Ecological Indicators 88 (2018) 425–438

Contents lists available at ScienceDirect

Ecological Indicators

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

Original Articles

Estimating soil organic carbon stocks using different modelling techniques in the semi-arid rangelands of eastern Australia

Bin Wang[a,*], Cathy Waters[b], Susan Orgill[a], Annette Cowie[c], Anthony Clark[b], De Li Liu[a], Marja Simpson[b], Ian McGowen[b], Tim Sides[a]

[a] NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, NSW 2650, Australia
[b] NSW Department of Primary Industries, Orange Agricultural Institute, NSW 2800, Australia
[c] NSW Department of Primary Industries, Trevenna Rd, Armidale, NSW 2351, Australia

ARTICLE INFO

Keywords:
Soil organic carbon stocks
Land fractional cover data
Remote sensing
Machine learning
Genetic algorithm
Stepwise regression

ABSTRACT

Soil organic carbon (SOC) is pivotal for biological, chemical and physical processes and provides vital information on changes in soil fertility and land degradation. Rangelands, accounting for about 81% of Australian land area, are significant carbon (C) stores and small increases in soil C sequestration over such a vast area represents a considerable climate change mitigation opportunity. Efficient modelling techniques to evaluate the potential to increase rangeland SOC stocks are vitally important to assess their role in the global carbon cycle and quantum abatement. This study aimed to evaluate boosted regression trees (BRT) and random forest (RF) models in predicting SOC stocks from available continuous remotely sensed variables using two feature selection techniques. Dominant variables that affect SOC stocks in the rangelands were also identified. Using field-based measurements of SOC stock collected from 564 data points across the study area and 28 GIS-based environmental variables including climate, topography, radiometry, vegetation and land fractional cover data, we employed stepwise regression (SR, linear approach) and genetic algorithm (GA, nonlinear approach) to select the most informative variables. These selected predictors were then used to train the BRT and RF models. In all, four models were evaluated: BRT using SR (SR_BRT); RF using SR (SR_RF); BRT using GA (GA_BRT) and RF using GA (GA_RF). In addition, BRT using all predictors (All_BRT) and the RF using all predictors (All_RF) were used as benchmarks to test the performance of the four models. Of the field-based data, 75% were used to train the model ("calibration dataset") and the remaining 25% were used to validate the prediction of SOC stocks ("validation dataset"). The results indicate that the RF exhibited a better performance in predicting SOC stocks than the BRT regardless of input variables. In addition, we verified that feature selection for both machine learning techniques is necessary for estimating SOC stocks because they can increase accuracy and save time. The GA_RF was the most reliable method to predict SOC stocks, with the lowest root mean square error (RMSE) and the highest R^2 values (7.44 Mg C ha^{-1} and 0.48, respectively), suggesting that the method of using GA-RF to generate a predictive model from measured data and remotely-sensed variables may provide a cost effective alternative to direct sampling to predict SOC stocks in the semi-arid rangelands of eastern Australia. The important variables for explaining the observed SOC stocks were rainfall, elevation, Prescott index (PI, a measure of water balance), and land fractional cover (bare ground fraction). The approach proposed here can be extended in areas where field observed data is scarce (e.g. rangelands) to produce more detailed information about SOC stocks. As such, the results of our study are of particular importance in Australian rangelands to provide a statistical and theoretical basis for producing digital SOC stock maps based on readily available remotely-sensed data, with potential for use in similar rangelands conditions internationally.

1. Introduction

Approximately half of the global land mass is occupied by rangelands including grasslands, shrub lands, deserts and tundra, which contain more than a third of above- and below-ground carbon (C) reserves (Schuman et al., 2002). Rangelands extend across low rainfall and variable climates, accounting for about 81% of Australian land area (http://www.environment.gov.au/land/rangelands) (Allen et al., 2013). It is estimated that Australia's rangeland soils contain between 34 and 48 Gt of carbon, representing a sequestration potential of 78 Mt

* Corresponding author.
E-mail address: bin.a.wang@dpi.nsw.gov.au (B. Wang).

https://doi.org/10.1016/j.ecolind.2018.01.049
Received 29 August 2017; Received in revised form 24 January 2018; Accepted 24 January 2018

Table 1
Summary of field surveys in south eastern Australia where SOC stocks (0–0.30 m) were compared under different land use and management. Note: there is some overlap in these studies due to the same data being used in multiple papers for different comparisons. Abbreviations include: average annual rainfall (AAR), soil pH in water (pH_w), Colwell phosphorus (Col P) and vapour pressure deficit (VPD).

Agricultural system and management comparison	Sites (total)	Significant difference (Y/N)	Comparisons where difference in OC stock were significant	Other significant factors influencing OC stock	Location and reference
Rain-fed crop and pasture					
Crop: continuous, rotation, pasture croppingPasture (perennial) : low *vs* high nutrient input Grazing: continuous *vs* rotational	354	N		Climate (AAR), elevation, silica content (texture)	NSW Badgery et al. (2013)
Crop: continuous, rotation, pasture cropping Pasture: introduced perennial	223	Y	Pasture > crop (continuous and rotation)	Climate (AAR), soil properties (silica content, pH_w, gravel, Col P) and pasture composition	NSW Badgery et al. (2014)
Pasture: introduced perennial *vs* native, perennial *vs* annual, phosphate applied *vs* no phosphate, pasture cropping *vs* control (crop or pasture) Grazing: continuous *vs* rotational grazing	44	Y	Phosphate applied > no phosphate (pasture)	Climate (AAR), previous land use	NSW and Victoria Chan et al. (2010)
Crop: continuous, rotationPasture: introduced perennial, native (voluntary mixed pastures)	200	Y	Pasture > crop[1]	Climate, topography and soil properties (EC, pH, Col P, CEC, exchangeable sodium (Na), potassium (K), calcium (Ca),magnesium (Mg) and aluminium (Al))	NSW Davy and Koen (2013)
Crop: continuous, rotation, stubble and tillage management Pasture: introduced perennial *vs* annual or mixed Grazing: continuous *vs* rotational	615	N		Climate (AAR VPD, temperature) and soil properties (silica content)	Victoria (Robertson et al., 2016)
Crop: continuousPasture: introduced perennial (tropical) and native	145	Y	Pasture > crop[2]	Climate, soil type and duration of land use	NSW Schwenke et al. (2013)
Crop: rotation (with cultivation) Pasture: introduced perennial, native pasture Remnant woodland	20	Y	Woodland > non-woodland	Soil type	NSW Wilson et al. (2011)
Crop: continuousPasture: introduced perennial (grass) or lucerne Grassy woodland	22[3]	Y	Non-crop > crop	Soil type	NSW Young et al. (2005)
Rain-fed crop and pasture, irrigated crop (cotton) or pasture					
Pasture: introduced perennial (species comparison), native, inputs: irrigated with dairy effluent *vs* mineral fertiliser Native forest	20[3]	Y	Introduced > native pasture Introduced pasture irrigated with dairy effluent > mineral fertiliser	Climate (AAR), previous land use	NSW Chan and McCoy (2010)
Crop: continuous, rotation, cotton (irrigated) Pasture: introduced perennial, native Grazing: low intensity, continuous, rotational Wooded: native forest, softwood plantation	1401	Y	Native forest and grass > introduced pasture ≅ grazing systems > softwood plantation ≅ crop rotation ≅ minimum till > cropping (tillage) > irrigated cotton ≅ continuous cropping	Climate and soil properties	NSW Hobley et al. (2015)
Crop: rotation, tillage and stubble management, cotton (irrigated) Pasture: introduced perennial, native Grazing: continuous *vs* rotational Organic amendments	780	Y	Pasture and crop rotation > continuous crop	Climate (AAR), longitude, elevation, soil pH and inputs (N fertiliser, soil conditioner, organic amendments)	NSW Rabbi et al. (2014)

[1] Only significant in Slopes region > 500 mm AAR (Davy and Koen, 2013).
[2] Only significant on Chromosols (Schwenke et al., 2013).
[3] Comparison 0–0.20 m (Chan and McCoy, 2010; Young et al., 2005).