A vegetation carbon isoscape for Australia built by combining continental-scale field surveys with remote sensing

Samantha E.M. Munroe*[1,2], Greg R. Guerin[1,2], Francesca A. McInerney[3,4], Irene Martín-Forés[1,2], Nina Welti[5], Mark Farrell[5], Rachel Atkins[3], & Ben Sparrow[1,2]

[1] School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

[2] Terrestrial Ecosystem Research Network (TERN), University of Adelaide, Adelaide, South Australia 5005, Australia

[3] School of Physical Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

[4] School of Earth and Environmental Sciences, University of Queensland, St. Lucia, QLD 4072, Australia

[5] CSIRO Agriculture and Food, Kaurna Country, Urrbrae, South Australia 5064, Australia

*corresponding author: samantha.munroe@adelaide.edu.au

Author Contributions: SM originally formulated the idea, SM, as well as FM, GG, IM, NW, and BS designed the study and developed the methodology. SM and GG analysed the data. MF & RA contributed data and results validation; SM wrote the manuscript; all other authors provided editorial advice.

Acknowledgements

We acknowledge the TERN Ecosystem Surveillance field team and the support of TERN by the Australian government through the National Collaborative Research Infrastructure Strategy. We thank Peter Scarth and David Summers for their advice on best practices for the remote sensing data in our analysis.

Abstract

Context: Maps of C_3 and C_4 plant abundance and stable carbon isotope values ($\delta^{13}C$) across terrestrial landscapes are valuable tools in ecology to investigate species distribution and carbon exchange. Australia has a predominance of C_4-plants, thus monitoring change in C_3:C_4 cover and $\delta^{13}C$ is essential to national management priorities.

Objectives: We applied a novel combination of field surveys and remote sensing data to create maps of C_3 and C_4 abundance in Australia, and a vegetation $\delta^{13}C$ isoscape for the continent.

Methods: We used vegetation and land-use rasters to categorize grid-cells (100 m^2) into woody (C_3), native herbaceous, and herbaceous cropland (C_3 and C_4) cover. Field surveys and environmental factors were regressed to predict native C_4 herbaceous cover. These layers were combined and a $\delta^{13}C$ mixing model was used to calculate site-averaged $\delta^{13}C$ values.

Results: Seasonal rainfall, maximum summer temperature, and soil pH were the best predictors of C_4 herbaceous cover. Comparisons between predicted and observed values at field sites indicated our approach reliably predicted generalised C_3:C_4 abundance. Southern Australia, which has cooler temperatures and winter rainfall, was dominated by C_3 vegetation and low $\delta^{13}C$ values. C_4-dominated areas included northern savannahs and grasslands.

Conclusions: Our isoscape approach is distinct because it incorporates remote sensing products that calculate cover beneath the canopy, the influence of local factors, and extensive validation, all of which are critical to accurate predictions. Our models can be used to predict C_4:C_3 abundance under climate change, which is expected to substantially alter current C_4:C_3 abundance patterns.

Keywords: Photosynthesis, C_4, C_3, isoscape, carbon

Declarations

- Funding: Primary financial support for this project was provided by the Australian government through the National Collaborative Research Infrastructure Strategy. Additional financial support for this project was from the AMP Foundation and the AMP Tomorrow Maker award presented to SM, and Australian Research Council Future Fellowship (FT110100100793) awarded to FAM.
- Conflicts of interest: All authors declare they have no conflict of interest
- Ethics approval: NA
- Consent to participate: NA
- Consent for publication: All authors have provided consent to publish
- Availability of data and material: All data and relevant materials are available via open access data applications, specifically via the TERN Data Discovery Portal (www.portal.tern.org.au/) or the R package *ausplotsR* (https://cran.r-project.org/web/packages/ausplotsR/index.html)
- Code availability: All analysis was performed in the R environment

Introduction

The spatial patterns of stable carbon isotope ratios ($\delta^{13}C$) across terrestrial landscapes, also known as $\delta^{13}C$ 'isoscapes', are used in a wide range of research applications (West et al. 2009). Most commonly, $\delta^{13}C$ isoscapes are used to study food web dynamics and animal migration (Hobson et al. 2010; Hobson and Wassenaar 2018; Vander Zanden et al. 2018). Animals tissues reflect the $\delta^{13}C$ value of their diet (Ben-David and Flaherty 2012; Kelly 2000; Tieszen et al. 1983). By comparing the carbon isotope ratios of an organism to its environment, we can deduce its likely place of origin (Flockhart et al. 2017; Hobson and Kardynal 2015; López-Calderón et al. 2017). Terrestrial $\delta^{13}C$ ratios can also be used to unravel carbon biogeochemical fluxes (i.e. carbon exchange between the biosphere and atmosphere; Still and Rastogi 2017), fractional plant productivity (Powell et al. 2012) and water use efficiency (Cernusak 2020; Frank et al. 2015). Given their vast utility, creating isoscapes has become a high priority in environmental research.

The primary determinant of average vegetation $\delta^{13}C$ values across terrestrial landscapes is the relative abundance of C_3 and C_4 plants (Still et al. 2003). C_3 plants include cool season grasses, most shrubs, and nearly all trees (Kellogg 2001; Sage 2016), whereas C_4 plants include warm-season grasses, many sedges, and some forbs and shrubs (Sage et al. 2012). The distribution of C_3 and C_4 plants reflects their divergent responses to climate. In hot and dry environments, C_3 plants experience increased rates of oxygen fixation by rubisco (photorespiration), a toxic and energetically expensive process, and diminishing returns in the trade-off between carbon uptake and water loss (Andrews and Lorimer 1987; Sage et al. 2012). In contrast, C_4 plants possess a unique set of adaptations that separate and concentrate CO_2 with rubisco, eliminating photorespiration and increasing productivity in hot and dry conditions (Kanai and Edwards 1999; Sage 2004). As a result, C_3 plants are typically less competitive in warm, arid climates. C_3 and C_4 plants also have a unique range of $\delta^{13}C$ values. Due to their distinct carbon fractionation processes during photosynthesis, the values of C_3 plants range from -37‰ to -20‰ $\delta^{13}C$ (mean= ~-27‰), and the values of C_4 plants range from -12‰ to -16‰ $\delta^{13}C$ (mean=~-13‰; Kohn 2010; O'Leary 1988). Therefore, knowledge of C_3 and C_4 cover can be used to estimate average plant $\delta^{13}C$ across terrestrial environments (Powell et al. 2012; Still and Powell 2010).

Remote sensing capabilities can be used to approximate C_3 and C_4 cover at a continental scale (Griffith et al. 2019; Powell et al. 2012; Still and Powell 2010). Satellite imagery enables the separation of woody (predominantly C_3) and herbaceous (mixed C_3 and C_4) plant cover. Climate masks or models can be used to predict the relative abundance of C_4 and C_3 cover in the herbaceous layer, and the $\delta^{13}C$ values of C_3 and C_4 plants can be applied to extrapolate the mean $\delta^{13}C$ value of vegetation in a given area. Cropland cover must also be considered because the photosynthetic pathway of cropland is dictated by humans, not climate. This technique has been applied to create terrestrial $\delta^{13}C$ isoscapes at the continental scale in Africa and America (Firmin 2016; Powell et al. 2012; Still and Powell 2010), although other continents undergoing profound land-use changes remain unassessed.

Field surveys can greatly enhance the accuracy of $\delta^{13}C$ isoscapes. Vegetation cover data from field surveys can be used to compare different C_4 cover-climate models and determine what approach should be used to predict the relative abundance of C_4 and C_3 herbaceous cover. Numerous models have been proposed to predict relative C_4 herbaceous cover, such as summer maximum temperatures (von Fischer et al. 2008) and seasonal rainfall patterns

(Murphy and Bowman 2007; Winslow et al. 2003). The most commonly employed approach is the physiological temperature crossover model (Collatz et al. 1998; Ehleringer 1978), which predicts C_4 plants will be more abundant in areas where the mean monthly temperature is greater than 22°C. The best approach may vary between regions, therefore selecting the most appropriate model for a specific area is essential for accurate isoscape predictions. Field surveys can also be used to model the modifying effects of local edaphic factors on C_4 cover (Griffith et al. 2015; Nippert and Knapp 2007), which is generally overlooked in large-scale analysis. They can be used to quantify the herbaceous cover under trees, which is often obscured, and thus excluded, from isoscapes built using standard remote sensing tools. Finally, but perhaps most crucially, field surveys can validate remote sensing predictions. Yet, systematic and comparable field surveys that span an entire continent are rare, and existing large-scale isoscapes have been largely constructed without the benefits of ground observations or extensive validation.

Australia is a continent with abundant C_4 vegetation due to the large expanses of C_4 grasslands, shrublands and savannahs (Hattersley 1983; Murphy and Bowman 2007; Sage 2016). Therefore, monitoring and predicting trends in C_4 abundance and $\delta^{13}C$ is important to national management priorities, such as fire modelling (Prober et al. 2007) and projecting changes in C_3 and C_4 abundance due to climate change (Corlett and Westcott 2013; Hasegawa et al. 2018). Despite this, no large-scale estimates of C_3 or C_4 vegetation cover or $\delta^{13}C$ values are available. This represents a significant gap in national research capacity. The Australian Terrestrial Ecosystem Research Network (TERN) is an environmental monitoring program funded through the Australian Government National Collaborative Research Infrastructure Strategy (NCRIS) that observes, records, and measures terrestrial ecosystem parameters and conditions for Australia over time. TERN has developed numerous remote sensing layers that estimate the relative distribution of vegetation cover across the country (see www.tern.org). TERN has also conducted over 700, one ha plot-based vegetation surveys across all major biomes and dryland habitats. These combined resources provide a novel opportunity to advance and validate remote sensing strategies for building large terrestrial isoscapes, and for the first time develop a $\delta^{13}C$ isoscape for Australia.

The goals of this paper were to create mapping products that represent the distribution of C_3 and C_4 vegetation in Australia, and construct a site-averaged vegetation $\delta^{13}C$ isoscape for the continent (including Tasmania) using a unique combination of field surveys and remote sensing tools. To create a terrestrial vegetation $\delta^{13}C$ isoscape, we adapted the methodology pioneered by Still and Powell (2010) and Powell et al. (2012), with key modifications that benefit from Australian ground survey data and advancements in remote sensing. To predict the relative cover of C_3 and C_4 vegetation, we used vegetation and climate rasters to (1) categorize grid-cells (100 m^2) into woody (C_3) and herbaceous (C_3 and C_4) components, (2) determine the extent of Australian cropland and assign each crop a photosynthetic type (i.e. C_3 or C_4), and (3) apply a % herbaceous C_4 cover~climate and edaphic model to predict proportional (%) C_3 and C_4 herbaceous cover. In contrast to other large-scale isoscapes, TERN remote sensing data and field surveys were used to account for the ground cover fraction beneath the vegetation canopy, and the influence of local-scale factors on C_4 abundance. Once relative C_3 and C_4 vegetation cover layers were generated, we used a $\delta^{13}C$ mixing model to determine the average vegetation $\delta^{13}C$ value in each grid-cell. We also conducted novel accuracy assessments of our final predictions across major vegetation groups and demonstrate the research potential of these data layers with an example of C_4-landscape analysis across all bioregions in Australia. Our results provide an alternative approach to constructing terrestrial $\delta^{13}C$ isoscapes that may better incorporate local-scale controls on C_3:C_4 abundance and enables the prediction of future changes in C_3 and C_4 distribution under various climate change scenarios. This is a critical feature of our methodology, as climate change is anticipated to drastically shift the competitive advantage of C_3 and C_4 plants across the continent.

Methods

Step 1: Estimate *% woody* and *% herbaceous* cover

Our Australian $\delta^{13}C$ vegetation isoscape was constructed using remote sensing vegetation data primarily sourced for the year 2015. Climate conditions in 2015 for Australia were considered average (i.e. not dry or wet), and fire occurrence and intensity were relatively low. This was also one of the most recent years for which exhaustive vegetation data were available. Thus, a 2015 isoscape should be a good representation of modern average conditions in Australia.

To create the isoscape, we adapted the methodology of Still and Powell (2010) and Powell et al. (2012) and partitioned Australian vegetation cover into C_3 and C_4 cover layers (Fig. 1). The *% woody cover* layer was generated from the Seasonal Persistent Green Cover product for Australia (Gill et al. 2017; Gill et al. 2015). This product is derived from Landsat 5 TM, Landsat 7 ETM+ and Landsat 8 OLI images acquired from the United States Geological Survey (USGS) and estimates the proportion (%) of green fractional cover (i.e. the fraction of ground covered by green vegetation) that does not entirely deteriorate within a year (see Supplemental Methods Table 1 for synopsis of all datasets). This primarily consists of woody vegetation (i.e. trees and shrubs). Estimates for Seasonal Persistent Green Cover and projected woody foliage cover (2000-2010) have been validated with field-measurements, providing an R^2 of 0.918 and a root mean square error (RMSE) of 0.070. The overall classification accuracy of the woody vegetation extent is 81.9%. Based on these results, we treated *% woody cover* as the most accurate estimate for any cover product in our analysis.

The *% herbaceous cover* layer was generated from the Seasonal Fractional Ground Cover product for Australia (Trevithick et al. 2014). The Seasonal Fractional Ground Cover product is derived from the Seasonal Fractional Cover time series and the Seasonal Persistent Green Cover product. It consists of three components, (1) % vegetated green (photosynthetically active) ground cover, (2) % vegetated non-green (i.e. non-photosynthetic) ground cover (primarily dead vegetation), and (3) % bare ground. These three components sum to 100%. The Seasonal Fractional Ground Cover is distinct from other remote sensing measures of fractional ground cover because it accounts for vegetation layering. The Seasonal Fractional Ground Cover includes the ground cover fraction that is visible to the satellite (i.e. viewed from above), but also applies a model to account for the ground cover fraction that may grow beneath the vegetation canopy. Essentially, the Seasonal Fractional Ground Cover predicts the ground cover under the canopy that is normally obscured from the view of the satellite. This provides a potentially more accurate representation of 'true' ground cover compared to other remote sensing data. Vegetated green and vegetated non-green ground cover were combined to estimate the total *% herbaceous cover* in each grid-cell. Vegetated non-green ground cover was included in *% herbaceous cover* to account for Australia's highly arid climate and ensure that wide spread senescent vegetation was incorporated into our calculations. Both *% woody* and *% herbaceous cover* predicts vegetation cover at medium resolution (30 m) for each calendar season (3 months) and are freely available from the TERN Landscape Monitoring's Remote Sensing Data Facility. To bring cover data to a scale consistent with the other data products, we resampled all vegetation raster layers to a resolution of 100 m x 100 m per pixel (1 ha). Values from each season were combined to calculate the annual mean *% woody* and *% herbaceous cover* (Fig. 2).

Estimates of Seasonal Fractional Ground Cover were restricted to areas of < 60 *% woody cover* because the model used to estimate the herbaceous cover under trees is not effective in dense forests. TERN plot data indicated in areas where tree cover was > 60%, herbaceous cover was limited and ranged from 0 to 25% (Supplemental Methods Figure 1). This is consistent with other work demonstrating increased canopy cover can reduce herbaceous cover due to reduced light availability in the understory (Cole and Weltzin 2005; Dormann et al. 2020). Therefore, in grid cells with > 60% *woody cover*, *% herbaceous cover* was presumed to be minimal and set to zero (see Supplemental Methods for full justification).

The *% woody cover* layer was designated 100% C_3 vegetation. This introduces a potential source of error because some groups of shrubs, in particular chenopods, may use either C_3 or C_4 photosynthesis (Akhani et al. 1997; Munroe et al. 2020b). However, chenopods are mostly evergreen and are likely largely incorporated into the *% woody cover* fraction (Scarth, personal communication). We were unable to identify an accurate way to distinguish and model C_4 chenopod shrub cover from other woody cover across Australia. Remote sensing does not relate well to chenopod vegetation (O'Neill 1996; Sparrow et al. 1997), and statistical analysis of TERN field plot found proportional C_4 chenopod distribution (relative to C_3) is not closely associated with climate in Australia (Munroe et al. in review). Consequently, we made the simplifying assumption that all woody cover is C_3.

Step 2: Incorporate agro-ecosystems

The photosynthetic pathway of cropland is determined by what type of crop is planted in each area. Therefore, the photosynthetic pathway of crops must be evaluated separately to natural vegetation. To accomplish this, we partitioned *% herbaceous cover* into *% natural herbaceous cover* and *% herbaceous crop cover* layers. This was achieved using the Catchment Scale Land Use of Australia (CLUM) dataset. The CLUM dataset is the most current, nationally consistent compilation of catchment scale land use data for Australia (current as of December 2018). It is a seamless raster dataset that combines land use data for all state and territory jurisdictions at a resolution of 50 metres. The CLUM dataset indicates a single dominant land use type for each grid-cell. Land use is classified according to the Australian Land Use and Management (ALUM) Classification version 8 (ABARES 2016). This dataset identifies cropping land across the country, and includes information on specific commodities (e.g. sugar, rice, cereals). Using CLUM, we determined the geographical extent of herbaceous cropland areas. We assumed that in cropland grid-cells, 100% of the *% herbaceous cover* was crops. Based on this assumption, *% herbaceous cover* was divided into *% natural herbaceous cover* and *% herbaceous crop cover* layers (Fig 3). Using the CLUM dataset, we then determined the likely commodity and photosynthetic type planted at each grid-cell in the *% herbaceous crop cover* layer.

Most identified crops in Australia were C_3 (e.g. wheat, barley, rice). The only specifically identified C_4 commodity was sugarcane. However, the generic ALUM classifications ‘cereal crops’ and ‘crops’, which were the most common and extensive crop designations in the CLUM dataset, may be C_3 or C_4 grain. To assess the likelihood of ‘cereal crops’ and ‘crops’ being C_3 or C_4, we consulted the Australian Bureau of Statistics (ABS), which conducts detailed agricultural censuses that quantify crop area, commodity type, production, and yield data for Australia, each state/territory, and sub-state regions. The most recent relevant agriculture census was for 2015/16 (ABS 2016). According to ABS (2016), the most common C_4 grain crops in Australia are sorghum and maize. Together, sorghum and maize only equalled approximately 2% of the total cropping area (ha) in Australia in 2015. Most sorghum and maize were grown in the so-called 'sorghum belt', which stretches across the southern cropping regions of Queensland and the northern cropping areas of New South Wales. Within this area, sorghum and maize represent less than 15% of the cropping area. In addition, sorghum is often seasonally rotated with wheat. Without more specific information on the cropping locations for sorghum and maize, and given its likely limited land cover in 2015, we determined that unspecified cropland should be assigned 100% C_3. Using these finalised C_3 and C_4 cropland assignments, *% herbaceous crop cover* was subdivided into *% herbaceous C_3 crop cover* and *% herbaceous C_4 crop cover* layers.

Step 3: Assign *% natural herbaceaous cover* layer proprotional C_3 and C_4 values

% natural herbaceous cover includes a mix of C_3 and C_4 plants whose relative abundance is dictated by climate and local environmental conditions. Therefore, to estimate the relative cover of C_3 and C_4 plants in each grid-cell of the *% natural herbaceous cover* layer, we applied a statistical model that accounts for their divergent responses to climate and edaphic factors. We used TERN vegetation survey data to compare various environmental models to identify the most accurate method for predicting proportional (%) herbaceous C_4 vegetation across Australia.

Step 3a. Create a model to predict proportional (%) herbaceous C_4 vegetation cover

We calculated proportional (%) herbaceous C_4 vegetation cover (relative to herbaceous C_3 and C_4 cover) at 700 one-hectare plots systemically surveyed using a point-intercept method by TERN between 2011 and 2019. A full description of TERN plot survey protocols is detailed in the TERN AusPlots Rangeland manual (Sparrow et al. 2020; White et al. 2012). The protocols most relevant to our analysis are documented in the Supplemental methods. TERN plot data were analysed in the R statistical environment (R Core Team 2019) and imported using the 'ausplotsR' package (Guerin et al. 2020; Munroe et al. 2020a), a package which enables the import and analysis TERN plot survey data. Herbaceous species cover (%) was calculated at each TERN plot using the *species_table* function. Species were assigned a photosynthetic pathway using Munroe et al. (2020b). Herbaceous species included the growth forms 'Forb', 'Hummock grass', 'Rush', 'Sedge', and 'Tussock grass'. Proportional herbaceous C_4 cover at TERN plots (Fig 4) was then calculated as a proportion of C_3 and C_4 herbaceous species cover by:

Eq 1 $$\text{Proportional herbaceous } C_4 \text{ cover} = C_4 \text{ herbaceous species cover} / (C_4 \text{ herbaceous species cover} + C_3 \text{ herbaceous species cover})$$

We then compiled a dataset of climatic and edaphic variables (Supplemental Methods Table 3) that are considered potential drivers of C_4 plant distribution (Griffith et al. 2015; Pau et al. 2013; Sage 2004). Climate data layers were sourced from Williams et al. (2010) and edaphic data from Gallant et al. (2018). We also considered the Collatz et al. (1998) crossover temperature model for comparison (Collatz et al. 1998; Ehleringer 1978). Using this approach, a particular month is determined to favour C_4 growth when the mean daytime temperature was > 22 °C and precipitation is ≥ 25 mm, while a particular month is determined to favour C_3 growth when the mean daytime temperature was ≤ 22 °C and precipitation is ≥ 25 mm. However, because large areas of Australia receive < 25 mm of precipitation per month, a traditional crossover approach may not be accurate (Murphy and Bowman 2007). Therefore, to apply the crossover temperature model consistently across the country, we regressed proportional C_4 herbaceous cover against the mean annual proportion of C_4 favoured months (> 22 °C and ≥ 25 mm rainfall), instead of the absolute number of C_4 favoured months (Munroe et al. in review). Climate data for the crossover approach were calculated using 1970–2018 records from the Australian Gridded Climate Data set (Bureau of Meteorology). Australian Gridded Climate Data were required to calculate monthly values for the crossover temperature model because unlike Williams et al. (2010), it provides daily data.

To relate proportional herbaceous C_4 cover at each plot to climate and soil data, we used a generalised additive model (GAM) approach. GAMs were chosen because they can accommodate non-linear effects (Wood 2006; Wood 2017) and can be specified to account for high spatial autocorrelation (see discussion below; Zuur et al. 2009). Because C_4 plot cover data was proportional with 'true' values of 0 and 1, we used a logistic error structure (Douma and Weedon 2019). The smooth functions of each variable were limited to five degrees of freedom. This allowed for nonlinearity in the data while avoiding overfitting. Models were limited to variables that had Pearson pairwise correlations < 0.8 and interaction terms were not included. Models were compared using a step-wise, forward-selection procedure and Akaike information criterion (AIC). Model fit was measured using R^2. Models were constructed using the *gamm* function in the mgcv package (Wood 2021).

Moran's I tests confirmed the presence of spatial autocorrelation in preliminary GAM residuals (Matthews et al. 2019). Spatial autocorrelation can reduce model precision and predictive power (Guélat and Kéry 2018; Mets et al. 2017). Spatial autocorrelation can be alleviated by either (a) including spatial coordinates (i.e. longitude, latitude) in the model as covariates, or by (b) accounting for spatial autocorrelation in model residuals. The former can be problematic because spatial coordinates typically co-vary with environmental variables. Therefore, we incorporated a correlation structure in the model residuals.

Step3b. Extrapolate proportional herbaceous C_4 and C_3 cover

Model AIC comparisons indicated the best model to predict proportional herbaceous C_4 cover included the ratio (log) of summer (Dec-Jan-Feb) to winter (Jun-Jul-Aug) rainfall (slrain1), the maximum temperature of the hottest month (maxtx), and soil pH-CaCl2 (PHC), sand content (%; SND), and available water capacity (%; AWC) as variables (R^2=0.7; Supplemental Results Table 4). As maxtx, slrain1 and PHC increased (i.e. pH becomes more alkaline), proportional herbaceous C_4 cover generally increased (Fig 5 a,b,c). The effects of sand content and AWC were nonlinear (Fig 5 e,f), where proportional herbaceous C_4 cover was predicted to be higher in plots where both soil sand content and AWC exhibited more extreme values. However, these nonlinear trends may have been driven by the relative paucity of data in areas with low sand content (<40%) and AWC (<12%). The resulting GAM was extrapolated across the Australian continent (Fig 6) and used to predict proportional herbaceous C_4 cover in each grid-cell of the *% natural herbaceous cover* layer and generate a *% natural herbaceous C_4 cover* layer. A *% natural herbaceous C_3 cover* layer was calculated by subtracting the *% natural herbaceous C_4 cover* layer from the original *% natural herbaceous cover* layer.

Step 4: Create final C_3 and C_4 vegetation layers

To finalise the C_3 and C_4 cover vegetation layers, all C_3 vegetation layers were summed to create a single *% C_3 cover* layer (Fig 7a).

Eq 2 *% C_3 crop cover + % natural herbaceous C_3 cover + % woody vegetation =% C_3 cover*

Similarly, C_4 vegetation layers were summed to create a single *% C_4 cover* layer (Fig 7b).

Eq 3 *% C_4 crop cover + % natural herbaceous C_4 cover = % C_4 cover*

Finally, both *% C_3* and *C_4 cover* layers were converted from *% cover* to *% vegetation*. This is because many areas will have a high percentage of bare ground that is irrelevant to calculating the final isoscape. The *% vegetation* was calculated as:

.

Eq 4 *% vegetation = % cover* of vegetation type / *% total vegetation cover*.

This resulted in the final two layers, *% C_3 vegetation* and *% C_4 vegetation* (Fig 7c,d)

Step 5: Calculate site-averaged vegetation $\delta^{13}C$ using a two end-member mixing model

The average vegetation $\delta^{13}C$ value for each grid-cell was calculated based on the final *% C_3 vegetation* and *% C_4 vegetation* layers and a $\delta^{13}C$ mixing model. End-members were derived from the literature. Previous work has indicated that understory plants in closed canopy environments have lower $\delta^{13}C$ values than open forests (Cheesman et al. 2020; Powell et al. 2012); however, the bulk of leaf mass resides in the upper canopy. Moreover, this effect is typically most exagerated in dense rainforest habitats, which represent a minute porportion of the total land area in Australia. Therefore, we opted not to apply a canopy cover correction to average vegetation $\delta^{13}C$ values because (a) there was enough data to calculate a reliable correction value, and (b) such a correction was not deemed useful at this resolution. Previous work has also applied different end-member $\delta^{13}C$ values for herbacous and woody C_3 vegetation (Firmin 2016). However work by Pate et al. (1998) and data from Munroe et al.

(2020b) did not indentify significant differences in $\delta^{13}C$ between C_3 herbaceous and C_3 woody species. Thus, for simplicity, using values from Munroe et al. (2020b), we calculated the mean ± sd $\delta^{13}C$ values for C_4 and C_3 (herbacous and woody) endmembers. The mean ± sd of $\delta^{13}C$ values for C_4 herbaceous plants was -13.8±1.1‰ (n=119), and for C_3 herbaceous/woody plants was -27.7 ± 2.3‰ (n=420).

The site-averaged vegetation $\delta^{13}C$ isoscape was then calculated using a Monte Carlo method and a simple mixing model:

Eq 5 $$\delta^{13}C_{leaf} = f\,C_{4veg} * (\delta^{13}C_{C4veg}) + f\,C_{3veg} *(\delta^{13}C_{C3veg})$$

$f\,C_{4veg}$= *% C_4 vegetation*
$f\,C_{3veg}$ =*% C_3 vegetation*

Different possible values of $\delta^{13}C_{C4veg}$ and $\delta^{13}C_{C3veg}$ from the range of possible $\delta^{13}C$ values (mean ± 2 * sd) determined from Munroe et al. (2020b) were randomly substituted into Eq 5 for 1000 iterations. The results were averaged to produce the final vegetation $\delta^{13}C$ isoscape. A standard deviation raster was created by calculating the standard deviation of the 1000 iterations of each grid cell (Fig. 8).

Step 6. Validation

To validate model outcomes and the final vegetation $\delta^{13}C$ isoscape, we calculated the root mean squared error (RMSE) of competing % herbaceous C_4 cover ~ climate models (Bataille et al. 2018). The RMSE of each model was calculated using 10-fold cross-validation where the original dataset was randomly split ten times between a training data set (90% of plots) and a testing dataset (10% of plots). To assess the accuracy of the final *% C_4 vegetation* layer, we compared the predicted *% C_4 vegetation* layer outputs to the proportional % C_4 vegetation cover observed at all TERN plots. We used a linear regression to quantify relationships between predicted and observed *% C_4 vegetation* values. We also compared the residual values of predicted and observed *% C_4 vegetation* in different major vegetation groups (MVG), as determined by onsite evaluations by TERN survey teams.

Finally, we compared predicted leaf-$\delta^{13}C$ values to soil organic matter (SOM) $\delta^{13}C$ values determined samples collected at TERN plots. SOM $\delta^{13}C$ values were provided from two separate projects. Soil samples were collected at 19 TERN plots between 2011 and 2013 and analysed in 2019 as part of a project testing different isotopic tools to predict % C_4 abundance (Atkins 2020). These plots are located along a North to South transect through central Australia (Supplemental Methods Figure 4). For this project, a single soil sample was collected from the top 3 cm of the soil profile at the same location in each plot. Additional SOM $\delta^{13}C$ values were provided from 32 TERN plots located along the Adelaide Geosyncline in South Australia as part of a project examining the relationship between soil isotopic composition and aridity (Farrell, unpublished data). In April and May 2016, 20 soil samples were taken at random within each plot from the 0-10 cm layer; the 20 samples were composited and homogenised in the field to yield a single representative 0-10 cm soil sample for each plot. Atkins (2020) 0-3 cm depth SOM $\delta^{13}C$ values were adjusted by 0.5‰ and Farrell 0-10 cm depth SOM $\delta^{13}C$ values by 1‰ to account for ^{13}C enrichment during decomposition in SOM (Krull and Bray 2005). Like *% C_4 vegetation* comparisons, we calculated the residuals for SOM-adjusted and predicted leaf $\delta^{13}C$ values and used a linear regression to compare predicted and measured results.

Applications

To demonstrate the analytical potential for landscape research with these vegetation data layers, we used the % *C_4* and *C_3 vegetation cover* layers and leaf-δ^{13}C isoscape to calculate the mean C_4 and C_3 cover and leaf-δ^{13}C values of 86 different continental Australian bioregions, as described by the interim Biogeographic Regionalisation for Australia version 7 (IBRA 7.0; Department of Agriculture, Water and the Environment, 2020). Bioregions are large, geographically distinct areas that share common characteristics such as climate, landform patterns, and plant and animal communities. These regions are used to help identify unique ecosystems within Australia. Thus, understanding differences in C_3 and C_4 cover between these regions is critical to identifying their unique attributes and vulnerabilities. Here we compared mean proportional C_3 and C_4 cover and leaf-δ^{13}C in each bioregion to trends slrain1 and % *woody* and *herbaceous cover*.

Results

Geographic distribution of vegetation δ^{13}C in Australia

Our stepwise procedures produced 9 data layers representing C_4 and C_3 distribution in both agricultural and native environments. Predicted *% C_3* and *C_4 vegetation* maps and the δ^{13}C leaf isoscape followed expected trends in C_3 and C_4 vegetation (Fig 7 and 8). Southern areas of the country, which are characterised by cooler temperatures and high winter rainfall, were dominated by large areas of C_3 cropland and woody vegetation, and thus had the most negative δ^{13}C values. Mid-western and eastern coastal regions also have a large proportion of C_3 vegetation, including a mix of forests, cropland, and herbaceous vegetation, and have correspondingly low δ^{13}C values. C_4-dominated and isotopically ^{13}C-enriched areas predominately included northern savannahs and grasslands.

The south to north transition from C_3 to C_4 dominated areas, and more negative to less negative $\delta^{13}C$ values, was abrupt. The clear demarcation between C_3 and C_4 habitats is consistent with the relatively rapid transition from winter to summer dominated rainfall patterns across the country. Central areas of Australia are arid and receive sporadic rainfall with high inter-annual variability. As a result, there is relatively low and sparse woody cover and conditions do not support most C_3 herbaceous plants. The apparent exception to this is the Simpson Desert, located in central Australia across South Australia and the Northern Territory. Although C_3 cover in the Simpson Desert was low and consistent with surrounding areas, this region has notably lower C_4 herbaceous cover compared to other nearby environments, leading to lower proportional C_4 vegetation cover and $\delta^{13}C$ values. This due to the extremely dry conditions (< 50 mm rainfall/year) in the desert which make it difficult for any herbaceous plants to grow.

Validation

As previously described, the best model to predict proportional herbaceous C_4 cover included slrain1, maxtx, PHC, SND, and AWC as variables. The proportional herbaceous C_4~climate GAM used to predict C_4 cover had a mean RMSE of 27.8% ± 2.0. Linear regression analysis comparing predicted and observed proportional herbaceous C_4 vegetation cover resulted in an adjusted-R^2 of 0.54 (Fig 9a). Comparisons between predicted and observed % C_4 vegetation (including woody cover) at TERN plots returned residuals ranging from -63.4 to 73.4% (mean ± sd = 9.1 ± 24.5) and a RMSE of 26.1%. This suggests that, on average, our approach overestimates relative C_4 cover. Linear regression analysis comparing predicted and observed proportional C_4 vegetation cover resulted in an adjusted-R^2 of 0.44 (Fig 9b).

Most TERN plots were located in Eucalypt woodlands, followed by Tussock grasslands, Chenopod shrublands, and Acacia woodlands. Comparisons of residuals between major vegetation group classifications revealed that residuals were smallest in heathlands, Eucalypt woodlands and forests, and tussock grasses, but were largest in Acacia- and Melaleuca-dominated habitats (Supplemental Results Table 2; Fig 10). The spread in the residuals for each MVG indicated that C_4 cover was generally overestimated in most habitats but was underestimated in grasslands.

Comparisons between predicted leaf and soil $\delta^{13}C$ isotope values returned a RMSE of 2.1‰. Residuals ranged from -5.40‰ to 5.44‰ with a mean value of 0.26‰ (±2.12). The line of best fit between these variables had a slope of 0.74, an intercept of -6.0, and an adjusted-R^2 of 0.71 (Fig 9c). These results indicate that on average the isoscape overestimated mean leaf $\delta^{13}C$ values (i.e. were less negative), which is consistent with comparisons between predicted and observed % C_4 vegetation.

IBRA Analysis

Bioregions with the greatest proportional C_3 cover were located Tasmania, southern Australia, and the Australian Alps (100% C_3 cover; Supplemental Results, IBRA Analysis). Bioregions with the greatest C_4 cover included the Central Kimberly, Mitchell Grass Downs, and Gulf Plains (> 75% C_4 cover). Across all bioregions, we found an increasing trend of proportional C_3 cover with increased % woody cover (Fig. 11a), but no relationship between increased herbaceous cover and proportional C_4 cover (Fig. 11b). There was also a clear non-linear relationship between slrain1 and mean proportional C_3 cover; where slrain1 increased, there was a rapid decline in % C_3 cover (Fig. 11c). This is mirrored by an increase in mean predicted leaf-$\delta^{13}C$ with increased slrainl.

Discussion

We leveraged a novel combination of field surveys and remote sensing data to create national C_3 and C_4 vegetation maps and a $\delta^{13}C$ vegetation isoscape for Australia. The good agreement between our predictions and observed values indicates our approach can provide valuable generalized depictions of C_4 and $\delta^{13}C$-leaf variation across diverse landscapes at large scales. Our approach benefits from recent advancements in remote sensing by being the first to incorporate vegetation layering, which is critical to accurate representations of C_3:C_4 trends. Our work also demonstrates the value of extensive field surveys when constructing and validating isoscape projections in different regions, by providing the unique ability to incorporate edaphic variables into large-scale models. This is particularly impressive considering the ground survey vegetation data used to construct the final outputs were collected by TERN over a period of 9 years, both before and after the 2015 remote sensing time-slice used to create the isoscape. Most of these plots have only been surveyed once and thus describe a snap-shot in time from a single season. Therefore, an average error rate of ~25% represents a significant level of overall accuracy. Comparisons between predicted leaf-$\delta^{13}C$ values to measured $\delta^{13}C$ soil values achieved a stronger correlation than comparisons to ground surveys. The stronger correlation may be because soil $\delta^{13}C$ represents long-term averages in relative C_4 vegetation cover. Our $\delta^{13}C$ validation results are consistent with the level of accuracy achieved by other $\delta^{13}C$ isoscapes developed using remote sensing techniques in North and South America (Powell et al. 2012, Firmin 2016). Overall, the relatively high level of accuracy in our C_4 and $\delta^{13}C$ predictions demonstrates remote sensing combined with field surveys can provide useful, generalized C_4 maps and $\delta^{13}C$ isoscapes, and informative estimations on C_3:C_4 vegetation cover over diverse landscapes in areas where data is limited.

Modelling herbaceous C_4 and C_3 distribution

The best model for predicting proportional C_4 herbaceous cover included maximum summer temperature and seasonal rainfall ratios as climate variables. This is consistent with previous work indicating both C_4 grass and sedge cover is predominantly correlated with January temperatures and proportional summer rainfall (Murphy and Bowman 2007; von Fischer et al. 2008). Interestingly, the crossover temperature model was one of the least accurate climate models and was difficult to apply consistently across Australia. These findings are consistent with Munroe et al (2022) and Xie et al (2022), who also found that seasonal rainfall ratios and summer temperatures were better predictors of C_4 grass abundance than the crossover temperature model. Although we acknowledge that the crossover approach was never intended to delineate fine-scale distribution patterns, our results demonstrate this approach is not the best method to determine C_4 distribution in Australia.

Local edaphic factors were also selected in the best fit model. Previous work has demonstrated local environmental factors can significantly modify herbaceous C_4 distribution (Griffith et al. 2015; Nippert and Knapp 2007; Wang et al. 2020). Our work suggests pH has a significant positive influence on relative C_4 herbaceous cover and should be considered even in continental models. The influence of alkaline-stress on C_4 versus C_3 plants is not well understood, but C_4 plants are thought to be more resistant to stress and therefore more tolerant to alkaline soil (Bromham et al. 2013; Sage 2004; Saslis-Lagoudakis et al. 2014). However, pH is often related or correlated with other climate and soil conditions like salinity and rainfall, thus the observed effect of pH may reflect underlying factors not included in our analysis (James et al. 2005; Saslis-Lagoudakis et al. 2014). Isolating the impacts of available water capacity and sand content is more difficult given its apparent nonlinear relationship to C_4 cover, but together they may indicate a significant impact of changes in local moisture availability, which can affect competitive dynamics between C_3 and C_4 species (Nippert and Knapp 2007; Sage 2004).

Limitations and Uncertainty

The proportional herbaceous C_4 cover model tended to underestimate C_4 cover in areas with high observed values, and overestimate cover in areas with low or zero measured herbaceous C_4 cover. There are several possible explanations for this pattern. Analysis revealed most TERN plots were dominated by either C_3 or C_4 herbaceous cover. Because mixed C_3-C_4 herbaceous environments were less common, they were invariably harder to predict. Lastly, most climate data were centred on the year 1990, which may be less applicable for more recent plots, leading to higher overall error rates. Most importantly, although we considered a range of local factors in our C_4 cover models, models did not include other factors which may also modify C_4 patterns but cannot currently be extrapolated at large scales, such as local disturbance, soil salinity, and competition between native and alien species (Griffith et al. 2015; Sage et al. 1999).

A critical source of potential error in our final vegetation maps was the *% woody vegetation* layer, generated using the Seasonal Persistent Green Cover product (Gill et al. 2017; Gill et al. 2015). While the overall accuracy of the Seasonal Persistent Green Cover product is impressive, Gill et al. (2017) noted that accuracy varied significantly between habitat types. This was evident when comparing C_4 cover model accuracy between different major vegetation groups. We found our C_4 estimates were least accurate in *Acacia*-dominated habitats. These higher error rates are consistent with Gill et al. (2017), who found most areas identified as *Acacia* forests, woodlands, and open woodlands were not mapped as forest. Instead, they were incorrectly classed as having very low or no woody cover. Gill et al.

(2017) suggested several explanations for this issue; vegetation cover in *Acacia*-dominated habitats can be sparse, which can make woody cover more difficult to detect. At thresholds of <10% woody cover it was difficult to distinguish woody and non-woody vegetation (Gill et al. 2017). Therefore, it can be more difficult to accurately assess woody cover, and proportional C_3 vegetation cover, in sparse areas. Some *Acacia* also have narrow, needle-like leaves which are harder to detect via satellite, whereas other *Acacia* species are known to drop their leaves in very dry conditions, resulting in a low minimum green cover-fraction over the course of the year. Finally, the understory is often visible through the sparse *Acacia* canopy. When the understory greens-up in response to rainfall, this can give the appearance of a highly variable time series in green cover for *Acacia* foliage, leading to its misclassification as non-woody. Unsurprisingly, the difficulties associated with measuring *Acacia* woody cover in Australia using remote sensing led to a high degree of variation C_3 and C_4 cover estimates in *Acacia*-dominated habitats.

C_3:C_4 estimates were also less accurate in chenopod shrublands. Accurately estimating C_4 cover in these environments may be more difficult because chenopod shrublands are often sparsely vegetated (Gill et al. 2017). Our approach also assumed all shrub cover had a C_3 pathway. But as previously discussed, C_4 chenopods can be locally common in Australian shrublands. As a result, our approach may underestimate C_4 cover in these habitats. However, our model residuals indicate C_4 cover is more likely to be overestimated in chenopod shrublands, which suggests our assumption that all shrubs are C_3 is not the main source of error in these habitats. More likely, it is the difficulty associated with accurately assessing woody cover in these sparse environments.

Other potential sources of error include the high degree of variation in $\delta^{13}C$ values between different C_3 species and environmental conditions (Kohn 2010). For example, rainfall, soil pH, and leaf nitrogen area are all significant drivers of variation in global C_3 $\delta^{13}C$ values (Cornwell et al. 2018). Variation in $\delta^{13}C$ values within the canopy will also affect the overall accuracy of $\delta^{13}C$ isoscapes (Cheesman et al. 2020), however it is difficult to effectively quantify and model these different sources of variation across Australia at this time. Unsurprisingly, areas with the greatest standard deviation in $\delta^{13}C$ values were areas dominated by C_3 vegetation reflecting the greater variability in the carbon isotopic composition among C_3 plants.

Future Improvements

The accuracy of the $\delta^{13}C$ isoscape hinges on three main components; (1) estimates of woody and herbaceous cover, (2) the C_3:C_4 herbaceous cover model, and (3) the endmember values in the $\delta^{13}C$-leaf mixing model. Gill et al. (2017) outlines multiple ways to improve estimates of woody cover. The proportional herbaceous C_4 cover ~ climate model could be improved as TERN increases its plot network and environmental representation. For example, establishing plots in Tasmania or increasing the number of plots with more equal C_3:C_4 ratios would improve model outcomes by increasing the amount of data from cool climates and transitional habitats. TERN has also begun to regularly revisit existing plots to monitor change over time. Revisits could be used to calculate average C_4 cover over multiple years and seasons, which would make the plot data a more appropriate validation tool for average C_4 vegetation and isoscape projections. This would also enable the creation of more seasonally specific isoscapes, rather than a static annual average. More specific information on crop commodities, namely the location of maize and sorghum, would also improve the accuracy of C_3 and C_4 vegetation layers.

The $\delta^{13}C$ endmembers were based on $\delta^{13}C$ values from Munroe et al. (2020b). These values were measured from species collected at TERN plots, making them a useful metric with which to calculate Australian vegetation $\delta^{13}C$ endmembers. However, the plants measured by Munroe et al. (2020b) were not necessarily dominant or wide spread. Measuring the $\delta^{13}C$ of the most common plants in TERN plots, and incorporating a wider range of herbaceous and woody species, may help create endmembers that are better representations of dominant Australian plant $\delta^{13}C$ values. Testing specimens that were collected under different conditions (e.g. rainfall or soil pH) would enable expansion of the current mixing model to account for different climate conditions when predicting $\delta^{13}C$ values, particularly in C_3 species (Cornwell et al. 2018).

Applications

The terrestrial carbon isoscape and C_3 and C_4 maps presented here have numerous valuable applications. As demonstrated in this study, C_3, C_4 and $\delta^{13}C$ maps can be used to quantify and compare C_3 and C_4 distribution across different bioregions at a landscape scale. Such analysis would not be possible without these data. Isoscapes are also enormously useful in the study of food web dynamics and animal migration (Hobson et al. 2010; Vander Zanden et al. 2018; Wunder 2010). These maps could also be used to calculate fractional productivity of different photosynthetic pathways (Powell et al. 2012).

TERN's expansive plot network provides the opportunity to not only identify, but also quantify discrepancies between predicted and observed C_4 and C_3 cover. Indeed, our work has already demonstrated the importance of some edaphic factors in controlling C_4 distribution. As more data becomes available, further comparisons across a wider range of factors will be possible. Similarly, differences in predicted $\delta^{13}C$ values and local vegetation can be used to examine the influence of local factors, such as water stress or drought, on $\delta^{13}C$ values (Ehleringer 1993; Mårtensson et al. 2017; Tieszen 1991).

Climate change is anticipated to drastically shift the competitive advantage of C_3 and C_4 plants in Australia and globally, leading to substantial changes in species distribution (Corlett and Westcott 2013; Hasegawa et al. 2018). This will likely drive significant bottom-up changes to the structure and diversity of faunal communities (Haddad et al. 2009; Haveles et al. 2019; Warne et al. 2010). Using our underlying climate models, C_3 and C_4 abundance can be extrapolated under future conditions and areas that are most vulnerable to extreme changes in C_3 and C_4 cover can be identified. Our models identified maximum temperature and seasonal water availability as the two most significant climate factors driving C_3 and C_4 herbaceous cover in Australia. Based on these findings, we would expect to see considerable expansion of C_4 suitable climate-zones in southern Australia. Historically, southern Australia has a Mediterranean climate, with dry summers and higher winter rainfall. However, the climate in southern Australia is expected to become increasingly dry, with hotter temperatures and more frequent heatwaves (Keywood et al. 2017; Suppiah et al. 2006), conditions that are better suited to C_4 species. These models will also improve our ability to quantify potentially improved conditions for invasive species, such as the invasive C_4 buffel grass, *Cenchrus ciliaris* L. (de Albuquerque et al. 2019; Lawson et al. 2004). Forecasting native C_3 and C_4 abundance can also enable proactive environmental management in Australia's changing climate, such as identifying suitable locations for future C_4 and C_3 crops (Cullen et al. 2009) or important refuge areas for native plant communities (Graham et al. 2019; Selwood and Zimmer 2020).

Conclusion

We have applied a novel combination of detailed ground survey, climate, and remote sensing data to create and evaluate the first Australian vegetation $\delta^{13}C$ isoscape. These results have a wide range of applications, including the study of animal migration, food web patterns, spatial and temporal variation in plant productivity and habitat structure, carbon exchange, and the impact of water stress on plant communities. Our continued ability to test and validate these models as new TERN plots and isotope data become available provides a unique opportunity to develop future improvements. The C_3, C_4 and isoscape maps presented here were created to support the study of Australian ecosystems and have enormous value to broader ecological research. It is our intention to curate and update these outputs where possible as new TERN plots and isotope data become available.

References

ABARES (2016) The Australian Land Use and Management Classification Version 8. Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra

ABS (2016) Agricultural Commodities, Australia, 2015-16. In: ABS (ed). ABS, Canberra

Akhani H, Trimborn P, Ziegler H (1997) Photosynthetic pathways in Chenopodiaceae from Africa, Asia and Europe with their ecological, phytogeographical and taxonomical importance. Plant Syst Evol 206:187-221

Atkins R (2021) Soil carbon isotopic proxies for determining the photosynthetic pathway of floral communities: A method inter-comparison. Doctoral dissertation, University of Adelaide

Bataille CP, von Holstein ICC, Laffoon JE, Willmes M, Liu X-M, Davies GR (2018) A bioavailable strontium isoscape for Western Europe: A machine learning approach. PLOS ONE 13:e0197386

Ben-David M, Flaherty EA (2012) Stable isotopes in mammalian research: a beginner's guide. J Mammal 93:312-328

Bromham L, Saslis-Lagoudakis CH, Bennett TH, Flowers TJ (2013) Soil alkalinity and salt tolerance: adapting to multiple stresses. Biol Lett 9:20130642

Cernusak LA (2020) Gas exchange and water-use efficiency in plant canopies. Plant Biol 22:52-67

Cheesman AW, Duff H, Hill K, Cernusak LA, McInerney FA (2020) Isotopic and morphologic proxies for reconstructing light environment and leaf function of fossil leaves: a modern calibration in the Daintree Rainforest, Australia. Am J Bot 107:1165-1176

Cole PG, Weltzin JF (2005) Light limitation creates patchy distribution of an invasive grass in eastern deciduous forests. Biol Invasions 7:477-488

Collatz GJ, Berry JA, Clark JS (1998) Effects of climate and atmospheric CO_2 partial pressure on the global distribution of C_4 grasses: present, past, and future. Oecologia 114:441-454

Corlett RT, Westcott DA (2013) Will plant movements keep up with climate change? Trends Ecol Evol 28:482-488

Cornwell WK, Wright IJ, Turner J et al (2018) Climate and soils together regulate photosynthetic carbon isotope discrimination within C_3 plants worldwide. Glob Ecol Biogeogr 27:1056-1067

Cullen BR, Johnson IR, Eckard RJ et al (2009) Climate change effects on pasture systems in south-eastern Australia. Crop Pasture Sci 60:933-942

de Albuquerque FS, Macías-Rodríguez MÁ, Búrquez A, Astudillo-Scalia Y (2019) Climate change and the potential expansion of buffelgrass (*Cenchrus ciliaris* L., Poaceae) in biotic communities of Southwest United States and northern Mexico. Biol Invasions 21:3335-3347

Department of Agriculture, Water and the Environment (2020), Interim Biogeographic Regionalisation for Australia (Regions - States and Territories) v. 7 (IBRA)

Dormann CF, Bagnara M, Boch S et al (2020) Plant species richness increases with light availability, but not variability, in temperate forests understorey. BMC Ecology 20:43

Douma JC, Weedon JT (2019) Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression. Method Ecol Evo 10:1412-1430

Ehleringer JR (1978) Implications of quantum yield differences on the distributions of C_3 and C_4 grasses. Oecologia 31:255-267

Ehleringer JR (1993) 11 - Carbon and Water Relations in Desert Plants: An Isotopic Perspective. In: Ehleringer J. R., Hall A. E., Farquhar G. D. (eds), Stable Isotopes and Plant Carbon-water Relations. Academic Press, San Diego, pp. 155-172

Firmin SM (2016) The spatial distribution of terrestrial stable carbon isotopes in North America, and the impacts of spatial and temporal resolution on static ecological models. University of Denver

Flockhart DTT, Brower LP, Ramirez MI et al (2017) Regional climate on the breeding grounds predicts variation in the natal origin of monarch butterflies overwintering in Mexico over 38 years. Glob Change Biol 23:2565-2576

Frank DC, Poulter B, Saurer M et al (2015) Water-use efficiency and transpiration across European forests during the Anthropocene. Nat Clim Change 5:579-583

Gallant J, Austin J, Williams K et al (2018) 9s soil and landform for continental Australia analysis of biodiversity pattern: aggregated from 3s data. v1. CSIRO,

Gill T, Johansen K, Phinn S, Trevithick R, Scarth P, Armston J (2017) A method for mapping Australian woody vegetation cover by linking continental-scale field data and long-term Landsat time series. Int J Remote Sens 38:679-705

Gill T, Johansen K, Scarth P, Armston J, Trevithick R, Flood N (2015) Persistent Green Vegetation Fraction. In: Held A., Phinn S., Soto-Berelov M., Jones S. (eds), AusCover Good Practice Guidelines: A technical handbook supporting calibration and validation activities of remotely sensed data products. . TERN AusCover, pp. 139-160

Graham V, Baumgartner JB, Beaumont LJ, Esperón-Rodríguez M, Grech A (2019) Prioritizing the protection of climate refugia: designing a climate-ready protected area network. J Environ Plan Manag 62:2588-2606

Griffith DM, Anderson TM, Osborne CP, Strömberg CA, Forrestel EJ, Still CJ (2015) Biogeographically distinct controls on C_3 and C_4 grass distributions: merging community and physiological ecology. Glob Ecol Biogeogr 24:304-313

Griffith DM, Rebecca L Powell, Firmin S, Cotton J, Still CJ (2019) grassmapr, an R package to predict C_3/C_4 grass distributions and model terrestrial $\delta^{13}C$ isoscapes. R Package version 1 edn.,

Guélat J, Kéry M (2018) Effects of spatial autocorrelation and imperfect detection on species distribution models. Method Ecol Evo 9:1614-1625

Guerin GR, Saleeba T, Munroe S, Blanco-Martin B, Martín-Forés I, Tokmakoff A (2020) ausplotsR: TERN AusPlots analysis package. R Package version 1.2 edn.,

Haddad NM, Crutsinger GM, Gross K, Haarstad J, Knops JMH, Tilman D (2009) Plant species loss decreases arthropod diversity and shifts trophic structure. Ecol Lett 12:1029-1039

Hasegawa S, Piñeiro J, Ochoa-Hueso R et al (2018) Elevated CO_2 concentrations reduce C_4 cover and decrease diversity of understorey plant community in a Eucalyptus woodland. J Ecol 106:1483-1494

Hattersley P (1983) The distribution of C_3 and C_4 grasses in Australia in relation to climate. Oecologia 57:113-128

Haveles AW, Fox DL, Fox-Dobbs K (2019) Carbon isoscapes of rodent diets in the Great Plains USA deviate from regional gradients in C_4 grass abundance due to a preference for C_3 plant resources. Palaeogeogr Palaeoclimatol Palaeoecol 527:53-66

Hobson KA, Barnett-Johnson R, Cerling T (2010) Using isoscapes to track animal migration. In: West J., Bowen G., Dawson T., Tu K. (eds), Isoscapes. Springer, Dordrecht, pp. 273-298

Hobson KA, Kardynal KJ (2015) An isotope ($\delta^{34}S$) filter and geolocator results constrain a dual feather isoscape (δ^2H, $\delta^{13}C$) to identify the wintering grounds of North American Barn Swallows. The Auk 133:86-98

Hobson KA, Wassenaar LI (2018) Tracking animal migration with stable isotopes, 2nd Edition. Academic Press, London

James JJ, Tiller RL, Richards JH (2005) Multiple resources limit plant growth and function in a saline-alkaline desert community. J Ecol 93:113-126

Kellogg EA (2001) Evolutionary history of the grasses. Plant Physiol 125:1198-1205

Kelly JF (2000) Stable isotopes of carbon and nitrogen in the study of avian and mammalian trophic ecology. Can J Zool 78:1-27

Keywood M, Hibberd M, Emmerson K (2017) Australia state of the environment 2016: atmosphere, independent report to the Australian Government Minister for the Environment and Energy, Australian Government Department of the Environment and Energy, Canberra. Canberra,

Kohn MJ (2010) Carbon isotope compositions of terrestrial C_3 plants as indicators of (paleo)ecology and (paleo)climate. PNAS 107:19691-19695

Krull EG, Bray, SS (2005) Assessment of vegetation change and landscape variability by using stable carbon isotopes of soil organic matter. Aus J Bot, 53:651-661 doi:10.1071/bt04124

Lawson B, Bryant M, Franks A (2004) Assessing the potential distribution of buffel grass (*Cenchrus ciliaris* L.) in Australia using a climate-soil model. Plant Prot Q 19:155-163

López-Calderón C, Hobson KA, Marzal A et al (2017) Wintering areas predict age-related breeding phenology in a migratory passerine bird. J Avian Biol 48:631-639

Mårtensson L-M, Carlsson G, Prade T, Kørup K, Lærke PE, Jensen ES (2017) Water use efficiency and shoot biomass production under water limitation is negatively correlated to the discrimination against 13C in the C3 grasses Dactylis glomerata, Festuca arundinacea and Phalaris arundinacea. 113:1-5

Matthews JL, Diawara N, Waller LA (2019) Quantifying spatio-temporal characteristics via Moran’s statistics. In: Diawara N. (ed), Modern statistical methods for spatial and multivariate data. Springer International Publishing, Cham, pp. 163-177

Mets KD, Armenteras D, Dávalos LM (2017) Spatial autocorrelation reduces model precision and predictive power in deforestation analyses. Ecosphere 8:e01824

Munroe S, Guerin G, Saleeba T et al (2020a) ausplotsR: An R package for rapid extraction and analysis of vegetation and soil data collected by Australia’s Terrestrial Ecosystem Research Network. EcoEvoRxiv.

Munroe S, McInerney F, Andrae J et al (2020b) The photosynthetic pathways of plant species surveyed in Australia’s national terrestrial monitoring network.

Munroe SEM, McInerney FA, Guerin GR et al (2022) Plant families exhibit unique trends in C_4 richness and abundance. EcoEvoRxiv doi:10.32942/osf.io/vxu6s.

Murphy BP, Bowman DM (2007) Seasonal water availability predicts the relative abundance of C_3 and C_4 grasses in Australia. Glob Ecol Biogeogr 16:160-169

Nippert JB, Knapp AK (2007) Soil water partitioning contributes to species coexistence in tallgrass prairie. Oikos 116:1017-1029

O'Leary MH (1988) Carbon isotopes in photosynthesis. Bioscience 38:328-336

O'Neill AL (1996) Satellite-derived vegetation indices applied to semi-arid shrublands in Australia. Aust Geogr 27:185-199

Pate JS, Unkovich MJ, Erskine PD, Stewart GR (1998) Australian mulga ecosystems –^{13}C and ^{15}N natural abundances of biota components and their ecophysiological significance. Plant Cell Environ 21:1231-1242

Pau S, Edwards EJ, Still CJ (2013) Improving our understanding of environmental controls on the distribution of C_3 and C_4 grasses. Glob Change Biol 19:184-196

Powell RL, Yoo E-H, Still CJ (2012) Vegetation and soil carbon-13 isoscapes for South America: integrating remote sensing and ecosystem isotope measurements. Ecosphere 3:1-25

Prober SM, Thiele KR, Lunt ID (2007) Fire frequency regulates tussock grass composition, structure and resilience in endangered temperate woodlands. Austral Ecol 32:808-824

Sage RF (2004) The evolution of C_4 photosynthesis. New Phytol 161:341-370

Sage RF (2016) A portrait of the C_4 photosynthetic family on the 50th anniversary of its discovery: species number, evolutionary lineages, and hall of fame. J Exp Bot 68:11-28

Sage RF, Sage TL, Kocacinar F (2012) Photorespiration and the evolution of C_4 photosynthesis. Ann Rev Plant Biol 63:19-47

Sage RF, Wedin DA, Li M (1999) The biogeography of C_4 photosynthesis: patterns and controlling factors. In: Sage R. F. andMonson R. K. (eds), C_4 plant biology. Academic Press, Sydney, pp. 313-373

Saslis-Lagoudakis CH, Hua X, Bui E, Moray C, Bromham L (2014) Predicting species' tolerance to salinity and alkalinity using distribution data and geochemical modelling: a case study using Australian grasses. Ann Bot 115:343-351

Selwood KE, Zimmer HC (2020) Refuges for biodiversity conservation: a review of the evidence. Biol Conserv 245:108502

Sparrow AD, Friedel MH, Smith DMS (1997) A landscape-scale model of shrub and herbage dynamics in Central Australia, validated by satellite data. Ecol Model 97:197-216

Sparrow BD, Foulkes JN, Wardle GM et al (2020) A vegetation and soil survey method for surveillance monitoring of rangeland environments. Front Ecol Evol 8

Still C, Rastogi B (2017) What drives carbon isotope fractionation by the terrestrial biosphere? J Geophys Res Biogeosci 122:3108-3110

Still CJ, Berry JA, Collatz GJ, DeFries RS (2003) Global distribution of C_3 and C_4 vegetation: carbon cycle implications. Global Biogeochem Cy 17:1006

Still CJ, Powell RL (2010) Continental-scale distributions of vegetation stable carbon isotope ratios. Isoscapes. Springer, pp. 179-193

Suppiah R, Preston B, Whetton P et al (2006) Climate change under enhanced greenhouse conditions in South Australia.

Tieszen LL (1991) Natural variations in the carbon isotope values of plants: Implications for archaeology, ecology, and paleoecology. 18:227-248

Tieszen LL, Boutton TW, Tesdahl KG, Slade NA (1983) Fractionation and turnover of stable carbon isotopes in animal tissues: Implications for $\delta^{13}C$ analysis of diet. Oecologia 57:32-37

Trevithick R, Scarth P, Tindall D, Denham R, Flood N (2014) Cover under trees: RP64G Synthesis Report. Department of Science. Information Technology, Innovation and the Arts, Brisbane,

Vander Zanden HB, Nelson DM, Wunder MB, Conkling TJ, Katzner T (2018) Application of isoscapes to determine geographic origin of terrestrial wildlife for conservation and management. Biol Conserv 228:268-280 von Fischer JC, Tieszen LL, Schimel DS (2008) Climate controls on C_3 vs. C_4 productivity in North American grasslands from carbon isotope composition of soil organic matter. Glob Change Biol 14:1141-1155

Wang K, Zhong S, Sun W (2020) Clipping defoliation and nitrogen addition shift competition between a C_3 grass (*Leymus chinensis*) and a C_4 grass (*Hemarthria altissima*). Plant Biol 22:221-232

Warne RW, Pershall AD, Wolf BO (2010) Linking precipitation and C_3–C_4 plant production to resource dynamics in higher-trophic-level consumers. Ecology 91:1628-1638

West JB, Bowen GJ, Dawson TE, Tu KP (2009) Isoscapes: understanding movement, pattern, and process on Earth through isotope mapping. Springer, Dordrecht

White A, Sparrow B, Leitch E et al (2012) AUSPLOTS rangelands survey protocols manual. The University of Adelaide Press, Adelaide

Williams K, Stein J, Storey R et al (2010) 0.01 degree stack of climate layers for continental analysis of biodiversity pattern. version 1.0 edn. CSIRO

Winslow JC, Hunt Jr ER, Piper SC (2003) The influence of seasonal water availability on global C_3 versus C_4 grassland biomass and its implications for climate change research. Ecol Model 163:153-173

Wood S (2006) GAMMs with R. Chapman & Hall/CRC, New York, NY

Wood S (2021) Mixed GAM computation vehicle with automatic smoothness estimation. R Package version 18-35 edn.

Wood SN (2017) Generalized additive models: an introduction with R. CRC press, Boca Raton

Wunder MB (2010) Using isoscapes to model probability surfaces for determining geographic origins. In: West J. B., Bowen G. J., Dawson T. E., Tu K. P. (eds), Isoscapes. Springer, Dordrecht, pp. 251-270

Xie Q, Huete A, Hall CC, Medlyn BE, Power SA, Davies JM, et al. Satellite-observed shifts in C_3/C_4 abundance in Australian grasslands are associated with rainfall patterns. Remote Sens Environ. 273:112983. doi: https://doi.org/10.1016/j.rse.2022.112983.

Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) GLMM and GAMM. Mixed effects models and extensions in ecology with R. Springer New York, New York, NY, pp. 323-341

Figures

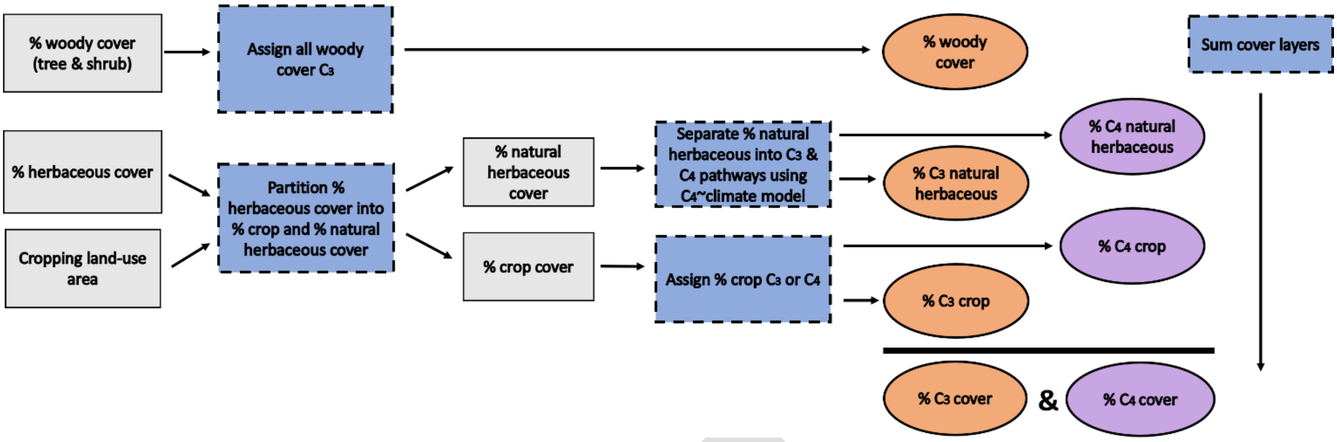

Fig 1. Conceptual diagram of the procedures used to create each C_3 and C_4 vegetation cover layer. Grey boxes specify generic vegetation layers, blue boxes specify steps in the methodology, orange ovals are the resulting C_3 vegetation cover layers, purple ovals are C_4 vegetation cover layers. All C_3 and C_4 layers were summed to create a total '% C_4 cover' and '% C_3 cover' layer

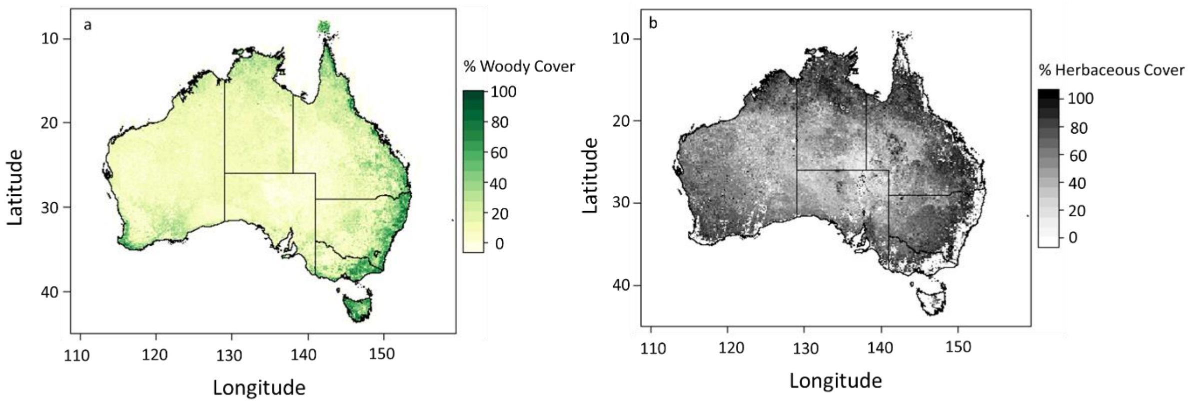

Fig. 2. Mean Australian (a) % woody cover (tree and shrub) and (b) % herbaceous cover in 2015

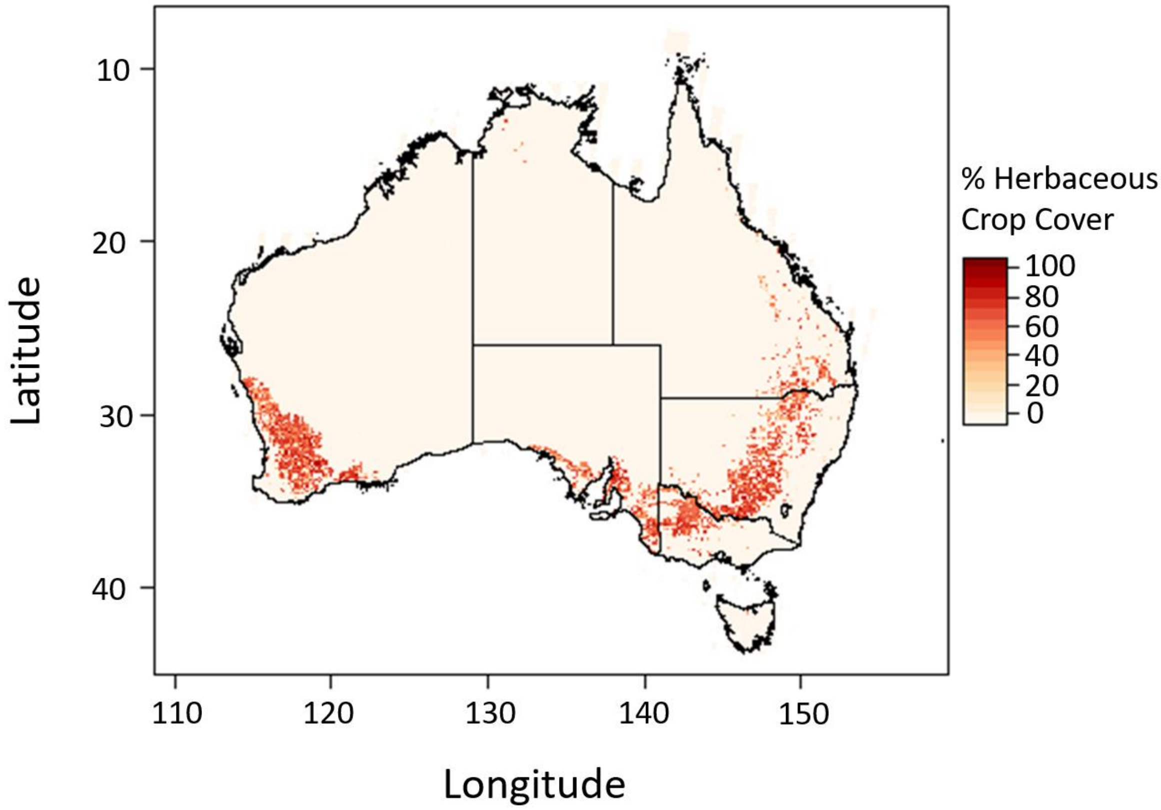

Fig. 3. Australian % herbaceous crop cover as of December 2018

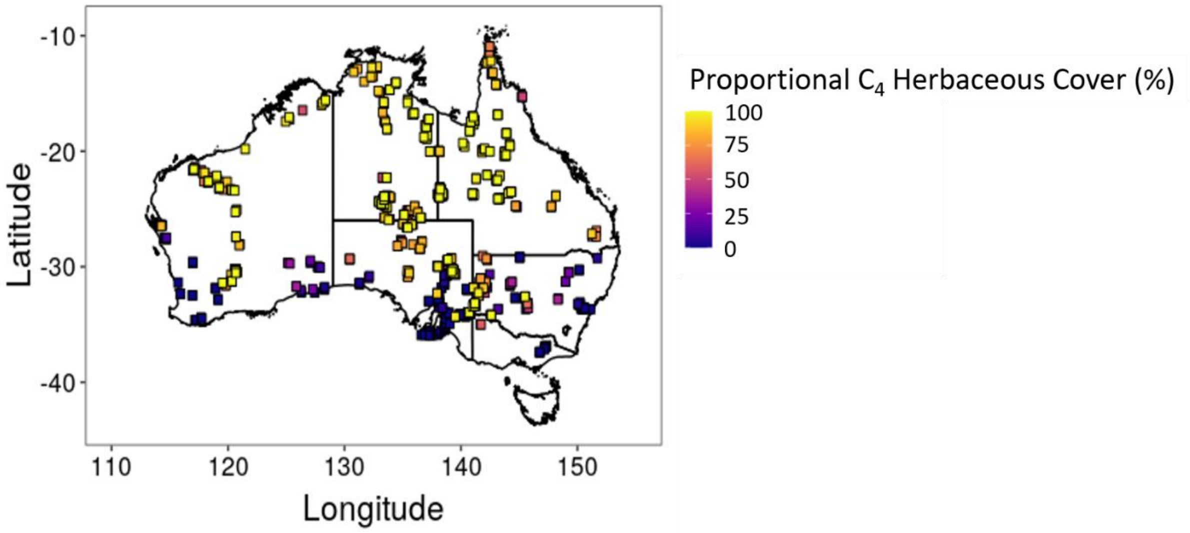

Fig. 4. Proportional (%) herbaceous C_4 cover (relative to herbaceous C_3 and C_4 cover) at TERN plots

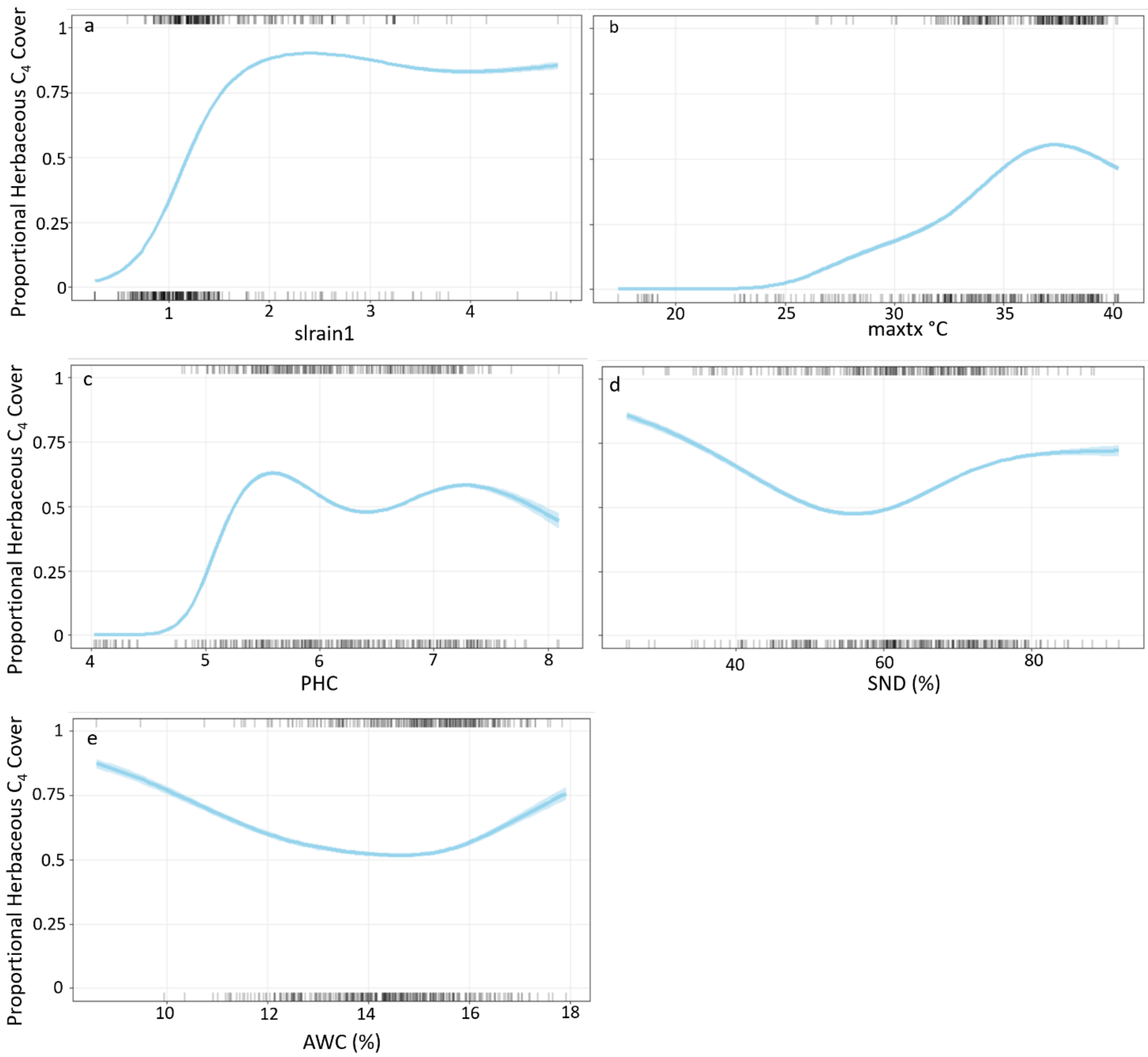

Fig. 5. Predicted proportional herbaceous C_4 Cover (relative to herbaceous C_3 and C_4 herbaceous cover) against the explanatory variables (a) slrain1 (The ratio (log) of summer (Dec-Jan-Feb) to winter (Jun-Jul-Aug) rainfall totals) (b) maxtx (Maximum temperature hottest month (°C) , (c) PHC (Soil pH-CaCl2) (d) AWC (soil available water capacity %), and (e) SND (soil sand content %) derived from a GAM model constructed using TERN vegetation survey plot data. Blue lines are predicted outcomes of the model. Rugs were drawn to indicate observations with positive residuals (top of the plot) or negative residuals (bottom of the plot). Independent variables not depicted on the x-axis are held constant at their median value

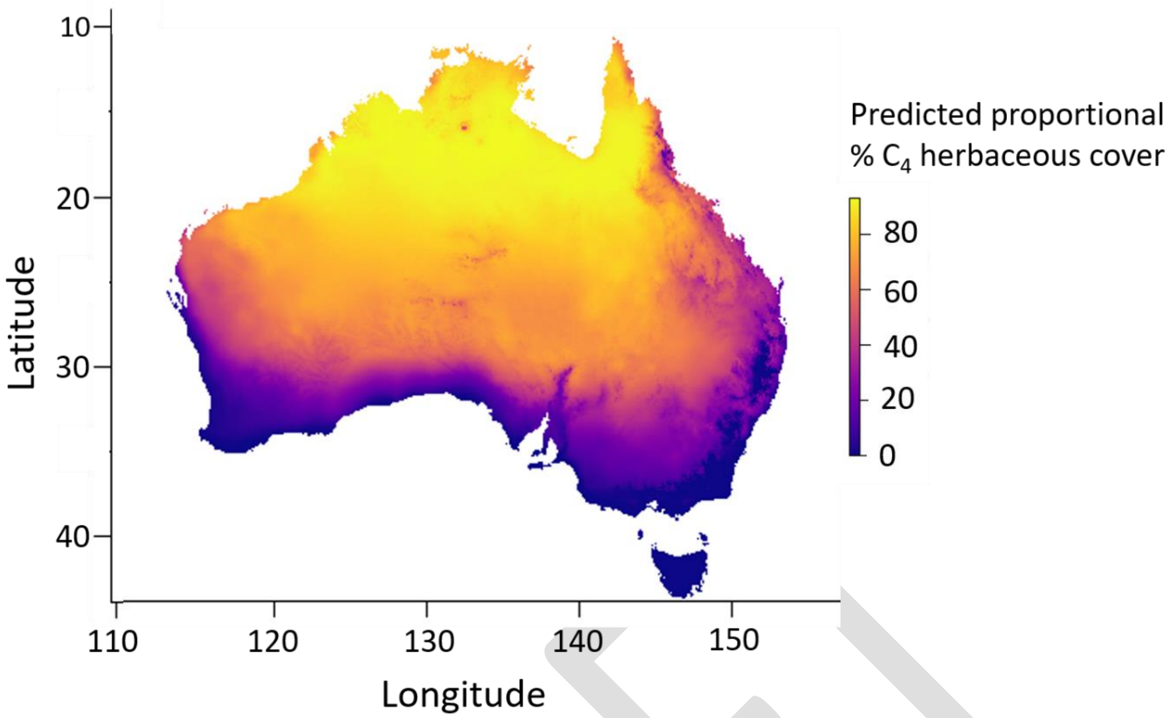

Fig 6. Predicted proportional (%) herbaceous C_4 cover (relative to herbaceous C_3 and C_4 herbaceous cover) extrapolated across Australia

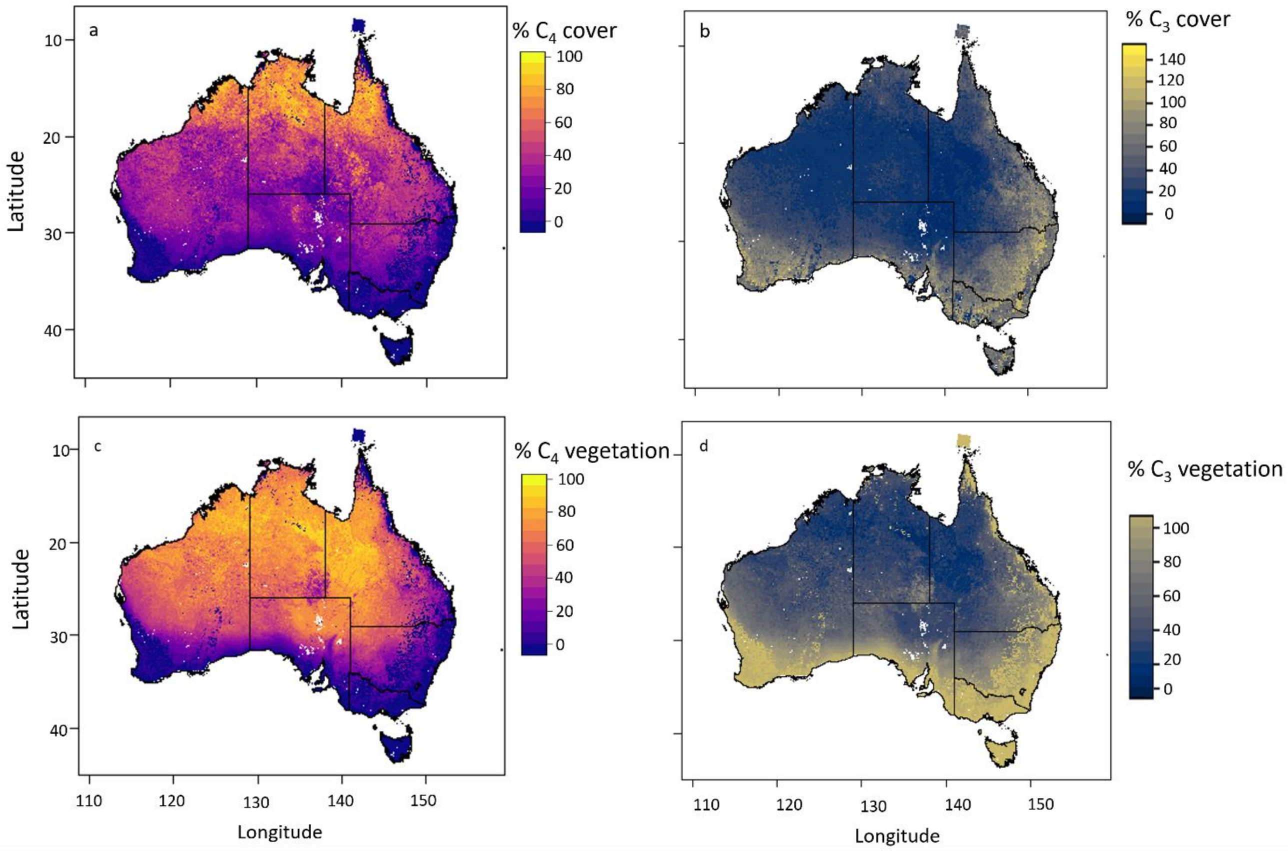

Fig. 7. % (a) C_4 and (b) C_3 cover, and proportional (c) C_4 and (d) C_3 vegetation cover (proportional to total vegetation) in 2015

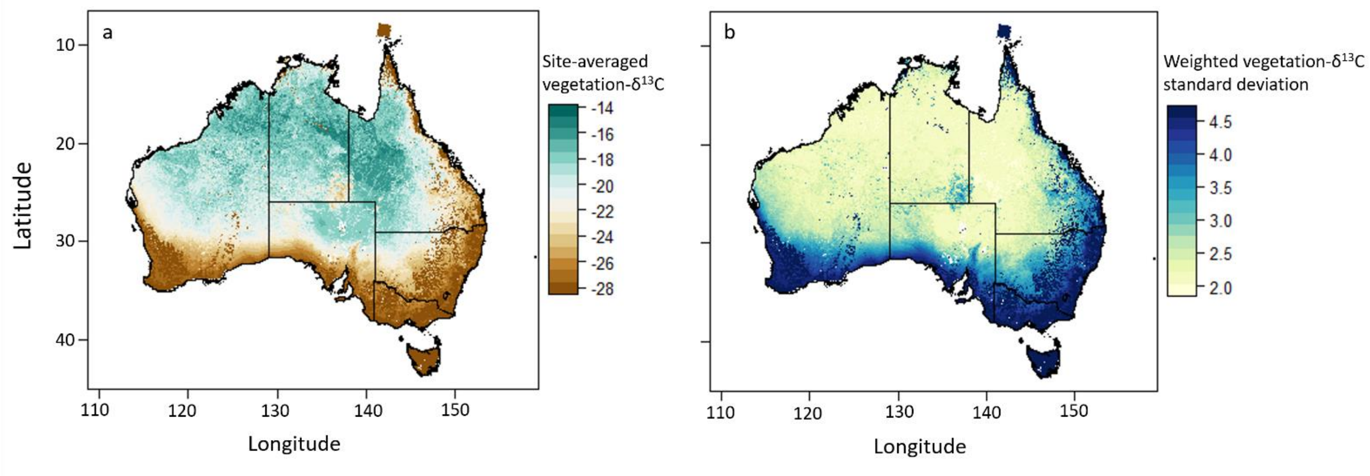

Fig. 8. a) Vegetation $\delta^{13}C$ isoscape of Australia corresponding to the year 2015 and b) weighted mean standard deviation of site-averaged $\delta^{13}C$ values

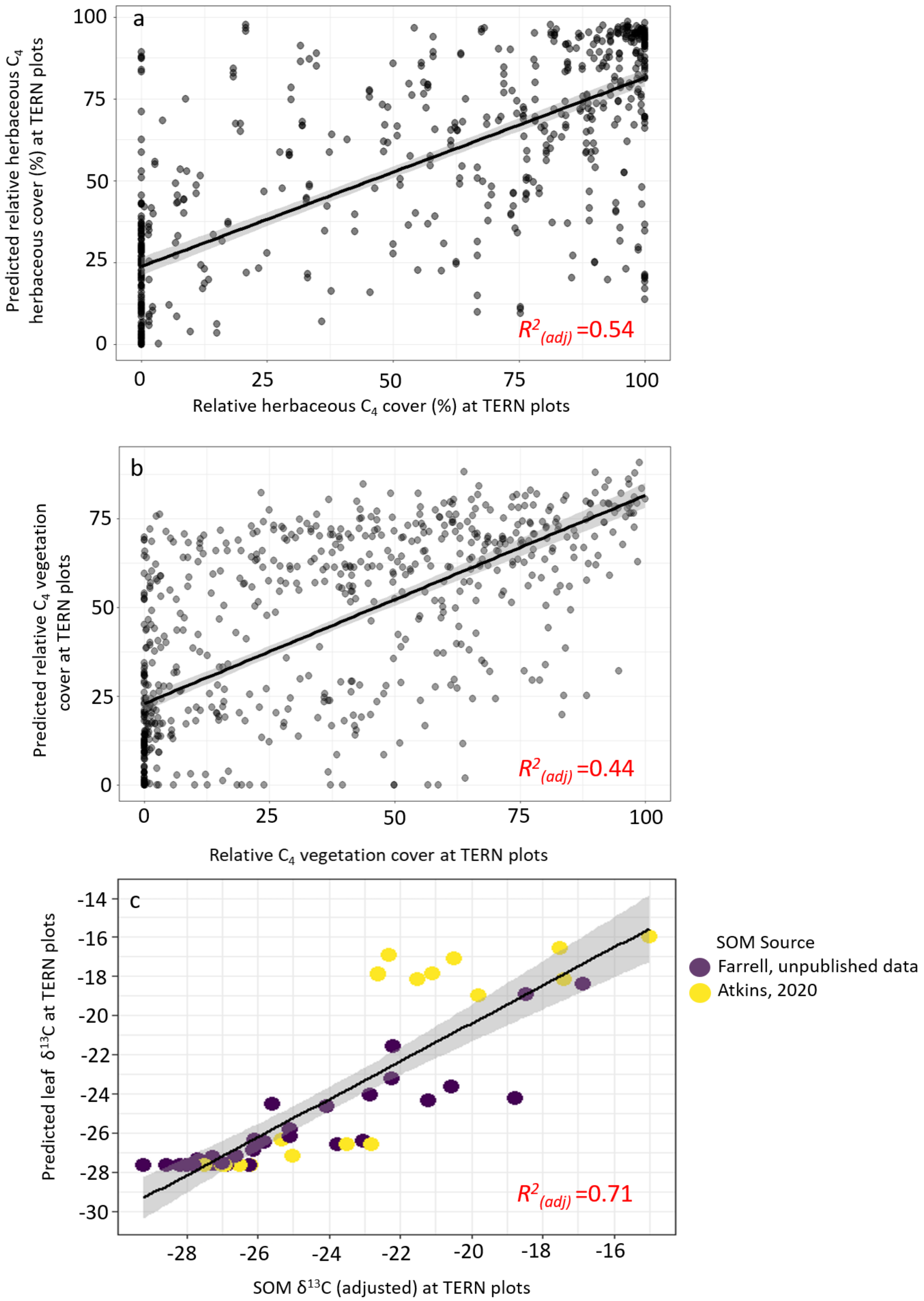

Fig 9. (a) Scatter plot of observed versus predicted relative % herbaceous C_4 Cover (relative to herbaceous C_3 cover) at TERN plots from 10-fold cross validation testing dataset, (b) predicted and observed relative C_4 vegetation cover (including woody cover) at all TERN plots, c) predicted leaf-$\delta^{13}C$ and measured Soil Organic Matter $\delta^{13}C$ at select TERN plots

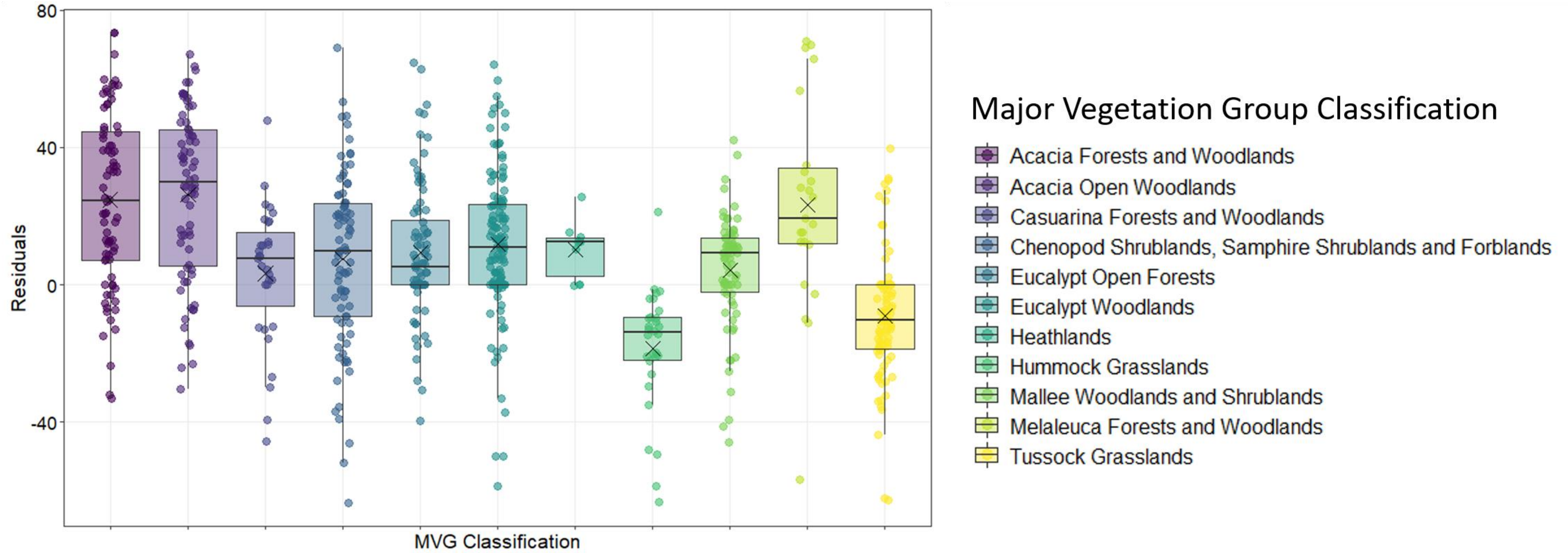

Fig. 10 Residuals of predicted and observed % C_4 vegetation cover (relative to total cover and including woody cover) at all TERN plots in major vegetation group (MVG) classifications. The box defines the second and third quartiles (likely range of variation), the vertical lines are the upper and lower quartiles. The black bands are the median residual values, the black **X** is the mean residual value for each classification.

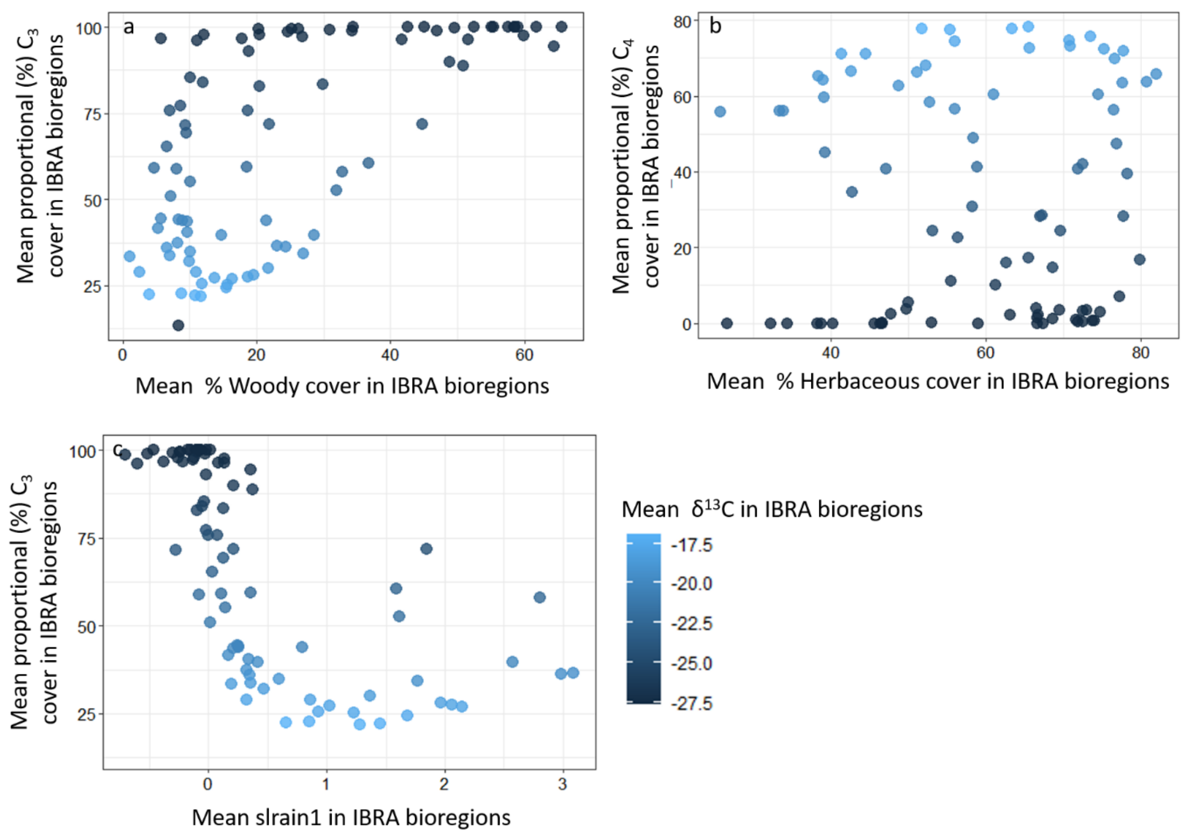

Fig. 11. (a) Scatterplot of predicted mean proportional C_3 Cover versus mean % woody cover (tree and shrub) across 86 different continental Australian bioregions, as described by the interim Biogeographic Regionalisation for Australia version 7 (IBRA 7.0; Department of Agriculture, Water and the Environment, 2020); (b) Scatterplot of predicted mean proportional C_4 Cover versus mean % herbaceous cover in different IBRA 7.0 ; (c) Scatterplot of predicted mean proportional C_3 cover versus mean slrain1 (The ratio (log) of summer (Dec-Jan-Feb) to winter (Jun-Jul-Aug) rainfall totals) across IBRA 7.0

Supplemental Matieral Captions

1. Supplemental Methods

2. Supplemental Results

3. Supplemental Results, IBRA Analysis