the Sydney Informatics Hub, a Core Research Facility of the University of Sydney. They also acknowledge funding by Department of Industry, Science, Energy and Resources (DISER) to support aspects of the work.

REFERENCES


