Soil carbon isotopic proxies for determining the photosynthetic pathway of floral communities: A method inter-comparison

Thesis submitted in accordance with the requirements of the University of Adelaide for an Honours Degree in Environmental Geoscience

Rachel Amber Atkins

April 2020

Word count: 7981

THE UNIVERSITY of ADELAIDE

SOIL CARBON ISOTOPIC PROXIES FOR DETERMINING THE PHOTOSYNTHETIC PATHWAY OF FLORAL COMMUNITIES: A METHOD INTER-COMPARISON

SOIL CARBON ISOTOPIC TOOLS FOR VEGETATION RECONSTRUCTION

ABSTRACT

C_3 and C_4 plants have a unique range of carbon isotope values as a result of their distinct photosynthetic processes. These differences are retained within leaf wax *n*-alkanes and bulk soil organic matter after entering the soil. As a result, soil organic matter and *n*-alkanes are used as isotopic tools for determining the proportion of C_4 versus C_3 vegetation cover for palaeoenvironmental reconstruction. However, unlike *n*-alkanes, soil organic matter is susceptible to isotopic enrichment due to decomposition, which affects the $\delta^{13}C$ values. The validity of these two methods as accurate estimations for proportional C_4 vegetation cover is yet to be tested in a comparative setting. In this study, these two methods were compared with ground vegetation surveys to determine whether they can be used interchangeably for palaeoenvironmental reconstruction. Surface soil samples were collected from 20 plots along a North to South transect through central Australia. These samples were analysed to determine their carbon isotopic composition, which was used to estimate the proportional C_4 cover at each locality. I hypothesised that both soil organic matter and *n*-alkanes would accurately reflect spatial trends seen in the C_3 and C_4 community. I also hypothesised that high proportional C_4 cover would correlate with high grass cover and low proportional C_4 cover would correlate with high tree cover. However, proportional C_4 cover would produce different relationships with climate variables depending on the method. Soil organic matter and *n*-alkane-derived proportional C_4 cover were positively correlated with the vegetation survey data. Therefore, all three approaches provided the same general geographical trend. All three methods also produced similar relationships with proportional grass and tree cover and climate variables. These results demonstrate that either isotopic approach can be used to document large-scale geographic trends in proportional C_4 cover without concern for variance associated with a particular method.

KEYWORDS

Stable carbon isotopic composition, leaf wax *n*-alkanes, soil organic matter, vegetation survey, C_4 plants, palaeoenvironmental proxies.

TABLE OF CONTENTS

Abstract i
Keywords i
List of Figures and Tables 3
1. Introduction 5
2. Background 7
2.1 C_3 and C_4 plants 7
2.2 Stable carbon isotopes in plants 11
2.3 Bulk soil organic matter 14
2.4 Long chain *n*-alkanes 17
2.5 Aims and hypotheses 19
3. Methods 20
3.1 Sample collection 20
3.2 Soil preparation 22
3.3 Bulk soil organic matter $\delta^{13}C$ analysis 23
3.4 Long chain *n*-alkane analysis 23
3.5 Proportional C_4 cover 24
4. Observations and Results 28
4.1 Bulk and compound specific carbon isotopic composition 28
4.2 Proportional grass, tree and chenopod cover 32
4.3 Correlations between proportional C_4 cover and climate 34
5. Discussion 36
5.1 Comparing estimates of proportional C_4 cover relative to vegetation surveys 36
5.2 Relationship between proportional C_4 cover and proportional grass, tree and chenopod cover 39
5.3 Relationship between proportional C_4 cover and climatic variables 40
5.4 Implications for palaeoenvironmental reconstruction 40
6. Conclusions 41
Acknowledgments 42
References 42
Appendix A. Plot descriptions 47
Appendix B. Step by step method 48
B.1 Sample preparation 48
B.2 Acidification 48
B.3 Centrifuge and freeze drier 48

B.4 Sample weighing..... 49
Appendix C. RStudio code..... 50
Appendix D. Carbon isotopic composition and standard deviation..... 56
Appendix E. Proportional cover and climate data..... 57

LIST OF FIGURES AND TABLES

Figure 1. A conceptual diagram modified from Sikes and Ashley (2007) demonstrating how carbon isotopic composition ($\delta^{13}C$) changes with proportional C_4 cover (% C_4). Herbivores are influenced by their food. Thus, faunal communities also change with floral composition. 11
Figure 2. Carbon isotopic composition ($\delta^{13}C$) displayed by C_3 and C_4 plants. Values were derived from the bulk tissue of 1000 specimens. The atmospheric CO_2 value at present is approximately 407 ppm and C_3 plants fractionate more than C_4 plants, thus they exhibit a more negative mean $\delta^{13}C$ value of -27‰, whereas C_4 plants have a more positive mean $\delta^{13}C$ value of -13‰. Figure taken from Glaser, 2005. 13
Figure 3. Location map of the 20 TERN plots across central Australia with the ecoregions as context. Ecoregion data was sourced from the World Wide Fund for Nature (2001). 21
Figure 4. Location map of the 20 TERN plots across central Australia with mean annual temperature (MAT) as context. Data was averaged across the 1970 to 2018 records from the Australian Gridded Climate Data set from the Australian Government Bureau of Meteorology (2018). 27
Figure 5. Location map of the 20 TERN plots across central Australia with seasonal water availability (SWA) as context. Data was averaged across the 1970 to 2018 records from the Australian Gridded Climate Data set from the Australian Government Bureau of Meteorology (2018). 28
Figure 6. The carbon isotopic composition ($\delta^{13}C$) of *n*-alkanes compared to the carbon isotopic composition of soil organic matter (SOM). The C_3 and C_4 end members for SOM and *n*-alkanes were also included to show the portion of $\delta^{13}C$ data that fell within the end members. 30
Figure 7. Difference between the proportional C_4 cover (% C_4) derived from the vegetation survey, soil organic matter (SOM) and *n*-alkanes compared to one another for each plot. 30
Figure 8. Proportional C_4 cover (% C_4) derived from bulk soil organic matter (SOM) and *n*-alkane carbon isotopic ($\delta^{13}C$) mixing models for each plot compared to % C_4 acquired from plot vegetation surveys. Solid lines are the predicted outputs of the zero-inflated beta regression models and the shaded bands are the standard error of the model. 31
Figure 9. Proportional C_4 cover (% C_4) derived from the *n*-alkane carbon isotopic ($\delta^{13}C$) mixing model for each plot compared to % C_4 derived from the soil organic matter (SOM) $\delta^{13}C$ mixing model for each plot. A Spearman's rank correlation (rho (p-value)) was also included. 32
Figure 10. Proportional C_4 cover (% C_4) derived from the vegetation survey and % C_4 derived from bulk soil organic matter (SOM) and *n*-alkane carbon isotopic ($\delta^{13}C$) mixing models for each plot plotted by (a) grass cover, (b) tree cover and (c) chenopod cover, derived from the vegetation survey data. Solid lines are the predicted outputs of the zero-inflated beta regression models and the shaded bands are the standard error of the model. 33
Figure 11. Proportional C_4 cover (% C_4) derived from the vegetation survey and % C_4 derived from bulk soil organic matter (SOM) and *n*-alkane carbon isotopic ($\delta^{13}C$) mixing models for each plot plotted versus mean annual temperature (MAT) averaged across the 1970 to 2018 records from the Australian Gridded Climate Data set from the

Australian Government Bureau of Meteorology (2018). Solid lines are the predicted outputs of the zero-inflated beta regression models and the shaded bands are the standard error of the model..... 35
Figure 12. Proportional C_4 cover (% C_4) derived from the vegetation survey and % C_4 derived from bulk soil organic matter (SOM) and *n*-alkane carbon isotopic ($\delta^{13}C$) mixing models for each plot plotted versus seasonal water availability (SWA) averaged across the 1970 to 2018 records from the Australian Gridded Climate Data set from the Australian Government Bureau of Meteorology (2018). Solid lines are the predicted outputs of the zero-inflated beta regression models and the shaded bands are the standard error of the model..... 35

Table 1. Vegetation survey proportional C_4 cover (% C_4) was obtained from the Terrestrial Ecosystem Research Network (TERN) database. Carbon isotopic composition ($\delta^{13}C$) and % C_4 derived from bulk soil organic matter (SOM) and *n*-alkanes for each plot. ……..29
Table 2. Results of zero-inflated beta regression analysis of the proportional C_4 cover (% C_4) derived from soil organic matter (SOM), *n*-alkanes and the vegetation survey versus proportional grass, tree and chenopod cover derived from the vegetation survey data. μ and ν beta coefficients are provided for each model, as well as p-values which denote the statistical significance of the predictor for each sub-model (pseudo-R^2 values are also provided for each model). ……34
Table 3. Results of zero-inflated beta regression analysis of the proportional C_4 cover (% C_4) derived from soil organic matter (SOM), *n*-alkanes and the vegetation survey versus mean annual temperature (MAT) and seasonal water availability (SWA) average across the 1970 to 2018 records from the Australian Gridded Climate Data set from the Australian Government Bureau of Meteorology (2018). μ and ν beta coefficients are provided for each model, as well as p-values which denote the statistical significance of the predictor for each sub-model (pseudo-R^2 values are also provided for each model). ……36

1. INTRODUCTION

Plants photosynthesise using three different photosynthetic pathways; C_3, C_4 and crassulacean acid metabolism (CAM). Of these pathways, plants predominantly use C_3 photosynthesis. However, C_4 species account for a quarter of the total global productivity (Ehleringer, Cerling, & Helliker, 1997). The balance of C_3 and C_4 vegetation has changed over geological time and will continue to in the future due to climate (Spicer, 1993). Determining the climatic and ecological drivers of modern spatial patterns in C_4 plants can allow us to estimate the composition of past C_3 and C_4 vegetation cover and reconstruct palaeoenvironments (Hartman, 2011). The ability to accurately estimate the proportion of C_3 and C_4 vegetation cover both in modern and ancient soils will provide us with an understanding of how changes in floral communities effect carbon storage, agriculture, migration and vertebrate evolution (Ehleringer, Cerling, & Helliker, 1997). This knowledge will also allow us to determine the impact of global warming on the distribution of plants, which are the basis for most ecosystems (Chapin, Rincon, & Huante, 1993).

Organic carbon fixed by plants is incorporated into the soil by their roots and is also released by plant tissue during decay (Pausch & Kuzyakov, 2017). Due to physiological and structural differences between the different pathways, C_3 and C_4 plants have unique carbon isotopic compositions ($\delta^{13}C$) (Dawson et al., 2002; West et al., 2006). Therefore, photosynthetically fixed carbon deposited in soil by plants can be identified as C_3, C_4, or mixed in origin based on its carbon isotopic composition (Staddon, 2004). Because plant material is the major source of carbon for the soil environment (Bird & Pousai, 1997), proportional changes in C_3 and C_4 biomass over time and space can be

reconstructed using the $\delta^{13}C$ values derived from bulk soil organic matter (SOM) or leaf wax *n*-alkanes (once they have entered the soil) (Rao et al., 2008). Numerous studies have used SOM and *n*-alkane soil $\delta^{13}C$ values to evaluate changes in relative C_4 cover (Bird et al., 1995; Hartman, 2011; Krull & Bray, 2005; Rao et al., 2008) and the environmental conditions (Bird et al., 2003; Wynn & Bird, 2008) that influence variations observed in carbon isotopic composition.

Bulk SOM $\delta^{13}C$ values have been found to accurately reflect regional changes in proportional C_4 cover (Krull & Bray, 2005; Rao et al., 2008). However, varying decomposition rates between C_3 and C_4-derived carbon and further decomposition of the bulk soil is thought to hinder the ability of SOM $\delta^{13}C$ values to identify differences in vegetation biomass (Bird & Pousai, 1997). Therefore, *n*-alkanes have been explored as an alternative assessment of vegetation change as they are more robust to decomposition and their $\delta^{13}C$ values are less prone to variations caused by decomposition (Li et al., 2017). However, *n*-alkanes have the ability to become airborne, which causes the local $\delta^{13}C$ values to appear more C_3 or C_4 dominated than the vegetation cover due to the incorporation of *n*-alkanes from distant sites (Schreuder et al., 2018). Despite these potential problems, minimal research has been conducted to test the validity of SOM and *n*-alkane $\delta^{13}C$ values as accurate estimations for C_3 and C_4 vegetation cover.

The primary aim of this study was to determine if SOM and *n*-alkane $\delta^{13}C$ values accurately reflect spatial trends in proportional C_4 cover (% C_4). This was achieved by analysing surface soil samples from 20 plots along a North to South transect through

central Australia that included a diverse mix of C_3-dominated, mixed, and C_4-dominated habitats. For both SOM and *n*-alkane $\delta^{13}C$ values, % C_4 estimates were calculated using a two-source mixing model and then compared to % C_4 estimates based on ground surveys of vegetation cover. I evaluated the ability of SOM and *n*-alkanes to reflect both regional trends in % C_4 cover as well as within-site accuracy (relative to % C_4 vegetation cover). In addition, I also examined the relationship between % C_4 and proportional grass, tree and chenopod cover, as well as the relationship between % C_4 and climate variables to determine whether different approaches would lead to different C_4-climate trends. This will help to determine if each approach can be used interchangeably and compared between studies and regions when examining C_4 distribution over time and space.

2. BACKGROUND

2.1 C_3 and C_4 plants

Plants utilise three types of photosynthetic pathways; the primitive C_3 photosynthetic pathway, and the more evolutionarily recent C_4 and CAM pathways (Ehleringer, Cerling, & Helliker, 1997). C_3 photosynthesis is the most commonly exhibited photosynthetic pathway, with a global coverage of 87.4 million km^2 (Still et al., 2003). While only 3% of total plant species is characterised by the C_4 photosynthetic pathway (Still et al., 2003), C_4 plants contribute approximately 25% of global terrestrial primary production (Ehleringer, Cerling, & Helliker, 1997). CAM has evolved independently multiple times from C_3 ancestors and is found in 6% of vascular plant species, such as succulents (Silvera et al., 2010). However, CAM photosynthesis is less dominant in

plants, therefore, environmental studies mainly focus on C_3 and C_4 plants. Trees and most shrubs predominantly use the C_3 photosynthetic pathway (Schaedle, 1975), while grasses and Chenopodiaceae, a suite of drought and salt tolerant shrubs, include C_3 and C_4 plants (Shomer-Ilan, Nissenbaum, & Waisel, 1981).

C_3 and C_4 plants are adapted to distinct climate conditions as a result of their unique photosynthetic processes. In particular, C_4 plants are more efficient at converting solar energy into glucose at high temperatures, compared to C_3 plants (Collatz, Berry, & Clark, 1998). During the fixation of carbon by rubisco (RuBP), C_3 plants experience high levels of photorespiration. Photorespiration occurs when RuBP functions as an oxygenase to generate 2-phosphoglycolate (2-PG). This process reduces the efficiency of carbon fixation due to a net loss of carbon dioxide (CO_2) for the plant (Bauwe, Hagemann, & Fernie, 2010; Bräutigam & Gowik, 2016). Photorespiration is heightened in warmer conditions as the specificity of RuBP for CO_2 relative to oxygen (O_2) decreases with increasing temperature (Peterhansel et al., 2010). The photosynthetic efficiency of C_4 plants, however, is not influenced by photorespiration due to their ability to concentrate CO_2 with RuBP, aided by their Kranz anatomy (Lundgren, Osborne, & Christin, 2014). Kranz anatomy describes the wreath-like formation of the mesophyll cells surrounding the bundle sheath cells (Sedelnikova, Hughes, & Langdale, 2018). In C_4 plants carbon fixation and the Calvin cycle are separated into the mesophyll and bundle sheath cells, respectively, which allows RuBP to metabolise more CO_2 in the absence of O_2 (O'Leary, 1988). However, C_4 photosynthesis is more energetically expensive than the C_3 photosynthetic pathway, making it the less efficient photosynthetic pathway in cooler climates (Osmond, 1981).

Over the past several decades, a substantial body of research has investigated what factors influence the distribution and abundance of C_3 and C_4 plants (Ehleringer, 1978; Farquhar, Buckley, & Miller, 2002). Elevated CO_2 levels, like those experienced during the Mesozoic (252 to 66 Ma), favoured the dominance of the C_3 pathway. High CO_2 levels help supress photorespiration, even in warmer climates (Sage & Stata, 2015). However, since the late Oligocene (approximately 23 Ma) the concentration of atmospheric CO_2 has fallen, while O_2 levels have risen from 18% to 21% as of present (Berner, VandenBrooks, & Ward, 2007). As a result, the efficiency of C_3 photosynthesis decreased. Thus, it is widely believed that the reduction of atmospheric CO_2 promoted the evolutionary diversification of C_4 lineages (Arakaki et al., 2011).

The distribution of C_3 and C_4 plants is also controlled by temperature and rainfall. Ehleringer (1978) proposed that temperature acts on the quantum yield of C_3 and C_4 plants. The quantum yield is the ratio of moles of CO_2 assimilated to moles of photosynthetically active radiation (available light required for photosynthesis) absorbed by a leaf. The quantum yield is dependent on both CO_2 levels and temperature. As temperature increases, the quantum yield in C_3 plants decreases as a result of photorespiration. The quantum yield in C_4 plants, however, is unaffected across a range of temperatures due to their Kranz anatomy and their carbon concentrating mechanism. Based on this hypothesis, at modern CO_2 levels, the quantum yield of C_4 grasses surpasses that of C_3 grasses at a temperature of approximately 22°C (Collatz, Berry, & Clark, 1998; Ehleringer, 1978). Thus, C_4 plants have a competitive advantage over C_3 plants in warm environments, where they are more abundant (Still et al., 2003).

C_4 plants are also more water-use efficient as they are able to maintain carbon fixation while reducing water loss through their stomates (Sage, Christin, & Edwards, 2011). Therefore, they dominate regions with warm, summer rainfall (Murphy & Bowman, 2007). In contrast, C_3 plants experience a trade-off between CO_2 uptake and water loss, where the more CO_2 they take in through their stomates, the more water they lose (Akita & Moss, 1972). Therefore, C_3 plants favour cool-temperate environments with high rainfall (Murphy & Bowman, 2007). Arid environments have expanded globally due to the general cooling of Earth as a result of decreased atmospheric CO_2 (Zachos, Dickens, & Zeebe, 2008). This has provided new ecological opportunities for C_4 plants, allowing plants, such as grasses, to expand and display ecological dominance in tropical and subtropical environments (Arakaki et al., 2011).

The knowledge of what factors control the distribution and abundance of C_3 and C_4 plants is important as plants form the basis of most ecosystems and provide unique habitats for organisms (Cardinale et al., 2011). Different photosynthetic types impact the structure of an ecosystem as they have a "bottom-up" effect, where species at subsequent trophic levels are affected by the primary producers. This influences the distribution, migration and evolution of vertebrates (Ehleringer, Cerling, & Helliker, 1997) (Figure 1). For example, the expansion of C_4 plants led to a global rise in C_4-dominated grasslands, which promoted the evolution of grassland-adapted species (Bouchenak-Khelladi et al., 2009). Additionally, by reconstructing past vegetation cover using carbon isotopic analysis we are able to see how the distribution of prehistoric C_3 and C_4 plants was influenced by environmental factors to better understand how similar conditions could impact floral communities today (Still et al., 2003). However, the best

climate variables to predict C_4 distribution is not always agreed upon (Hattersley, 1983). Although it is generally agreed that temperature and rainfall affect C_4 distribution, specific factors and how they interact is still unclear. For example, crossover temperature, minimum summer temperature and water supply are all argued to be the dominant driver of C_3 and C_4 distribution (Collatz, Berry, & Clark, 1998; Hattersley, 1983; Teeri & Stowe, 1976;). Thus, this is an ongoing field of research and the relationship between climate and C_4 expansion is still debated. This highlights the importance of researching and comparing different carbon isotopic approaches to estimate proportional C_4 cover as these methods are key to understanding the impacts of global warming, through palaeoenvironmental reconstruction.

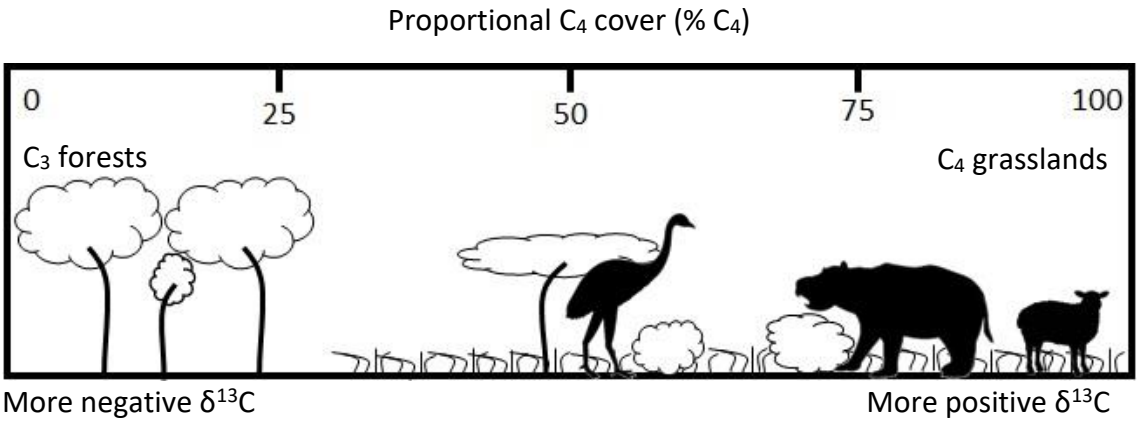

Figure 1. A conceptual diagram modified from Sikes and Ashley (2007) demonstrating how carbon isotopic composition ($\delta^{13}C$) changes with proportional C_4 cover (% C_4). Herbivores are influenced by their food. Thus, faunal communities also change with floral composition.

2.2 Stable carbon isotopes in plants

Isotopes of an element that do not undergo radioactive decay are referred to as stable isotopes (West et al., 2006). Stable isotope abundances are expressed as ratios in delta (δ) notation, relative to a standard using the following equation (West et al., 2006):

$$\delta^{13}C = \left(\left(\frac{R_{sample}}{R_{standard}}\right) - 1\right) \times 1000 \quad (1)$$

Where $\delta^{13}C$ is the carbon isotopic composition of a sample. R_{sample} is the ratio of the concentration of $^{13}C/^{12}C$ in the sample, and $R_{standard}$ is the ratio in the standard. The two naturally occurring stable isotopes of carbon are ^{12}C and ^{13}C, which contribute 98.9% and 1.1% to the total stable carbon pool, respectively (Farquhar, Ehleringer, & Hubick, 1989; Glaser, 2005; O'Leary, 1988).

Due to the physiological and structural differences between the different photosynthetic pathways, C_3 and C_4 plants have unique carbon isotopic values. This is due to the different isotopic fractionation processes that occur during photosynthesis for each type of pathway. Isotopic fractionation is defined as the relative partitioning of heavy and light isotopes between two phases in a system and is a consequence of the small mass-dependant differences in the physical and chemical behaviour of isotopes (Dawson et al., 2002). In the case of plants, fractionation occurs when CO_2 is fixed from the atmosphere and used to create glucose. As a result, plants have a much more negative isotopic value than the surrounding atmosphere as they preferentially take up lighter carbon isotopes. This preference is the result of CO_2 with the heavy carbon isotopes having a greater bond strength, making it more difficult to break during photosynthesis (Glaser, 2005). However, because the processes of C_4 photosynthesis occur in two cells, RuBP is physically isolated from the stomatal opening to the atmosphere and cannot fractionate in C_4 species as much as C_3 (O'Leary, 1988). Therefore, C_3 and C_4 plants exhibit a different range of isotopic values (Figure 2). For the bulk tissue, the vast

majority of C_3 plants exhibit a range of $\delta^{13}C$ values from -22 to -32‰ and average of about -27‰, while the vast majority of C_4 plants display a range of $\delta^{13}C$ values from -11 to -18‰ and average of about -13‰ (Figure 2) (Glaser, 2005; West et al., 2006).

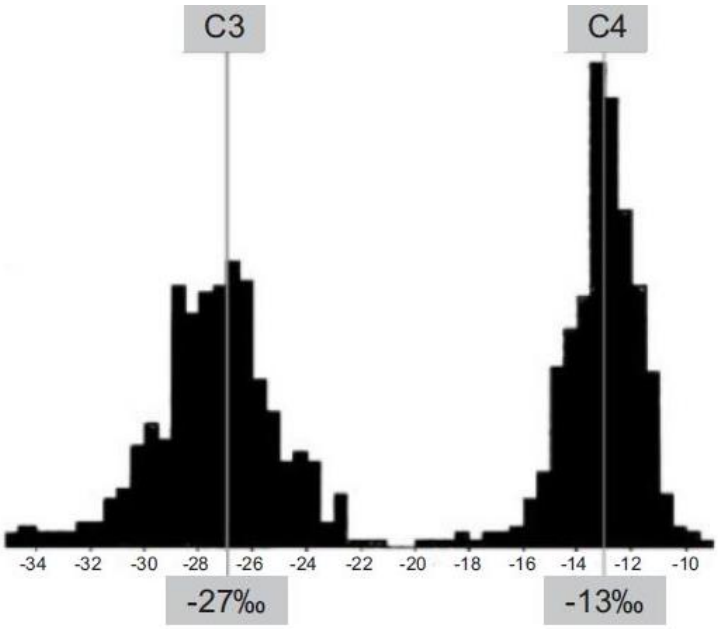

Figure 2. Carbon isotopic composition ($\delta^{13}C$) displayed by C_3 and C_4 plants. Values were derived from the bulk tissue of 1000 specimens. The atmospheric CO_2 value at present is approximately 407 ppm and C_3 plants fractionate more than C_4 plants, thus they exhibit a more negative mean $\delta^{13}C$ value of -27‰, whereas C_4 plants have a more positive mean $\delta^{13}C$ value of -13‰. Figure taken from Glaser, 2005.

Due to these isotopic differences, the stable carbon isotope analysis of plants and soils is commonly used to study the spatial and temporal changes in C_3 and C_4 vegetation abundance (Dawson et al., 2002; West et al., 2006). Palaeoenvironmental information can be derived from $\delta^{13}C$ values of prehistoric C_3 and C_4 plant carbon, which is often preserved through the burial of ancient organic matter, such as plant leaf waxes, within soil, and used to reconstruct the proportional C_4 cover of past ecosystems (Meyers, 1994; Rao et al., 2008).

The $\delta^{13}C$ values of ancient organic matter can also provide information on past climate conditions. If the proportional C_4 cover is known and we are able to model the C_4-climate relationship, then we can predict past climate conditions. Similarly, it can help us to predict the response of floral communities to future climate change (Salzmann, Haywood, & Lunt, 2009). For example, Salzmann, Haywood, and Lunt (2009) used palaeobotanical data from numerous sites across the globe to determine how vegetation from the middle Pliocene (approximately 3 Ma) could be used as an insight for global warming. They concluded that the distribution of vegetation 3 Ma was similar to that of modern-day vegetation under elevated atmospheric CO_2. Thus, reconstructing the distribution of C_3 and C_4 plants is a viable predictor of the impact of climatic conditions on vegetation cover. SOM and leaf wax *n*-alkanes are two forms of organic matter that can be used for this purpose.

2.3 Bulk soil organic matter

Plant carbon derived from the fixation of photosynthetic CO_2 acts as the main source of carbon in the soil organic matter, which is the major terrestrial carbon reservoir (Bird & Pousai, 1997). When leaf litter falls to the soil surface the $\delta^{13}C$ values of different plants are incorporated into the SOM relative to the above-ground plant biomass, preserving an average $\delta^{13}C$ value that broadly reflects the % C_4 at the time of burial (Pausch & Kuzyakov, 2017). The incorporation of carbon into the soil occurs sequentially over time and by measuring the $\delta^{13}C$ of the soil at different depths, evidence for current and past floral communities can be provided (Ambrose & Sikes, 1991). The proportional C_4 biomass can be calculated from SOM $\delta^{13}C$ values using a two-end member mixing model. This determines the proportion of the carbon isotopic composition that is C_3 and

C_4-derived (Phillips, 2012). SOM $\delta^{13}C$ values are also able to provide assessments of spatial trends in proportional C_4 biomass and continental scales (Bird & Pousai, 1997).

Field studies have been conducted to assess the relationship between vegetation changes observed over time and the corresponding $\delta^{13}C$ values of bulk SOM. One such study conducted by Krull and Bray (2005) compared vegetation changes recorded by SOM $\delta^{13}C$ values with the same changes captured by aerial photographs. They found that while SOM $\delta^{13}C$ values can be used to reconstruct intermediate and long-term changes in the proportion of C_3 and C_4 vegetation, it is only applicable to areas where the vegetation change is older than ten years, otherwise the change in carbon signal is not strong enough to detect. The site of interest also has to have a clear distinction between the photosynthetic pathways of grasses and trees (Krull and Bray, 2005). SOM $\delta^{13}C$ values have also been found to accurately reflect the proportion of C_3 and C_4 plants when comparing the vegetation cover of completely different biomes (Wang et al., 2008). Therefore, SOM $\delta^{13}C$ values are good at detecting general regional trends. However, while bulk SOM is a good option for isotopic analysis, the method is still associated with some issues.

Small variations in vegetation at a local scale are difficult to detect using SOM $\delta^{13}C$ values as numerous and complex influencing factors can affect $\delta^{13}C$ (Bird & Pousai, 1997; Krull & Bray, 2005; Saiz et al., 2015; Wynn & Bird, 2007; Wynn, Bird, & Wong, 2005). The use of SOM to examine changes in C_3 and C_4 plant biomass relies on the assumption that there is no inherent difference between the behaviour of C_3 and C_4-derived organic carbon once it enters the soil. However, isotopic data from SOM

suggests an imbalance between the vegetation cover and the fraction of C_4-derived carbon represented in the soil (Wynn & Bird, 2007). This disparity is partly caused by the more rapid decomposition of C_4-derived organic carbon compared to its C_3 counterpart (Wynn & Bird, 2007), which leads to an underestimation of the proportional C_4 biomass. Additionally, SOM further fractionates due to the microbial respiration of CO_2 from the soil. The breakdown of organic matter causes an enrichment of ^{13}C in the soil, making the vegetation appear more C_4 dominated. The process of decomposition further produces depth-enriched $\delta^{13}C$ profiles, as the degree of decomposition (age of SOM) increases with depth (Wynn, Bird, & Wong, 2005). Therefore, older SOM $\delta^{13}C$ values are associated with greater uncertainty due to the processes that occur during decomposition (Krull & Bray, 2005). The isotopic composition of the bulk soil is also modified by edaphic factors, such as texture. C_4-derived carbon has a greater preservation potential in the finer soil fractions compared to C_3 carbon (Bird & Pousai, 1997). However, differently sized soil particles also support specific microbial communities that may influence the rate of decomposition, regardless of the preferential incorporation of C_3 or C_4-derived carbon in a particular sized particle (Hemkemeyer et al., 2018). In addition to the varying rates of decomposition, more C_3-derived carbon is produced per unit area of ground cover than that of C_4 species in mixed C_3 and C_4 environments due to the greater biomass of C_3 plants (Bird & Pousai, 1997). Furthermore, pools of SOM $\delta^{13}C$ values that have been stabilised for millions of years may be impacted by more recent shifts in the C_3 and C_4 cover (Saiz et al., 2015).

Variations observed in SOM $\delta^{13}C$ values are often addressed by applying a simple correction (such as 1‰ (Bird & Pousai, 1997)) to the end members prior to the mixing

model for the effects of isotopic fractionation during decomposition (Bird & Pousai, 1997; Krull & Bray, 2005; von Fischer, Tieszen, & Schimel, 2008). However, this simplification does not account for the interaction between multiple sources of variation, such as faunal input, contribution from outside sources and fluctuating rates of decomposition. Furthermore, the amount of ^{13}C enrichment is variable between habitats and environmental conditions. Thus, a single, universal correction value introduces uncertainty when attempting to constrain the relationship between SOM and proportional C_4 cover.

2.4 Long chain *n*-alkanes

Long chain *n*-alkanes, constituents of plant leaf waxes, are long-lived in the sedimentary record, allowing them to serve as an alternative tool for estimating the proportions of C_3 and C_4 vegetation cover (Bush & McInerney, 2015; Li et al., 2017). Long chain *n*-alkanes, containing 25 to 35 carbon atoms, are produced nearly exclusively by terrestrial plants, making them valuable biomarkers for reconstructing palaeovegetation change (Diefendorf & Freimuth, 2017). Leaf-derived *n*-alkanes make a large contribution to the organic content of the soil due to the greater biomass of leaves and the higher concentration of *n*-alkanes in leaf waxes compared to other plant tissue (Diefendorf & Freimuth, 2017; Gamarra & Kahmen, 2015; Howard et al., 2018). The chain length distribution of *n*-alkanes varies depending on the preferential production of different chain lengths by the corresponding plant. While all plants produce a range of chain lengths, trees and grasses produce more C_{29} and C_{33} chain lengths, respectively (Howard et al., 2018). *n*-Alkanes have carbon isotope ratios that are representative of the photosynthetic pathway of the plant that produced them (Garcin et al., 2014). For *n*-alkanes with a chain length of C_{31} the average $\delta^{13}C$ values in C_3 and C_4 plants are -

35.3‰ and -17.8‰, respectively (Garcin et al., 2014). Unlike SOM, *n*-alkanes are resistant to degradation and less affected by environmental factors. Thus, the effect of early diagenesis does not interfere with their carbon isotopic signatures during burial in the soil (Li et al., 2017). While the abundance of *n*-alkanes decreases during early diagenesis this process has little to no effect on their carbon isotopic composition (Li et al., 2017). This allows *n*-alkane $\delta^{13}C$ values to serve as plant biomarkers, which can further be used to determine the proportion of C_3 and C_4-derived carbon for palaeoenvironmental reconstruction (Bird et al., 1995; Bush & McInerney, 2014; Garcin et al., 2014).

Studies have been conducted to validate the reliability of *n*-alkane $\delta^{13}C$ values as valuable palaeoclimate and palaeoenvironmental proxies. For example, Bird et al. (1995) used the carbon isotopic composition of terrestrial *n*-alkanes in marine sediment samples to determine if they reflect the proportion of C_3 and C_4-derived carbon in the adjacent vegetation. They found that the *n*-alkanes accurately represented the floral community. Thus, *n*-alkanes are suited for observing general trends and regional C_3 and C_4 vegetation biomass (Bird, Kracht, & Chivas, 1994; Howard et al., 2018). However, small changes at a local scale are difficult to accurately detect from terrestrial *n*-alkanes due to certain limitations. *n*-Alkanes have the potential to become airborne, which would lead to the incorporation of *n*-alkanes from distant sites (Diefendorf & Freimuth, 2017; Howard et al., 2018; Schreuder et al., 2018). This additional input would cause the $\delta^{13}C$ values to appear more C_3 or C_4 than the actual vegetation biomass of the site. While *n*-alkanes are recognised as useful plant biomarkers for palaeoenvironmental reconstruction, the method requires more testing to determine the accuracy of *n*-alkane

$\delta^{13}C$ values. Hence the importance of researching carbon isotopic tools in a comparative study. Even though there are clear differences between SOM and *n*-alkanes the two methods are often used interchangeably with little confirmation of the validity of using SOM and *n*-alkanes in such a way. These aforementioned differences between the two approaches may lead to inconsistencies amongst the $\delta^{13}C$ values acquired by each method for the same site. Despite this, the methods are often presumed to provide similar $\delta^{13}C$ values and trends.

2.5 Aims and hypotheses

While research into the proportion of photosynthetic types within Australia has been conducted using bulk SOM (Krull & Bray, 2005) and *n*-alkane $\delta^{13}C$ values (Bird et al., 1995), these methods are yet to be tested extensively in a comparative setting. Using both techniques is often prohibitively expensive, hence palaeoenvironmental reconstruction studies typically utilise a single method. However, it is uncertain how truly comparable bulk SOM and *n*-alkane $\delta^{13}C$ values are. Thus, the overarching aim of this project was to compare bulk SOM and *n*-alkane $\delta^{13}C$ values as a tool to estimate C_3 versus C_4 vegetation cover in Australia. This was tested by comparing the proportional C_4 cover estimates derived from bulk SOM and *n*-alkane $\delta^{13}C$ values to ground surveys of vegetation cover for 20 plots along a North to South transect through central Australia. I hypothesised that:

- Both SOM and *n*-alkane $\delta^{13}C$ values will accurately reflect spatial trends seen in the C_3 and C_4 community. However, *n*-alkanes will be more strongly correlated with the vegetation cover as they are not affected by the degradation that occurs in SOM.

- High proportional C_4 cover will correlate with increased grass cover, as both arid-adapted and tropical grasses are C_4, and low proportional C_4 cover will correlate with increased tree cover, as trees use the C_3 photosynthetic pathway.
- Each method will produce a different relationship between C_4 cover and climatic variables. SOM and *n*-alkane $\delta^{13}C$ values are impacted by different environmental conditions, such as decomposition. Therefore, when using proportional C_4 cover estimated from these approaches to define the relationship between C_4 plants and climate, different results will be observed depending on the method.

This research will help to determine if each approach can be used interchangeably and compared when examining C_4 distribution over time and space and in relation to climate. If SOM, *n*-alkanes and vegetation surveys are interchangeable, then any of the three techniques can be chosen to reconstruct vegetation cover in palaeoenvironmental studies. Their comparability will also allow us to more readily compare studies from different locations and times.

3. METHODS

3.1 Sample collection

Vegetation surveys were conducted by the Terrestrial Ecosystem Research Network (TERN), using the AusPlots Rangelands survey methodology (White et al., 2012) at 20 plots along a North to South transect through central Australia between 2011 and 2013 (Figure 3). TERN plots (one hectare, 100 x 100m) are permanently established sites in a homogenous area of vegetation. Transects (10 x 100m long) were laid out within each

plot in a 5 x 5 grid pattern and species occurrences were recorded at each point (1m) within the transect, providing the vegetation cover at 1010 locations per plot. The soil sample acquired from the top 3cm at location 5 in each plot was air dried and stored in calico bags (Howard et al., 2018) and used for this study.

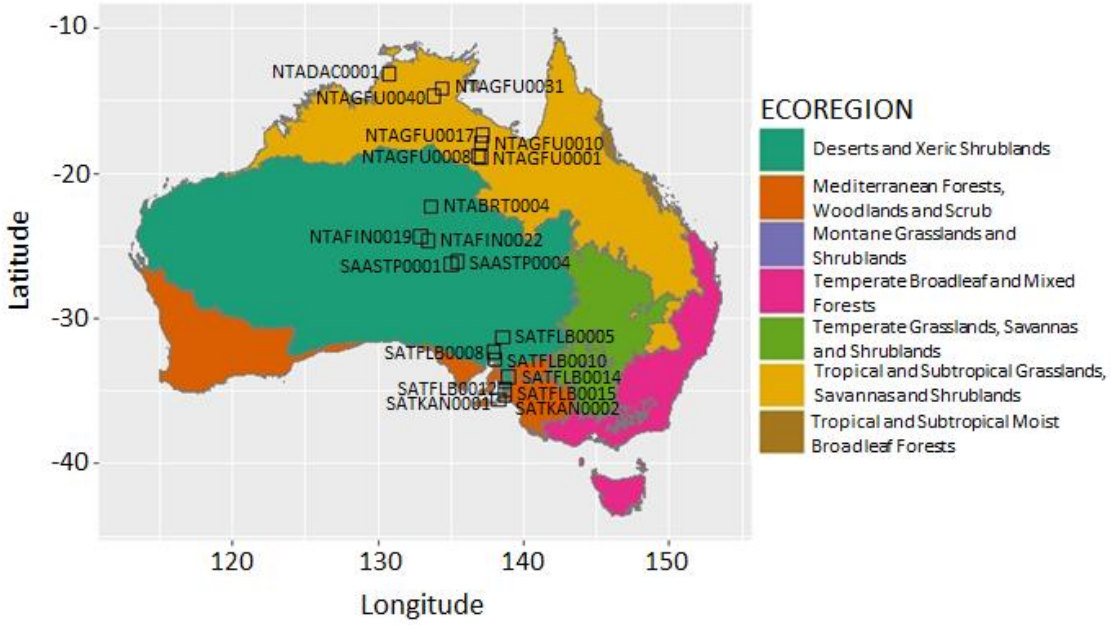

Figure 3. Location map of the 20 TERN plots across central Australia with the ecoregions as context. Ecoregion data was sourced from the World Wide Fund for Nature (2001).

The plots used in this study were the same as those utilised by Howard et al. (2018), whose distribution transitioned through several Australian ecoregions (see Appendix A for detailed plot descriptions). The northern most ecoregion consisted of tropical and subtropical grasslands, savannas and shrublands characterised by large expanses of land that do not receive enough rainfall to support extensive tree cover. Deserts and xeric shrublands make up the central ecoregion, where temperature extremes are prevalent, and evaporation exceeds rainfall. Although harsh, the climatic diversity of this ecoregion supports a rich array of mostly ephemeral habitats. The southernmost

ecoregion includes Mediterranean forests, woodlands and shrubs characterised by hot, dry summers and cool, moist winters. They feature a broad diversity of animals and plants, which are mostly fire adapted (World Wide Fund for Nature [WWF], 2001).

3.2 Soil preparation

Organic matter and rock fragments were removed from the samples by hand before the soil was finely ground with a mortar and pestle (< 2mm). All vessels and tools were cleaned between each sample to remove organic contaminants and prevent cross-contamination. This included washing the equipment with water and diluted soap, rinsing it three times each with tap water, reverse osmosis (RO) water and methanol then allowing it to completely dry under the fume cupboard.

Ground soil samples were acidified to remove inorganic carbon. Each sample was treated with 5% hydrochloric acid (HCl) and checked for bubbling from acidification of calcium carbonate. The samples were placed on a shaker table for an additional 30 minutes to ensure the soil and acid were mixed thoroughly. Once removed, the caps were loosened to release any pressure build up within the tubes containing the soil-acid mixture and the samples were left in the fume cupboard for 24 hours. The samples were then centrifuged in an Eppendorf Centrifuge 5810 for 7 minutes at 3500 rpm. The acid was carefully decanted. To remove any remaining acid, RO water was added, the samples were agitated and then centrifuged, and the supernatant was decanted, with this step repeated three times. The samples were then frozen and dried in the freeze drier for 48 hours. See Appendix B for a step by step method.

3.3 Bulk soil organic matter $\delta^{13}C$ analysis

The samples were re-ground to ensure sufficient mixing and weighed into tin boats and analysed for bulk carbon isotope ratios using a EuroVector Elemental Analyser – coupled to a Nu Horizon Isotope Ratio Mass Spectrometer (EA-IRMS). Samples were analysed for bulk carbon isotope ratios by Tony Hall at the University of Adelaide. Stable isotope ratios were expressed in δ notation as deviations from standards in parts per mil (‰) using Equation 1. $\delta^{13}C$ was reported relative to Vienna Pee Dee Belemite carbonate (VPDB). Glycine, glutamic acid and TPA were used as internal standards in this analysis, which were calibrated against international standards with a standard deviation of better than 0.1‰, based on 19 glycine, 7 glutamic acid and 3 TPA replicates.

3.4 Long chain *n*-alkane analysis

In this study leaf wax *n*-alkanes with a chain length of C_{31} were used as it is the most representative chain length, being produced more evenly between grasses and trees, than C_{29} and C_{33} (Uno et al., 2016). Lipid extraction of the soil samples was performed in a previous study conducted by Howard et al. (2018), using a Thermo Scientific Dionex Accelerate Solvent Extractor 350 using dichloromethane:methonal (DCM:MeOH; 9:1; v:v). Hexane was used to separate the non-polar, aliphatic hydrocarbon fraction from the total lipid extract, which was then concentrated under nitrogen gas.

The carbon isotope ratio of the *n*-alkanes was analysed using the Thermo Scientific Delta V Plus coupled to a Trace GC Ultra by Andrew Masteron at Northwestern

University. Sample values were standardised using *n*-alkane mixture A, (Arnnt Schimmelmann, University of Indianna) and are reported relative to VPDB.

3.5 Proportional C_4 cover

Point-intercept data collected by TERN vegetation surveys was acquired from the TERN plot database. Point-intercept data were analysed in the R statistical environment (R Core Team, 2019) and imported into R using the 'ausplotsR' package (Guerin, Saleeba, & Tokmakoff, 2019). The code utilised in this study to assign photosynthetic pathways and determine the proportional C_4 cover was developed and provided by Samantha Munroe (Appendix C).

Species ground cover (%) was calculated at each TERN plot using the *species_table* function. Species were then assigned a photosynthetic pathway using a TERN plant species photosynthetic database (Munroe et al., in prep). Proportional C_4 cover was then calculated at each plot as a proportion of C_3 and C_4 species cover by:

$$Proportional\ C_4\ cover = \frac{C_4\ species\ cover}{(C_4\ species\ cover + C_3\ species\ cover)} \quad (2)$$

Bulk SOM and *n*-alkane $\delta^{13}C$ values were converted to % C_4 using a two-end member mixing model in Equation 3 (Phillips, 2012).

$$\%\ C_4 = 1 - \frac{(\delta^{13}C_{sample} - \delta^{13}C_4)}{(\delta^{13}C_3 - \delta^{13}C_4)} \quad (3)$$

Where $\delta^{13}C_{sample}$ is the carbon isotopic composition derived from the SOM or *n*-alkanes, and $\delta^{13}C_4$ and $\delta^{13}C_3$ are the end members for C_4 and C_3 plants, respectively. For bulk SOM, the end members were -11.5‰ for C_4 and -27.5‰ for C_3. These values were adjusted from plant $\delta^{13}C$ values from Bird and Pousai (1997) to account for the effects of decomposition. Bird and Pousai (1997) used a 1‰ correction, however, the soil for this study was sampled from the top 3cm so a 0.5‰ correction was applied instead. The *n*-alkane % C_4 were calculated with end members of -17.8‰ for C_4 and -35.3‰ for C_3, from Garcin et al. (2014).

The proportional C_4 values derived from bulk SOM, *n*-alkanes and the vegetation survey were compared using zero-inflated beta regression models (a type of generalised linear model or glm), which were constructed using the *gamlss* function in R (Stasinopoulos et al., 2020). A zero-inflated beta regression was the most appropriate option to analyse these data because its error structure is designed to accommodate proportional or percentage data that ranges from 0 to 1 and includes values of 0. The zero-inflated beta regression model consist of two sub-models, a beta regression model (μ), which is a linear model that evaluates the relationship between variables between values of (0, 1), and a binomial model (ν), which models the probability of zero (i.e. 0% C_4 cover). The sub-models are then combined to estimate % C_4 cover over all possible values of the independent variable. Zero-inflated beta regression models were compared using a pseudo-R^2 (Nagelkerke R^2), which was used to measure the explained variance or "goodness-of-fit" of each model. Nagelkerke R^2-values range from 0 to 1, where values closer to 1 indicate a stronger correlation between two variables. A Spearman's rank correlation was used to compare SOM and *n*-alkane % C_4 values. A simplified

correlation was used to compare soil isotope metrics because there is no causal (i.e. cause and effect) relationship between these variables.

Additional data such as the percentage of grass, tree and chenopod vegetation cover, mean annual temperature (MAT) (Figure 4) and C_4 growing seasonal water availability (SWA) (Figure 5) were also extracted for each plot. Climate data were based on the 1970 to 2018 records from the Australian Gridded Climate Data set (Australian Government Bureau of Meteorology [BoM], 2018, accessed through the TERN Data Discovery Portal). Seasonal water availability was calculated according to Murphy and Bowman (2007) and is defined as the proportion of annual precipitation that occurs during the C_4 growing season, which is constrained by temperature. C_3 growth was considered possible in any month when the mean daily minimum temperature was ≥ -1 ºC, and the mean daily maximum temperature was ≥ 10 ºC and < 24 ºC. C_4 growth was considered possible in months where the mean daily maximum was ≥ 21 ºC and less than the temperature defined by:

$$T\ (℃) = \max\ (27, (1.745\ T_{annual} + 11.143)) \quad (4)$$

Where max (x, y) is the maximum of terms x and y, and T_{annual} is the mean annual temperature in degrees Celsius (°C). Seasonal water availability was then calculated as:

$$SWA = \frac{Precipitation\ in\ C_4\ growing\ season}{Precipitation\ in\ C_4\ growing\ season + precipitation\ in\ C_3\ growing\ season} \quad (5)$$

The correlations between latitude, the percentage of grass, tree and chenopod cover, MAT, SWA and proportional C_4 values were visualised and assessed using zero-inflated beta regression models as described above.

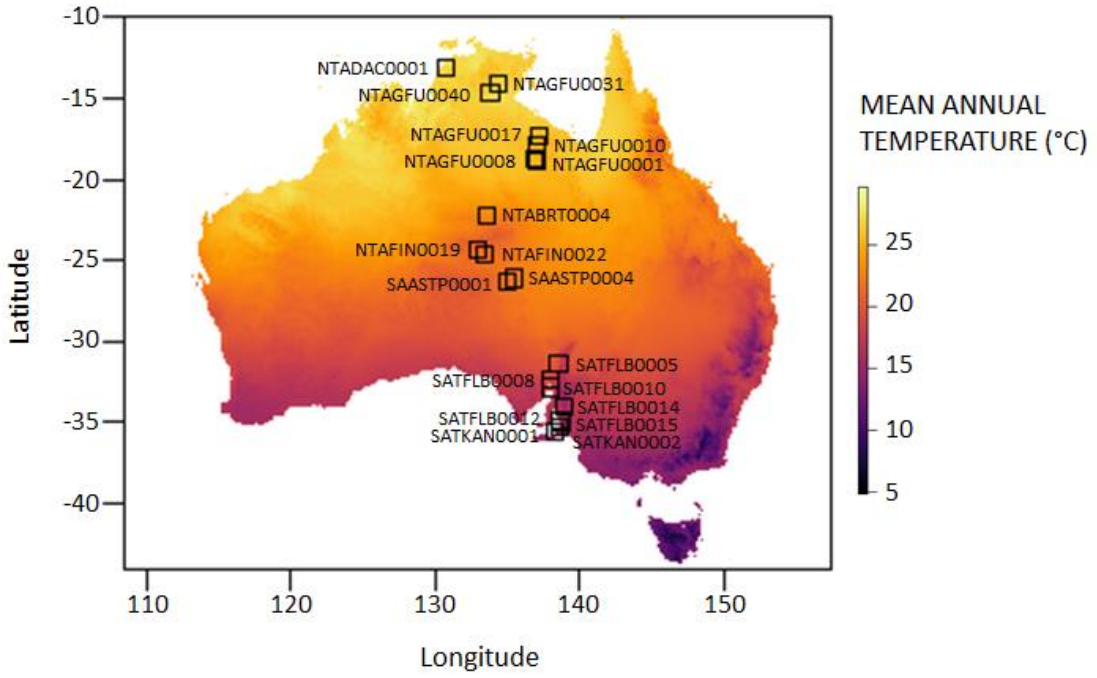

Figure 4. Location map of the 20 TERN plots across central Australia with mean annual temperature (MAT) as context. Data was averaged across the 1970 to 2018 records from the Australian Gridded Climate Data set from the Australian Government Bureau of Meteorology (2018).

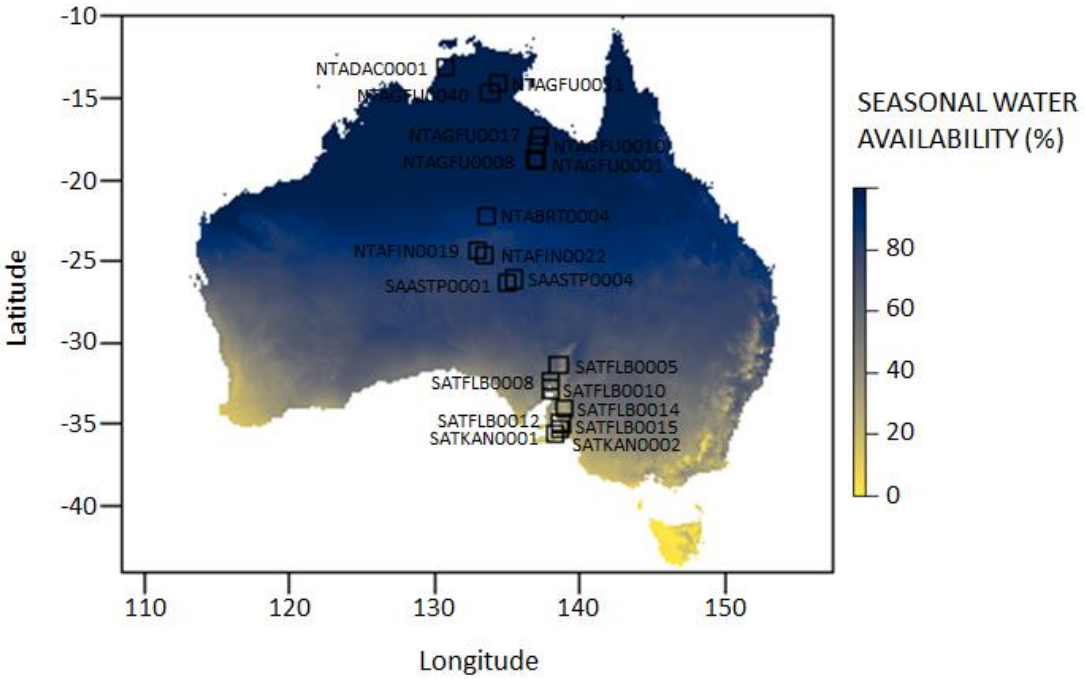

Figure 5. Location map of the 20 TERN plots across central Australia with seasonal water availability (SWA) as context. Data was averaged across the 1970 to 2018 records from the Australian Gridded Climate Data set from the Australian Government Bureau of Meteorology (2018).

4. OBSERVATIONS AND RESULTS

4.1 Bulk and compound specific carbon isotopic composition

$\delta^{13}C$ values for each run were output as values with units of per mil (‰) (Appendix D). SOM $\delta^{13}C$ values for each plot ranged from -15.5‰ to -28.0‰ with a mean of -23.2‰ and a standard deviation of 3.5‰, while *n*-alkane $\delta^{13}C$ values ranged from -22.3‰ to -35.3‰ with a mean of -31.0‰ and a standard deviation of 3.5‰ (Table 1). The $\delta^{13}C$ values of SOM and *n*-alkanes were visualised as a scatter graph with the C_3 and C_4 end members. With one plot as an exception, most of the data fell between the chosen end members (Figure 6). The $\delta^{13}C$ values were converted to proportional C_4 cover using a simple mixing model (Equation 3). The % C_4 derived from SOM $\delta^{13}C$ values ranged from 0.0% to 75.0%. The % C_4 derived from *n*-alkane $\delta^{13}C$ ranged from 0.0% to 74.3%

(Table 1). 84% of the plots had estimates from SOM and *n*-alkanes that were within 20% of one another, with the mean difference between *n*-alkane and SOM-derived % C_4 being negligible (-2.6%). However, in 3 of the 19 plots, the difference in predicted estimates of % C_4 was >20% (Figure 7). The average difference between the % C_4 derived from the vegetation survey and SOM and between the vegetation survey and *n*-alkanes for all 19 plots was 5.2% and 7.8%, respectively. This indicates that on average the proportional C_4 cover was higher for the vegetation survey than SOM and *n*-alkanes. The *n*-alkane $\delta^{13}C$ values for plot NTAGFU0031 suggested contamination from a petroleum source so it has been excluded from the subsequent analysis. However, the $\delta^{13}C$ values and % C_4 derived from bulk SOM for plot NTAGFU0031 has been reported in Table 1.

Table 1. Vegetation survey proportional C_4 cover (% C_4) was obtained from the Terrestrial Ecosystem Research Network (TERN) database. Carbon isotopic composition ($\delta^{13}C$) and % C_4 derived from bulk soil organic matter (SOM) and *n*-alkanes for each plot.

Plot name	Vegetation survey % C_4	Average bulk SOM $\delta^{13}C$ (‰)	SOM % C_4	Average *n*-alkane $\delta^{13}C$ (‰)	*n*-Alkane % C_4
NTABRT0004	37.1	-21.6	36.9	-32.2	17.7
NTADAC0001	65.2	-26.7	5.0	-32.1	18.3
NTAFIN0019	70.4	-20.3	45.0	-30.6	26.9
NTAFIN0022	27.8	-23.1	27.5	-32.0	18.9
NTAGFU0001	76.0	-18.0	59.4	-26.6	49.7
NTAGFU0008	92.7	-15.5	75.0	-22.8	71.4
NTAGFU0010	68.1	-22.8	29.4	-22.3	74.3
NTAGFU0017	36.7	-21.0	40.6	-32.3	17.1
NTAGFU0031	45.8	-24.3	20.0	Contaminated	
NTAGFU0040	31.8	-22.0	34.4	-29.7	32.0
SAASTP0001	35.2	-23.1	27.5	-31.4	22.3
SAASTP0004	25.7	-17.9	60.0	-30.1	29.7
SATFLB0005	0.0	-25.8	10.6	-33.0	13.1
SATFLB0008	47.8	-23.3	26.3	-29.9	30.9
SATFLB0010	0.0	-24.0	21.9	-33.3	11.4
SATFLB0012	0.0	-27.0	3.1	-34.0	7.4
SATFLB0014	0.0	-25.5	12.5	-32.5	16.0
SATFLB0015	0.0	-28.0	0.0	-35.3	0.0
SATKAN0001	0.0	-27.4	0.6	-34.7	3.4
SATKAN0002	0.0	-27.5	0.0	-34.2	6.3

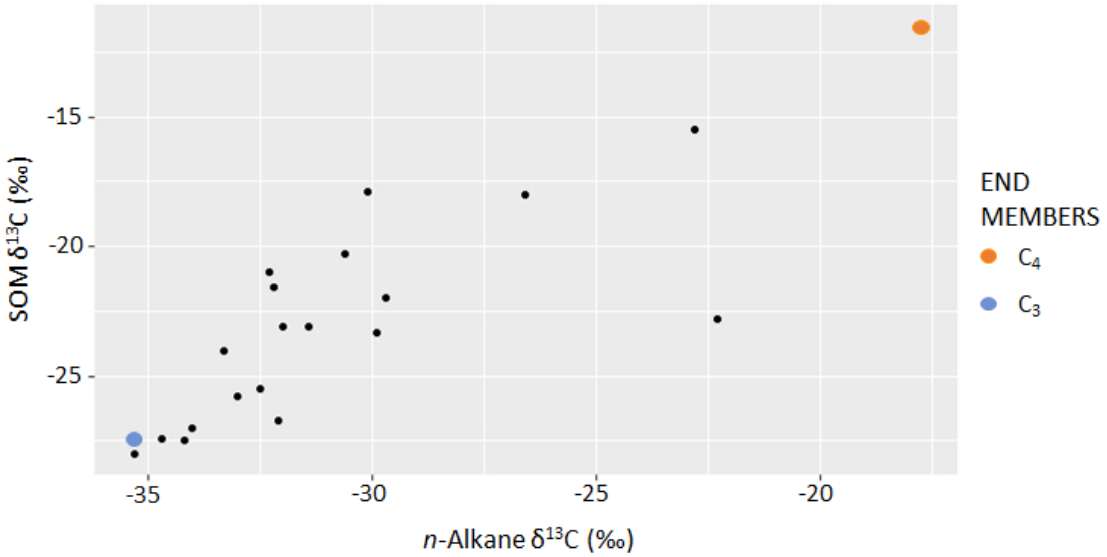

Figure 6. The carbon isotopic composition ($\delta^{13}C$) of *n*-alkanes compared to the carbon isotopic composition of soil organic matter (SOM). The C_3 and C_4 end members for SOM and *n*-alkanes were also included to show the portion of $\delta^{13}C$ data that fell within the end members.

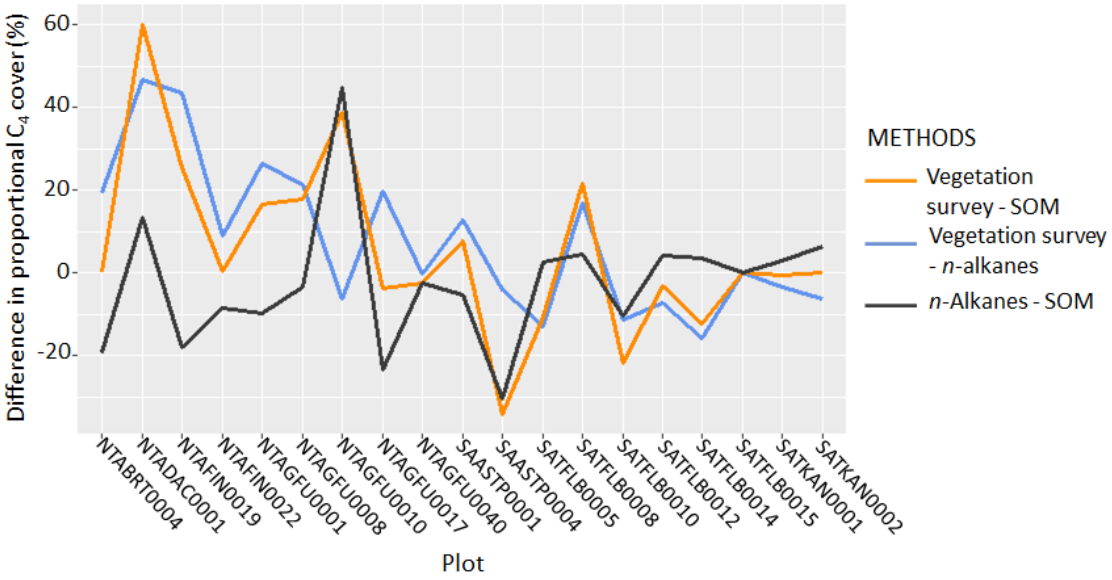

Figure 7. Difference between the proportional C_4 cover (% C_4) derived from the vegetation survey, soil organic matter (SOM) and *n*-alkanes compared to one another for each plot.

A positive correlation was observed between SOM-derived, *n*-alkane-derived and vegetation survey-derived % C_4 (Figure 8). However, the C_4 estimates produced using *n*-alkanes and SOM vary at low and high % C_4 values derived from the vegetation survey. When the vegetation survey recorded proportional C_4 cover of 0%, indicating that a site is purely dominated by C_3 plants, proportional C_4 cover estimates derived from SOM and *n*-alkanes were as high as 25% C_4. When the vegetation survey recorded high proportional C_4 cover, proportional C_4 cover estimates derived from SOM and *n*-alkanes were only 75%. The *n*-alkane % C_4 exhibited a stronger correlation with the vegetation survey, indicated by the pseudo-R^2 value of 0.67. This value is higher than the pseudo-R^2 value of 0.49 for SOM % C_4. Covariance can also be noted between *n*-alkane and SOM-derived % C_4 in Figure 9. The high correlation between the two methods is represented as a Spearman's rank correlation of 0.79 ($p<0.001$).

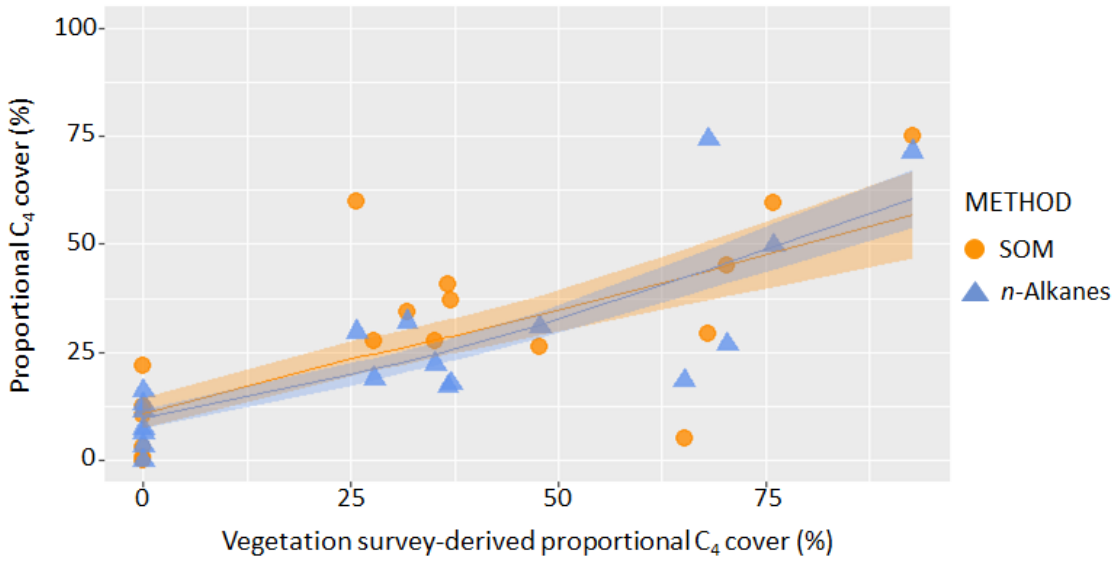

Figure 8. Proportional C_4 cover (% C_4) derived from bulk soil organic matter (SOM) and *n*-alkane carbon isotopic ($\delta^{13}C$) mixing models for each plot compared to % C_4 acquired from plot vegetation surveys. Solid lines are the predicted outputs of the zero-inflated beta regression models and the shaded bands are the standard error of the model.

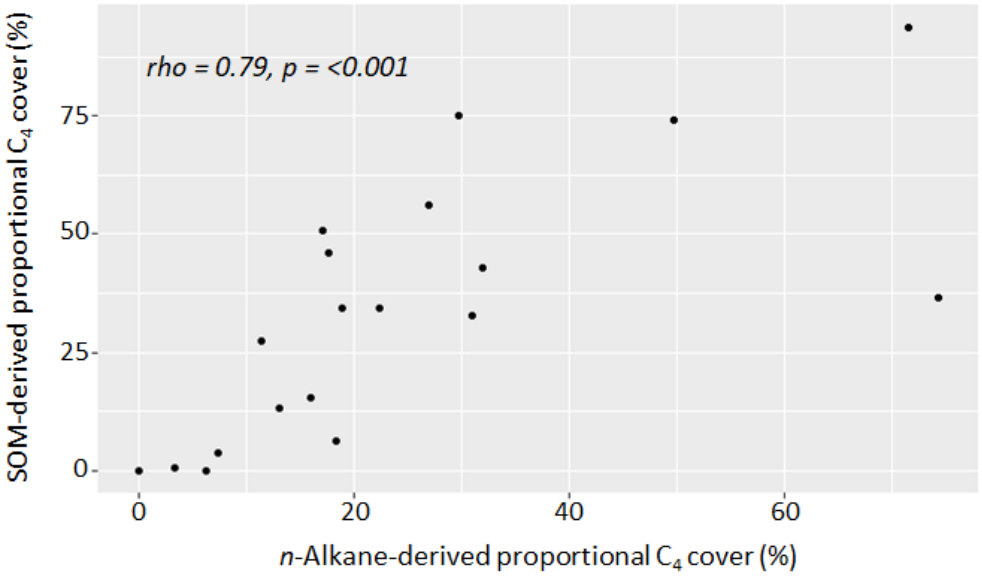

Figure 9. Proportional C_4 cover (% C_4) derived from the *n*-alkane carbon isotopic ($\delta^{13}C$) mixing model for each plot compared to % C_4 derived from the soil organic matter (SOM) $\delta^{13}C$ mixing model for each plot. A Spearman's rank correlation (rho (p-value)) was also included.

4.2 Proportional grass, tree and chenopod cover

A positive correlation was observed between estimated % C_4 from the vegetation survey, bulk SOM, and *n*-alkanes and proportional grass cover (Figure 10a). In contrast, as SOM and *n*-alkanes, and vegetation survey % C_4 increases, the proportion of tree cover decreases. The vegetation survey produces a classic binomial "S" shape due to the greater number of zeros within the data, compared to the SOM and *n*-alkanes. It can be noted from Figure 10a and b that % C_4 is better correlated with the proportion of grass cover than it is to tree cover for the vegetation survey and *n*-alkanes, but not for SOM. This correlation is reflected in the pseudo-R^2 values (Table 2). As expected, sites that are C_4 dominated have a greater proportion of grass cover compared to C_3 dominated sites that have more tree cover. No relationship is observed between % C_4 and the proportion of chenopod cover (Table 2) (Figure 10c). While model visualisation and R^2 values indicated a trend, p-values in some cases may not have been significant for some sub-models due to sample size.

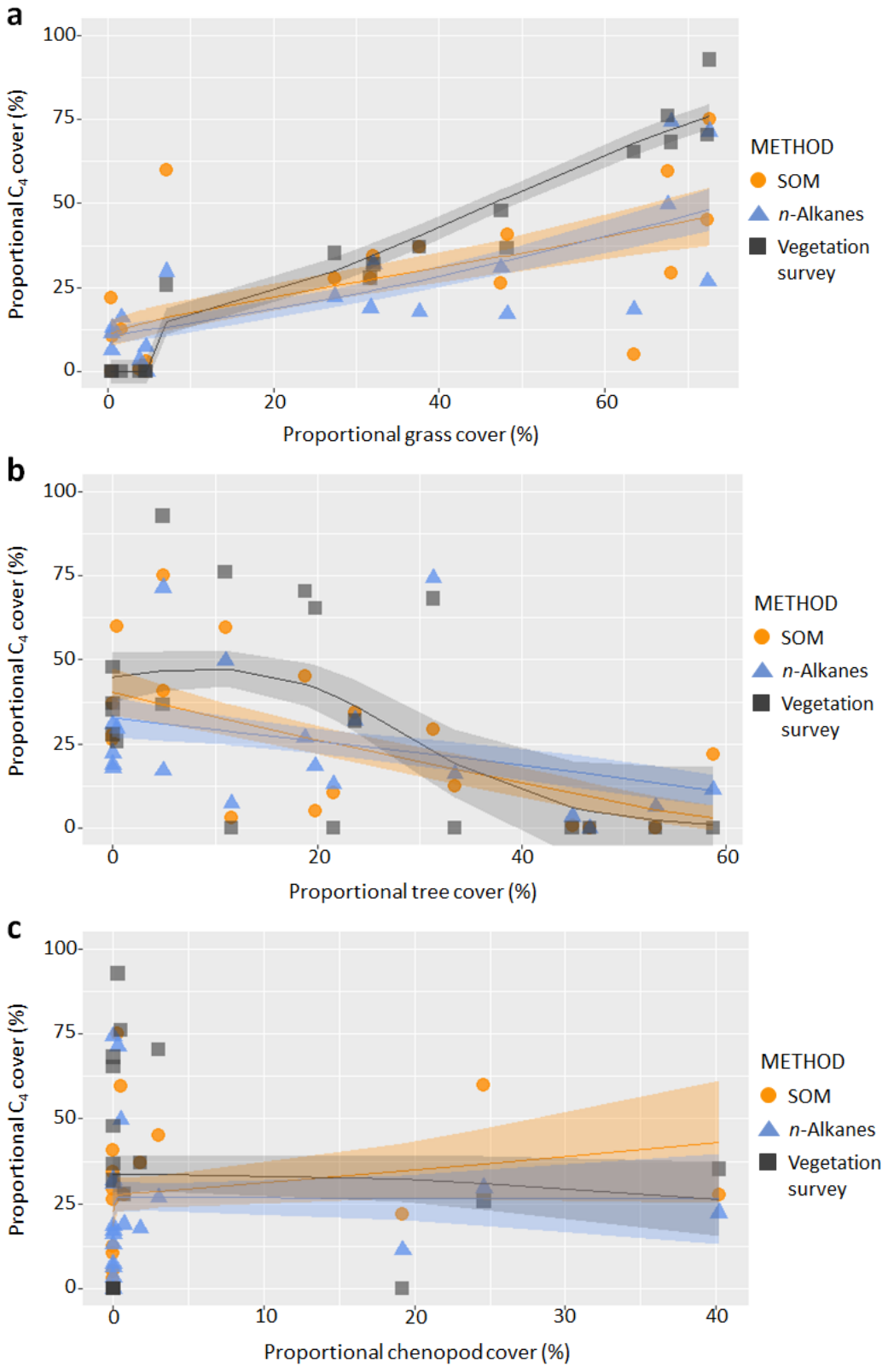

Figure 10. Proportional C_4 cover (% C_4) derived from the vegetation survey and % C_4 derived from bulk soil organic matter (SOM) and *n*-alkane carbon isotopic ($\delta^{13}C$) mixing models for each plot plotted by (a) grass cover, (b) tree cover and (c) chenopod cover, derived from the vegetation survey data. Solid lines are the predicted outputs of the zero-inflated beta regression models and the shaded bands are the standard error of the model.

Table 2. Results of zero-inflated beta regression analysis of the proportional C_4 cover (% C_4) derived from soil organic matter (SOM), *n*-alkanes and the vegetation survey versus proportional grass, tree and chenopod cover derived from the vegetation survey data. μ and ν beta coefficients are provided for each model, as well as p-values which denote the statistical significance of the predictor for each sub-model (pseudo-R^2 values are also provided for each model).

Proportional cover		SOM % C_4	*n*-Alkane % C_4	Vegetation survey % C_4
Grass	R^2	0.39	0.54	0.90
	μ coefficient (p-value)	1.99(0.023)	2.60(<0.001)	4.41(<0.001)
	ν coefficient (p-value)	-15.44(0.491)	-7.89(0.508)	-956.25(0.959)
Tree	R^2	0.43	0.20	0.51
	μ coefficient (p-value)	-3.15(0.044)	-1.58(0.154)	2.63(0.259)
	ν coefficient (p-value)	12.70(0.146)	8.40(0.287)	13.11(0.040)
Chenopod	R^2	0.15	0.06	0.11
	μ coefficient (p-value)	1.72(0.403)	-0.07(0.969)	-2.61(0.205)
	ν coefficient (p-value)	-3028.20(0.972)	-2790.45(0.967)	-3.28(0.556)

4.3 Correlations between proportional C_4 cover and climate

The % C_4 derived from the vegetation survey, bulk SOM and *n*-alkanes were compared to MAT and SWA to examine the correlation between the three methods and climate variables. The MAT and SWA for each plot can be found in Appendix E. Both MAT and SWA show a positive relationship with % C_4 derived from all three techniques. As temperature increases, the proportion of C_4 cover increases (Figure 11). This trend is also true for increasing SWA (Figure 12). As previously mentioned, the vegetation survey produces a classic binomial "S" shape due to the greater number of zeros within the data. Proportional C_4 cover is correlated strongly with both MAT and SWA (Table 3).

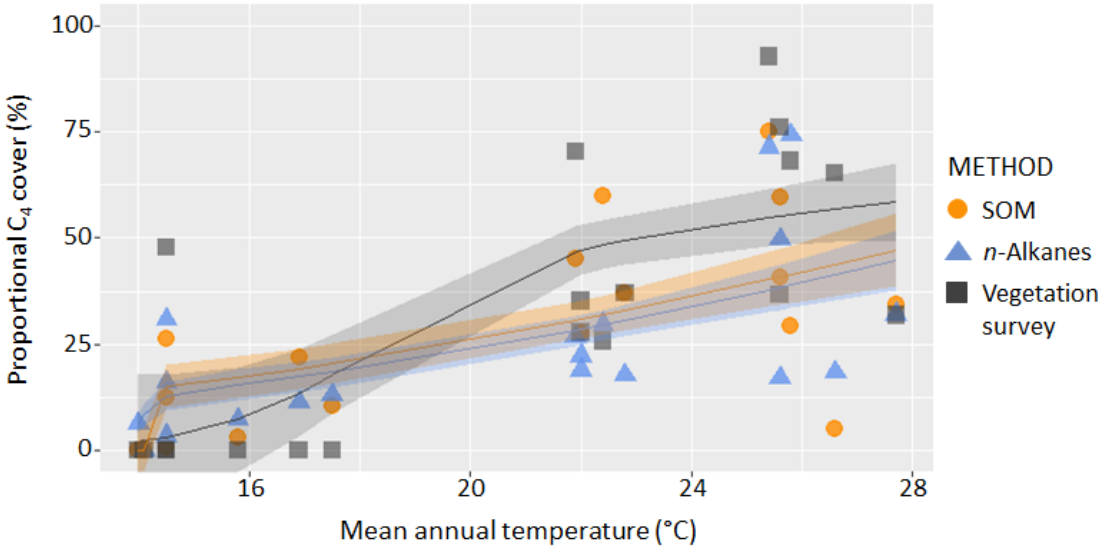

Figure 11. Proportional C_4 cover (% C_4) derived from the vegetation survey and % C_4 derived from bulk soil organic matter (SOM) and *n*-alkane carbon isotopic ($\delta^{13}C$) mixing models for each plot plotted versus mean annual temperature (MAT) averaged across the 1970 to 2018 records from the Australian Gridded Climate Data set from the Australian Government Bureau of Meteorology (2018). Solid lines are the predicted outputs of the zero-inflated beta regression models and the shaded bands are the standard error of the model.

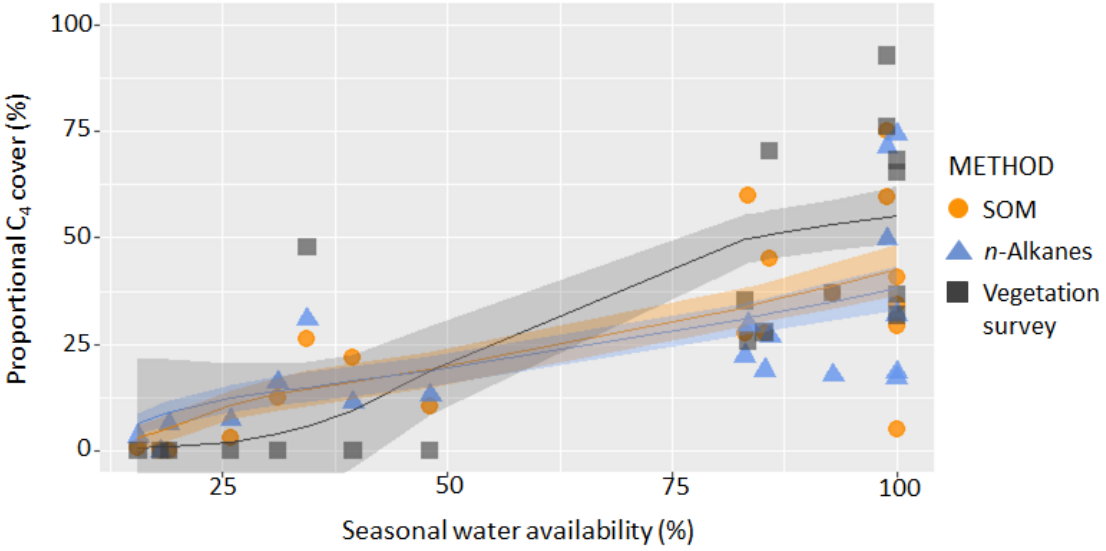

Figure 12. Proportional C_4 cover (% C_4) derived from the vegetation survey and % C_4 derived from bulk soil organic matter (SOM) and *n*-alkane carbon isotopic ($\delta^{13}C$) mixing models for each plot plotted versus seasonal water availability (SWA) averaged across the 1970 to 2018 records from the Australian Gridded Climate Data set from the Australian Government Bureau of Meteorology (2018). Solid lines are the predicted outputs of the zero-inflated beta regression models and the shaded bands are the standard error of the model.

Table 3. Results of zero-inflated beta regression analysis of the proportional C_4 cover (% C_4) derived from soil organic matter (SOM), *n*-alkanes and the vegetation survey versus mean annual temperature (MAT) and seasonal water availability (SWA) averaged across the 1970 to 2018 records from the Australian Gridded Climate Data set from the Australian Government Bureau of Meteorology (2018). μ and ν beta coefficients are provided for each model, as well as p-values which denote the statistical significance of the predictor for each sub-model (pseudo-R^2 values are also provided for each model).

Climate variables		SOM % C_4	*n*-Alkane % C_4	Vegetation survey % C_4
MAT	R^2	0.64	0.49	0.60
	μ coefficient (p-value)	0.12(0.017)	0.12(0.006)	0.06(0.422)
	ν coefficient (p-value)	-57.96(0.948)	-6.32(0.372)	-0.70(0.039)
SWA	R^2	0.59	0.50	0.64
	μ coefficient (p-value)	2.19(0.005)	1.91(0.005)	1.13(0.437)
	ν coefficient (p-value)	-25.47(0.257)	-24.12(0.447)	-11.42(0.068)

5. DISCUSSION

5.1 Comparing estimates of proportional C_4 cover relative to vegetation surveys

To better understand the influence of using different carbon isotopic tools on reconstructing the proportion of C_4 vegetation on a landscape, this study compared the proportional C_4 cover derived from bulk SOM and *n*-alkanes to vegetation surveys conducted for a number of plots across Australia. I hypothesised that SOM and *n*-alkane $\delta^{13}C$ values would accurately reflect broad spatial trends seen in the vegetation cover, which was supported by the results. Both SOM and *n*-alkane-derived % C_4 showed broadly similar trends to the vegetation survey. At a large scale, each method yielded similar geographical trends in proportional C_4 values. There was some evidence to suggest (i.e. the pseudo-R^2 values) that variations in the *n*-alkane-derived % C_4 were better explained by changes in the vegetation cover than SOM. This suggests there was a stronger relationship between *n*-alkanes and the vegetation survey than between SOM

and the vegetation survey. Previous studies recorded an inconsistent rate of decomposition amongst C_3 and C_4-derived carbon (Wynn & Bird, 2007), as well as an isotopic enrichment of ^{13}C due to microbial respiration in SOM (Krull & Bray, 2005; Wynn, Bird, & Wong, 2005). These processes influence the proportional C_4 cover estimated from SOM $\delta^{13}C$ values, which may account for the somewhat weaker relationship between SOM and vegetation survey-derived % C_4. However, it can be noted from the results that the proportional C_4 cover derived from SOM and *n*-alkanes covary, with 84% of the plots exhibiting similar % C_4. They also exhibited the same relationships with climate, where the % C_4 derived from all three methods increased with increasing MAT and SWA. Given the % C_4 models derived from each technique predicted similar % C_4 values across the transect, these results indicate that SOM and *n*-alkane $\delta^{13}C$ techniques can be used in place of one another to predict proportional C_4 cover at large scales. Furthermore, given that SOM is less expensive, and may be similarly effective as *n*-alkanes, it may be a better approach, in particular for large scale studies, as replicate samples can be taken to produce stronger relationships. However, this study only looked at surface soils (top 3cm) and as depth increases, SOM $\delta^{13}C$ values are more ^{13}C enriched due to decomposition (Wynn, Bird, & Wong, 2005). Therefore, determining end members and correctly calculating % C_4 becomes more difficult for older samples, such as 100 kyr, acquired from greater soil depths. Thus, *n*-alkanes may provide more accurate proportional C_4 cover for palaeoenvironmental reconstruction with older samples.

Despite following similar trends, *n*-alkanes and SOM % C_4 did not perfectly match the vegetation survey data. There are several potential explanations for these discrepancies.

The vegetation survey data is a direct measure of the above ground vegetation cover, whereas SOM and *n*-alkanes represent the vegetation biomass above and below ground. As a result, there may be a greater amount of root exudates that contribute to the carbon isotopic composition of SOM and *n*-alkanes which is not detected by the vegetation survey. Inputs from outside sources of *n*-alkanes is another potential explanation for different proportional C_4 cover derived from *n*-alkanes relative to the vegetation survey data. Aeolian transportation of leaf waxes from one site to another may give the $\delta^{13}C$ the appearance of being more or less C_4-derived (Schreuder et al., 2018). Data acquired from bulk SOM and *n*-alkanes is also time averaged, as it consists of old and new organic carbon inputs, unlike the vegetation survey which is a snapshot in time and can be seasonally biased as they are only conducted during a single season. Therefore, the vegetation survey represents more recent changes, which could include the impacts of fire activity and global warming, while SOM and *n*-alkanes represent the long-term history of the plot. Several plots in this study had experienced fire events over the past decade. It is possible that following an intense fire, secondary succession can occur, where pioneer species, such as C_4 grasses, grew after the plots were disturbed (Rego, Bunting, & da Silva, 1991). Therefore, the vegetation survey would reflect a high proportion of C_4 cover. The proportional C_4 cover detected by the vegetation survey may also be a result of global warming. The State of the Climate 2018 (CSIRO & Australian Government Bureau of Meteorology, 2018) report records a 1 ºC increase in temperature in Australia since 1910, leading to more frequent and extreme heat events. Such environmental conditions favour the expansion of C_4 plants (Edwards & Still, 2008).

5.2 Relationship between proportional C_4 cover and proportional grass, tree and chenopod cover

Proportional C_4 cover can be used to compare the geographical distribution of C_3 and C_4 plants at sites across the world. I hypothesised that high proportional C_4 cover would correlate with increased grass cover and low proportional C_4 cover would correlate with increased tree cover, which was supported by the results. Due to the climate preferences of C_3 and C_4 plants they are often geographically separated. As previously discussed, C_4 plants prefer warmer climates with summer rainfall, while C_3 plants are abundant in cool, wet environments (Murphy & Bowman, 2008; Still et al., 2003). Therefore, C_4 dominated communities are typically recorded in tropical and subtropical grasslands and savanna regions (Still et al., 2003), such as those located in northern Australia. C_3 plants are most diverse in the southern half of Australia (Hattersley, 1983). This can be observed in our results and is consistent with the findings of Hattersley (1983). Furthermore, Hattersley (1983) determined that 54% of Australia's native C_4 grasses can be found in the northern regions. This geographical distribution is also evident from our results, where proportional grass cover increased at plots in northern Australia, where proportional C_4 cover was high.

Chenopods, however, share no relationship with proportional C_4 cover. Chenopods also have a range of adaptations that allow them to exist in both arid and temperate environments, such as varying leaf structure (Shomer-Ilan, Nissenbaum, & Waisel, 1981), growth rate and the production of two types of seed, depending on the conditions (Shuyskaya et al., 2014). Therefore, the photosynthetic pathway of chenopods may not heavily influence their geographical distribution and they are not correlated with proportional C_4 cover (Munroe et al., in prep).

5.3 Relationship between proportional C_4 cover and climatic variables

Bulk SOM and *n*-alkane $\delta^{13}C$ values are also regularly used to determine how relative C_3 and C_4 vegetation cover is influenced by climate (Spicer, 1993). The relationship between C_4 plants and climatic variables are used to understand changes in proportional C_4 cover over space, in the future under climate change, and in palaeoclimate reconstruction (Still et al., 2003). While the intention of this study was not to determine how proportional C_4 cover was affected by climatic conditions, this work does demonstrate how each isotopic approach can produce C_4-climate relationships. I hypothesised that SOM, *n*-alkanes and the vegetation survey would produce different relationships between proportional C_4 cover and climatic variables. However, the results showed that all three methods shared similar trends with MAT and SWA and produced very similar C_4-climate relationships. This finding is consistent with previous environmental studies based in Australia and around the world (Andrae et al., 2018; Murphy & Bowman, 2007; Niu et al., 2005). Therefore, regardless of the method, the same general trend was acquired for both climatic variables. Thus, any of the three approaches can be used when conducting climate-based studies to determine the relationship between climatic variables and vegetation cover. However, as previously discussed, the age of the soil samples may impact the reliability of SOM to determine the relationship between climate and vegetation cover (Wynn, Bird, & Wong, 2005). Therefore, for ancient soils, *n*-alkanes may be the better approach as they are resistant to diagenesis.

5.4 Implications for palaeoenvironmental reconstruction

In a general sense, SOM, *n*-alkanes and the vegetation survey exhibited similar trends in proportional C_4 cover and produced similar relationships with climatic variables. The

similarity observed for all three methods provides confirmation that any approach can be used to estimate the proportional C_4 cover over large scales. SOM and *n*-alkane-derived % C_4 were more comparable to one another than with the vegetation survey. However, *n*-alkanes had a slightly stronger relationship with the vegetation survey than SOM. Additionally, SOM is more severely affected by decomposition with increasing depth, which could impact the ability of SOM to accurately reflect proportional C_4 cover. Therefore, when reconstructing palaeovegetation cover with ancient soil samples, *n*-alkane $\delta^{13}C$ values are likely a better isotopic tool for estimating proportional C_4 cover. The results of this study show the importance of comparative studies when considering the application of carbon isotopic tools for palaeoenvironmental reconstruction.

6. CONCLUSIONS

Carbon isotopic tools, such as bulk SOM and leaf wax *n*-alkanes are used to determine the photosynthetic pathway of plants for palaeoenvironmental reconstruction. This study utilised carbon isotopic composition values from SOM, *n*-alkanes and vegetation surveys to determine the ability of each method to estimate proportional C_4 cover and to test whether the three approaches were interchangeable. Bulk SOM, *n*-alkanes and the vegetation survey recorded similar trends over a large scale in the proportional C_4 cover through central Australia. All three methods also showed correlation between proportional C_4 cover and grass and tree cover, and climate variables, such as MAT and SWA. The similarity of the proportional C_4 cover yielded by these three methods provides flexibility when considering the preferred isotopic approach for the reconstruction of vegetation cover of regional areas. However, *n*-alkanes have a slightly

stronger correlation with the vegetation cover, compared to SOM. Furthermore, due to the increased severity of decomposition at greater depths within the soil, the reliability of SOM may be reduced when working with ancient samples. Therefore, for palaeoenvironmental reconstruction *n*-alkanes may be the more suitable approach for estimating proportional C_4 cover.

ACKNOWLEDGMENTS

I would like to express my gratitude to Francesca McInerney, Samantha Munroe and Ben Sparrow, my supervisor and co-supervisors, respectively, for this Honours thesis, and say thank you for their continuous support and sharing of their knowledge of isotopes, statistical analysis and palaeoenvironmental reconstruction. I would like to thank the Terrestrial Ecosystem Research Network for financial support, as well as supplying the samples and vegetation cover data used in this study. I would also like to thank Emrys Leitch, Michael Starkey, Nikki Francis and Lachlan Pink for taking me out into the field. I would like to express my appreciation for the organisation of this honours year to Derrick Hasterok and Katie Howard. I wish to acknowledge the assistance provided by Tony Hall and Andrew Masteron for the isotopic analysis of samples. I would like to offer my special thanks to Jonathan Smart for helping with the R coding. I would also like to extend my thanks to Christina Macdonald, Rob Klaebe and Sian Howard for their assistance in the lab. Finally, I wish to thank my family and partner, Zachary Tregenza, for their support and encouragement throughout this study.

REFERENCES

AKITA, S., & MOSS, D. N. (1972). Differential stomatal repsonse between C_3 and C_4 species to atmospheric CO_2 concentration and light. *Crop Science, 12*(6), 789-793. https://doi.org/DOI:10.2135/cropsci1972.0011183X001200060022x

AMBROSE, S. H., & SIKES, N. E. (1991). Soil carbon isotope evidence for Holocene habitat change in the Kenya Rift Valley. *Science, 253*(5026), 1402-1405. https://doi.org/DOI:10.1126/science.253.5026.1402

ANDRAE, J. W., MCINERNEY, F. A., POLISSAR, P. J., SNIDERMAN, J. M. K., HOWARD, S., HALL, P. A., & PHELPS, S. R. (2018). Initial expansion of C_4 vegetation in Australia during the Late Pliocene. *Geophysical Research Letters, 45*(10), 4831-4840. https://doi.org/DOI:10.1029/2018gl077833

ARAKAKI, M., CHRISTIN, P., NYFFELER, R., LENDEL, A., EGGLI, R., OGBURN, R. M., SPRIGGS, E., MOORE, M. J., & EDWARDS, E. J. (2011). Contemporaneous and recent radiations of the world's major succulent plant lineages. *Proceedings of the National Academy of Sciences, 108*(20), 8379-8384. https://doi.org/DOI:10.1073/pnas.1100628108

AUSTRALIAN GOVERNMENT BUREAU OF METEOROLOGY. (2018). *Gridded climate online*. http://www.bom.gov.au/climate/data-services/maps.shtml

BAUWE, H., HAGEMANN, M., & FERNIE, A. R. (2010). Photorespiration: Players, partners and origin. *Trends in Plant Science, 15*(6), 330-336. https://doi.org/DOI:10.1016/j.tplants.2010.03.006

BERNER, R. A., VANDENBROOKS, J. M., & WARD, P. D. (2007). Oxygen and evolution. *Science, 316*(5824), 557-558. https://doi.org/DOI:10.1126/science.1140273

BIRD, M., KRACHT, O., DERRIEN, D., & ZHOU, Y. (2003). The effect of soil texture and roots on the stable

carbon isotope composition of soil organic carbon. *Australian Journal of Soil Research, 41*(1), 77-94. https://doi.org/DOI:10.1071/SR02044

BIRD, M. I., GIRESSE, P., & CHIVAS, A. R. (1994). Effect of forest and savanna vegetation on the carbon-isotope composition of sediments from the Sanaga River, Cameroon. *Limnology and Oceanography, 39*(8), 1845-1854. https://doi.org/DOI:10.4319/lo.1994.39.8.1845

BIRD, M. I., & POUSAI, P. (1997). Variations of $\delta^{13}C$ in the surface soil organic carbon pool. *Global Biogeochemical Cycles, 11*(3), 313-322. https://doi.org/DOI:10.1029/97gb01197

BIRD, M. I., SUMMONS, R. E., GAGAN, M. K., ROKSANDIC, Z., DOWLING, L., HEAD, J., FIFIELD, L. K., CRESSWELL, R. G., & JOHNSON, D. P. (1995). Terrestrial vegetation change inferred from n-alkane $\delta^{13}C$ analysis in the marine environment. *Geochimica et Cosmochimica Acta, 59*(13), 2853-2857. https://doi.org/DOI:10.1016/0016-7037(95)00160-2

BOUCHENAK-KHELLADI, Y., VERBOOM, G. A., HODKINSON, T. R., SALAMIN, N., FRANCOIS§, O., NI CHONGHAILE, G., & SAVOLAINE, V. (2009). The origins and diversification of C_4 grasses and savanna-adapted ungulates. *Global Change Biology, 15*(10), 2397-2417. https://doi.org/DOI:10.11111/j.1365-2486.2009.01860.x

BRÄUTIGAM, A., & GOWIK, U. (2016). Photorespiration connects C_3 and C_4 photosynthesis. *Journal of Experimental Botany, 67*(10), 2953-62. https://doi.org/DOI:10.1093/jxb/erw056

BUSH, R. T., & MCINERNEY, F. A. (2015). Influence of temperature and C_4 abundance on *n*-alkane chain length distributions across the central USA. *Organic Geochemistry, 79*, 65-73. https://doi.org/DOI:10.1016/j.orggeochem.2014.12.003

CARDINALE, B. J., MATULICH, K. L., HOOPER, D. U., BYRNES, J. E., DUFFY, E., GAMFELDT, L., BALVANERA, P., O'CONNER, M. I., & GONZALEZ, A. (2011). The functional role of producer diversity in ecosystems. *American Journal of Botany, 98*(3), 572-592. https://doi.org/DOI:10.3732/ajb.1000364

CHAPIN, F. S., RINCON, E., & HUANTE, P. (1993). Environmental responses of plants and ecosystems as predictors of the impact of global change. *Journal of Biosciences, 18*(4), 515-524. https://doi.org/DOI:10.1007/DF02703083

COLLATZ, G. J., BERRY, J. A., & CLARK, J. S. (1998). Effects of climate and atmospheric CO_2 partial pressure on the global distribution of C_4 grasses: Present, past, and future. *Oecologia, 114*(4), 441-454. https://doi.org/DOI:10.1007/s004420050468

CSIRO., & AUSTRALIAN GOVERNMENT BUREAU OF METEOROLOGY. (2018). *State of the climate: Report at a glance*. http://www.bom.gov.au/state-of-the-climate/index.shtml

DAWSON ,T. E., MAMBELLI, S., PLAMBOECK, A. H., TEMPLER, P. H., & TU, K. P. (2002). Stable isotopes in plant ecology. *Annual Review of Ecology and Systematics, 33*(1), 507-559. https://doi.org/DOI:10.1146/annurev.ecolsys.33.020602.095451

DIEFENDORF, A. F., & FREIMUTH, E. J. (2017). Extracting the most from terrestrial plant-derived *n*-alkyl lipids and their carbon isotopes from the sedimentary record: A review. *Organic Geochemistry, 103*, 1-21. https://doi.org/DOI:10.1016/j.orggeochem.2016.10.016

EDWARDS, E. J., & STILL, C. J. (2008). Climate, phylogeny and the ecological distribution of C_4 grasses. *Ecology Letters, 11*(3), 266-276. https://doi.org/DOI:10.1111/j.1461-0248.2007.01144.x

EHLERINGER, J. R. (1978). Implications of quantum yield differences on the distributions of C_3 and C_4 grasses. *Oecologia, 31*(3), 255-267. https://doi.org/DOI:10.1007/BF00346246

EHLERINGER, J. R., CERLING, T. E., & HELLIKER, B. R. (1997). C_4 photosynthesis, atmospheric CO_2, and climate. *Oecologia, 112*(3), 285-299. https://doi.org/DOI:10.1007/s004420050311

FARQUHAR, G. D., BUCKLEY, T. N., & MILLER, J. M. (2002). Optimal stomatal control in relation to leaf area and nitrogen content. *Silva Fennica, 36*(3), 625-637. https://doi.org/DOI:10.14214/sf.530

FARQUHAR, G. D., EHLERINGER, J. R., & HUBICK, K. T. (1989). Carbon isotope discrimination and photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology, 40*, 503-537. https://doi.org/DOI:10.1146/annurev.pp.40.060189.002443

GAMARRA, B., & KAHMEN, A. (2015). Concentrations and δ^2H values of cuticular *n*-alkanes vary significantly among plant organs, species and habitats in grasses from an alpine and a temperate European grassland. *Oecologia, 178*(4), 981-98. https://doi.org/DOI:10.1007/s00442-015-3278-6

GARCIN, Y., SCHEFUß, E., SCHWAB, V. F., GARRETA, V., GLEIXNER, G., VINCENS, A., TODOU, G., SENE, O., ONANA, J., ACHOUNDONG, G., & SACHSE, D. (2014). Reconstructing C_3 and C_4 vegetation cover using *n*-alkane carbon isotope ratios in recent lake sediments from Cameroon, Western Central Africa. *Geochimica et Cosmochimica Acta, 142*, 482-500. https://doi.org/DOI:10.1016/j.gca.2014.07.004

GLASER, B. (2005). Compound-specific stable-isotope ($\delta^{13}C$) analysis in soil science. *Journal of Plant Nutrition and Soil Science, 168*(5), 633-648. https://doi.org/DOI:10.1002/jpln.200521794

GUERIN, G. R., SALEEBA, T., & TOKMAKOFF, A. (2019). *AusplotsR: TERN AusPlots analysis package* (Versin 1.0) [Computer software].

HARTMAN, G. (2011). Reconstructing Mid-Pleistocene paleovegetation and paleoclimate in the Golan Heights using the $\delta^{13}C$ values of modern vegetation and soil organic carbon of paleosols. *Journal of Human Evolution, 60*(4), 452-463. https://doi.org/DOI:10.1016/j.jhevol.2010.08.001

HATTERSLEY, P. W. (1983). The distribution of C_3 and C_4 grasses in Australia in relation to climate. *Oecologia, 57*(1), 113-128. https://doi.org/DOI:10.1007/BF00379569

HEMKEMEYER, M., DOHRMANN, A. B., CHRISTENSEN, B. T., & TEBBE, C. C. (2018). Bacterial preferences for specific soil particle size fractions revealed by community analyses. *Frontiers in Microbiology, 9*(149), 1-13. https://doi.org/DOI:10.3389/fmicb.2018.00149

HOWARD, S., MCINERNEY, F. A., CADDY-RETALIC, S., HALL, P. A., & ANDRAE, J. W. (2018). Modelling leaf wax *n*-alkane inputs to soils along a latitudinal transect across Australia. *Organic Geochemistry, 121*, 126-137. https://doi.org/DOI:10.1016/j.orggeochem.2018.03.013

KRULL, E. G., & BRAY, S. S. (2005). Assessment of vegetation change and landscape variability by using stable carbon isotopes of soil organic matter. *Australian Journal of Botany,* 53(7), 651-661. https://doi.org/DOI:10.1016/j.geoderma.2004.09.012

LI, R., FAN, J., XUE, J., & MEYERS, P. A. (2017). Effects of early diagenesis on molecular distributions and carbon isotopic compositions of leaf wax long chain biomarker *n*-alkanes: Comparison of two one-year-long burial experiments. *Organic Geochemistry, 104*, 8-18. https://doi.org/DOI:10.1016/j.orggeochem.2016.11.006

LUNDGREN, M. R., OSBORNE, C. P., & CHRISTIN, P. (2014). Deconstructing Kranz anatomy to understand C_4 evolution. *Journal of Experimental Botany, 65*(13), 3357-3369. https://doi.org/DOI:10.1093/jxb/eru186

MEYERS, P. A. (1994). Preservation of elemental and isotopic source identification of sedimentary organic matter. *Chemical Geology, 114*(3), 289-302. https://doi.org/DOI:10.1016/0009-2541(94)90059-0

MUNROE, S. E. M., MCINERNEY, F. A., ANDRAE, J., WELTI, N., GUERIN, G. R., ATKINS, R., & SPARROW, B. (in prep). Climate-driven trends in C_4 abundance differ between taxa.

MURPHY, B. P., & BOWMAN, D. M. J. S. (2007). Seasonal water availability predicts the relative abundance of C_3 and C_4 grasses in Australia. *Global Ecology and Biogeography, 16*(2), 160-169. https://doi.org/DOI:10.1111/j.1466-8238.2006.00285.x

NIU, S., YUAN, Z., ZHANG, Y., LIU, W., ZHANG, L., HUANG, J., & WAN, S. (2005). Photosynthetic responses of C_3 and C_4 species to seasonal water variability and competition. *Journal of Experimental Botany, 56*(421), 2867-2876. https://doi.org/DOI:10.1093/jxb/eri281

O'LEARY, M. H. (1988). Carbon isotopes in photosynthesis. *BioScience, 38*(5), 328-335. https://doi.org/DOI:10.2307/1310735

OSMOND, C. B. (1981). Photorespiration and photoinhibition: Some implications for the energetics of photosynthesis. *Biochimica et Biophysica Acta, 639*(2), 77-98. https://doi.org/DOI:10.1016/0304-4173(81)90006-9

PAUSCH, J., & KUZYAKOV, Y. (2017). Carbon input by roots into the soil: Quantification of rhizodeposition from root to ecosystem scale. *Global Change Biology, 24*(1), 1-12. https://doi.org/DOI:10.1111/gcb.13850

PETERHANSEL, C., HORST, I., NIESSEN, M., BLUME, C., KEBEISH, R., KURKCUOGLU, S., & KREUZALER, F. (2010). Photorespiration. In C. Somerville & E. Meyerowitz (Eds.), *Arabidopsis Book* (pp. 1-24). The American Society of Plant Biologists.

PHILLIPS, D. L. (2012). Converting isotope values to diet composition: The use of mixing models. *Journal of Mammalogy, 93*(2), 342-352. https://doi.org/DOI:10.1644/11-MAMM-S-158.1

R CORE TEAM. (2019). *R: A language and environment for statistical computing* (Version 3.6.1) [Computer Software]. The R Foundation for Statistical Computing.

RAO, Z., JIA, G., ZHU, Z., WU, Y., & ZHANG, J. (2008). Comparison of the carbon isotope composition of total organic carbon and long-chain *n*-alkanes from surface soils in eastern China and their significance. *Chinese Science Bulletin, 53*(24), 3921-3927. https://doi.org/DOI:10.1007/s11434-008-0296-3

REGO, F. C., BUNTING, S. C., & DA SILVA, J. M. (1991). Changes in understory vegetation following prescribed fire in maritime pine forests. *Forest Ecology and Management, 41*(1), 21-31. https://doi.org/DOI:10.1016/0378-1127(91)90117-E

SAGE, R. F., CHRISTIN, P., & EDWARDS, E. J. (2011). The C_4 plant lineages of planet Earth. *Journal of Experimental Botany, 62*(9), 3155-3169. https://doi.org/DOI:10.1093/jxb/err048

SAGE, R. F., & STATA, M. (2015). Photosynthetic diversity meets biodiversity: The C_4 plant example. *Journal of Plant Physiology, 172*, 104-119. https://doi.org/DOI:10.1016/j.jplph.2014.07.024

SAIZ, G., BIRD, M., WURSTER, C., QUESADA, C. A., ASCOUGH, P., DOMINGUES, T., SCHRODT, F., SCHWARZ, M., FELDPAUSCH, T. R., VEENENDAAL, E., DJAGBLETEY, G., JACOBSEN, G., HIEN, F., COMPAORE, H., DIALLO, A., & LLOYD, J. (2015). The influence of C_3 and C_4 vegetation on soil organic matter dynamics in contrasting semi-natural tropical ecosystems. *Biogeosciences, 12*(16), 5041-5059. https://doi.org/DOI:10.5194/bg-12-5041-2015

SALZMANN, U., HAYWOOD, A. M., & LUNT, D. J. (2009). The past is a guide to the future? Comparing Middle Pliocene vegetation with predicted biome distributions for the twenty-first century. *Philosophical Transactions, 367*(1886), 189-204. https://doi.org/DOI:10.1098/rsta.2008.0200

SCHAEDLE, M. (1975). Tree photosynthesis. *Annual Review of Plant Physiology, 26*(1), 101-115. https://doi.org/DOI:10.1146/annurev.pp.26.060175.000533

SCHREUDER, L. T., STUUT, J. W., KORTE, L. F., DAMSTE, J. S. S., & SCHOUTEN, S. (2018). Aeolian transport and deposition of plant wax *n*-alkanes across the tropical North Atlantic Ocean. *Organic Geochemistry, 115*, 113-123. https://doi.org/DOI:10.1016/j.orggeochem.2017.10.010

SEDELNIKOVA, O. V., HUGHES, T. E., & LANGDALE, J. A. (2018). Understanding the genetic basis of C_4 Kranz anatomy with a view to engineering C_3 crops. *Annual Review of Genetics, 52*, 249-270. https://doi.org/DOI:10.1146/annurev-genet-120417-031217

SHOMER-ILAN, A., NISSENBAUM, A., & WAISEL, Y. (1981). Photosynthetic pathways and the ecological distribution of the Chenopodiaceae in Israel. *Oecologia, 48*, 244-248. https://doi.org/DOI:10.1007/BF00347970

SHUYSKAYA, E. V., LI, E. V., RAHMANKULOVA, Z. F., KUZNETSOVA, N. A., TODERICH, K. N., & VORONIN, P. Y. (2014). Morphophysiological adaptation aspects of different *Haloxylon aphyllum* (Chenopodiaceae) genotypes along a salinity gradient. *Russian Journal of Ecology, 45*(3), 181-187. https://doi.org/DOI:10.1134/S1067413614030114

SIKES, N. E., & ASHLEY, G. M. (2007). Stable isotopes of pedogenic carbonates as indicators of paleoecology in the Plio-Pleistocene (upper Bed I), western margin of the Olduvai Basin, Tanzania. *Journal of Human Evolution, 53*(5), 574-594. https://doi.org/DOI:10.1016/j.jhevol.2006.12.008

SILVERA, K., NEUBIG, K. M., WHITTEN, W. M., WILLIAMS, N. H., WINTER, K., & CUSHMAN, J. C. (2010). Evolution along the crassulacean acid metabolism continuum. *Functional Plant Biology, 37*(11), 995-1010. https://doi.org/DOI:10.1016/j.ympev.2010.01.021

SPICER, R. A. (1993). Palaeoecoloy, past climate systems and C_3/C_4 photosynthesis. *Chemosphere, 27*(6), 947-978. https://doi.org/DOI:10.1016/0045-6535(93)90063-B

STADDON, P. L. (2004). Carbon isotopes in functional soil ecology. *Trends in Ecology and Evolution*, 19(3), 148-54. https://doi.org/DOI:10.1016/j.tree.2003.12.003

STASINOPOULOS, M., RIGBY, B., VOUDOURIS, V., AKANTZILIOTOU, C., ENEA, M., & KIOSE, D. (2020). *Generalised Additive Models for Location Scale and Shape* (Version 5.1-6) [Computer software].

STILL, C. J., BERRY, J. A., COLLATZ, G. J., & DEFRIES, R. S. (2003). Global distribution of C_3 and C_4 vegetation: Carbon cycle implications. *Global Biogeochemical Cycles, 17*(1), 1-14. https://doi.org/DOI:10.1029/2001GB001807

TEERI, J. A., & STOWE, L. G. (1976). Climatic patterns and the distribution of C_4 grasses in North America. *Oecologia, 23*(1), 1-12. https://doi.org/DOI:10.1007/BF00351210

UNO, K. T., POLISSAR, P. J., JACKSON, K. E., & DEMENOCAL, P. B. (2016). Neogene biomarker record of vegetation change in eastern Africa. *Proceedings of the National Academy of Science of the United States, 113*(23), 6355-6363. https://doi.org/DOI:10.1073/pnas.1521267113

VON FISCHER, J. C., TIESZEN, L. L., & SCHIMEL, D. S. (2008). Climate controls on C_3 vs. C_4 productivity in North American grasslands from carbon isotope composition of soil organic matter. *Global Change Biology, 14*(5), 1141-1155. https://doi.org/DOI:10.1111/j.1365-2486.2008.01552.x

WANG, G., FENG, X., HAN, J., ZHOU, L., TAN, W., & SU, F. (2008). Paleovegetation reconstruction using $\delta^{13}C$ of soil organic matter. *Biogeosciences, 5*(5), 1325-1337. https://doi.org/DOI:10.5194/bg-5-1325-2008

WEST, J. B., BOWEN, G. J., CERLING, T. E., & EHLERINGER, J. R. (2006). Stable isotopes as one of nature's ecological recorders. *Trends in Ecology and Evolution, 21*(7), 408-414. https://doi.org/DOI:10.1016/j.tree.2006.04.002

WHITE, A., SPARROW, B., LEITCH, E., FOULKES, J., FLITTON, R., LOWE, A. J., & CADDY-RETALIC, S. (2012). *AusPlots rangelands survey protocols manual*. The University of Adelaide Press.

WORLD WIDE FUND FOR NATURE. (2001). *Ecoregions.* https://www.worldwildlife.org/search?cx=003443374396369277624%3Av3nraqhmeyk&ie=UTF-8&x=australia+ecoregion

WYNN, J. G., & BIRD, M. I. (2007). C_4-derived soil organic carbon decomposes faster than its C_3 counterpart in mixed C_3/C_4 soils. *Global Change Biology, 13*(10), 2206-2217. https://doi.org/DOI:10.1111/j.1365-2486.2007.01435.x

WYNN, J. G., & BIRD, M. I. (2008). Environmental controls on the stable carbon isotopic composition of soil organic carbon: Implications for modelling the distribution of C_3 and C_4 plants, Australia. *Tellus B: Chemical and Physical Meteorology, 60*(4), 604-621. https://doi.org/DOI:10.1111/j.1600-0889.2008.00361.x

WYNN, J. G., BIRD, M. I., & WONG, V. N. L. (2005). Rayleigh distillation and the depth profile of $^{13}C/^{12}C$ ratios of soil organic carbon from soils of disparate texture in Iron Range National Park, Far North Queensland, Australia. *Geochimica et Cosmochimica Acta, 69*(8), 1961-1973. https://doi.org/DOI:10.1016/j.gca.2004.09.003

ZACHOS, J. C., DICKENS, G. R., & ZEEBE, R. E. (2008). An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature,* 451(7176), 279-283. https://doi.org/DOI:10.1038/nature06588

APPENDIX A. PLOT DESCRIPTIONS

Table A.1. Descriptions of the 20 plots accessed through the TERN database.

Plot name	Description
NTABRT0004	Grasslands dominate to the South and North of the plot. Grazing and weed impacts are minimal. Plot is long unburnt, but fire activity is present in the surrounding area.
NTADAC0001	Broad, slightly elevated plain with little bare soil extending to the West and gently sloping North-East in the East. Tree bark is charred, and charcoal is present on the ground, resembling burns of differing ages.
NTAFIN0019	Very homogenous vegetation with a high impact of introduced plant species, such as *Cenchrus ciliaris*. Fire has trickled into the South-East corner of the plot, however, the trees remain unmarred. The rest of the plot is long unburnt.
NTAFIN0022	Sandy intrusion in the South-East corner. Impact of introduced plant species is minimal.
NTAGFU0001	Flat plot located in a shallow drainage depression. Grazing impact is moderate, however, the grass layer is still intact. Weed impact is low. Plot is long unburnt due to the sparse vegetation cover.
NTAGFU0008	Gravelly plot with no real evidence of grazing or impact from weeds. Plot is long unburnt.
NTAGFU0010	Gravelly plot with moderate impact of grazing and minimal impact of introduced plant species. Fire activity within the last two to three years of sampling, having been evenly burnt throughout.
NTAGFU0017	Shallow valley that slopes up at the South-West and North-East corners. Impact of grazing and introduced plant species is low. Plot is long unburnt, however, *Melaleuca viridiflora* is much shorter in the southern section possibly due to previous fire activity.
NTAGFU0031	Flat plot sloping slightly to the North-East. Impact from grazing is low and the plot is long unburnt.
NTAGFU0040	Located in a valley between low hills. Impact of grazing is moderate, while impact of introduced plant species is minimal. Plot is long unburnt, evident from the unmarred trees.
SAASTP0001	Isolated channels to the South of the plot, with gravel strewn rills and minor drainage. Impact of introduced plants species is minimal.
SAASTP0004	Plot area leads into a dam. Cattle were also seen in small groups on site.
SATFLB0005	Located at the top of the range above Brachina Gorge on the cap of the slope dropping away to the North-East. The North-West corner drops off the edge of the ridge on the other side. Large amount of rock covered by cryptogam. Minimal bare ground and weeds. Plot is long unburnt.
SATFLB0008	Steep slope facing east to Dutchman's Stern Range with a drainage line running from the North-West to South-East. Impact of grazing and introduced plant species is low. *Spinifex* dominated. Plot is long unburnt.
SATFLB0010	Stony plot, with most stones covered in lichen. Impact of weeds is moderate. Plot is long unburnt.
SATFLB0012	Located on North-West facing side slopes of Black Hill. Impact of grazing and weeds is low. Plot is long unburnt.
SATFLB0014	Slope to the west of the plot, which becomes increasingly rocky towards the top. Impact from grazing and weeds is low. Plot is very long unburnt.
SATFLB0015	North to South slope in the centre of the plot falling to the North-West and North-East corners. Walking tracks run along the southern edge. Impact of weeds is high. Fire activity occurred in the last decade. Eucalypt trees have been scarred, however, the shrub layer has grown back.
SATKAN0001	Steeply sloping down to a valley in the North-East corner. Moderate impact of weeds. Plot is very long unburnt.
SATKAN0002	Gently sloping to the North-West. Shallow ephemeral creek is present in the North-West corner. Plot burnt within the last five years.

APPENDIX B. STEP BY STEP METHOD

B.1 Sample preparation

1. Mortar, pestle and tweezers were washed with diluted soapy water and rinsed three times with tap water, RO water and methanol. Excess methanol was drained into a beaker containing silica powder and equipment was left to dry in the fume cupboard.
2. Working surface was wiped down with methanol.
3. The soil sample was carefully emptied into the mortar and any organic material, such as twigs, seeds, leaves or fibrous roots were removed using tweezers and placed in a small vial. Rock fragments were also removed and discarded.
4. Soil was ground into a fine powder (< 2mm) with the pestle and placed back into the Eppendorf tube.
5. Steps 1 through 4 were repeated for all 20 soil samples.

B.2 Acidification

1. 50mL of hydrochloric acid was slowly incorporated into 900mL of RO water to make 1L of 5% hydrochloric acid.
2. The acid was added to the Eppendorf tubes (approximately halfway) of all 20 soil samples to remove any inorganic carbon.
3. Samples were shaken well to ensure the acid was thoroughly mixed with the soil and the caps were loosened to allow any pressure build up to be released.
4. An additional 40mL of acid was added to each sample, which were then placed on the shaker table for 30 minutes.
5. Samples were left for 24 hours.

B.3 Centrifuge and freeze drier

1. Samples were placed in the Centrifuge 5810 for seven minutes set at 3500 revolutions per minute (rpm).
2. Acid was carefully decanted into a waste container.
3. 40mL of RO water was added to the Eppendorf tubes of each sample, which were shaken vigorously to ensure the sediment was mixed well.
4. Samples were placed back onto the Centrifuge 5810 for seven minutes set at 3500 rpm.
5. RO water was carefully decanted.
6. Steps 3 through 5 were repeated two more times. The first RO water rinse was discarded in the acid waste container and the following two rinses were poured down the sink.
7. Samples were placed in the freezer and left for 24 hours.
8. Samples were placed in the freeze drier and left for 48 hours.

B.4 Sample weighing

1. Once removed from the freeze drier, the soil was re-ground with a mortar and pestle, ensuring equipment was thoroughly cleaned (as mentioned in the sample preparation) between each sample.
2. A small amount of soil was placed into a tin boat, that was previously tared on the scale.
3. The soil was weighed and adjusted accordingly depending on the amount needed for analysis. 30 – 40mg, 10 – 15mg and 3 – 4mg was weighed for red, yellow and dark brown soils, respectively.
4. Steps 1 through to 3 were repeated for all 20 soil samples.

APPENDIX C. RSTUDIO CODE

C_3 versus C_4 cover/abundance and distribution analysis
Developed by Samantha Munroe

```
#Libraries
library(ausplotsR)
library(rgbif)
library(data.table)
library(vegan)
library(betapart)
library(raster)
library(maptools)
library(plyr)
library(rgdal)
library(car)
library(dplyr)
library(tidyr)
library(reshape)
library(ggplot2)
library(mapproj)
library(viridis)
library(ALA4R)
library(maps)
library(mapdata)
library(ape)
library(phytools)
library(stringr)
library(forcats)
library(devtools)

#Cover calculations
#Step One: Subset and clean data
#Import the ausplots data from the 20 sample sites
setwd("D:/Honours/RStudio") #working directory
my.ausplots.data <- get_ausplots(my.Plot_IDs=c("NTADAC0001", "NTAGFU0031",
"NTAGFU0040", "NTAGFU0017", "NTAGFU0010", "NTAGFU0008",
"NTAGFU0001", "NTABRT0004", "NTAFIN0019", "NTAFIN0022",
"SAASTP0004", "SAASTP0001", "SATFLB0005", "SATFLB0008", "SATFLB0010",
"SATFLB0014", "SATFLB0012", "SATFLB0015", "SATKAN0002",
"SATKAN0001"), veg.vouchers=FALSE, site_info=T) #extracting sample sites
veg.PI<-my.ausplots.data$veg.PI #subset vegetation point intercept data
head(veg.PI) #shows the column headings and some data

#Calcualte species percent cover matrix
cover_matrix<- species_table(veg.PI, m_kind="percent_cover", cover_type="PFC")
write.csv(cover_matrix, file="cover_matrix_Pre2018.csv")
View(cover_matrix)
```

```
#"Herbarium determination" (aka species identification) includes a wide range of
taxonomic information, partly because there are not enough subheadings to record all
the interesting taxonomic details. For our purposes, we only want genus and species as
an identifier, because we are going to assume that different varieties and subspecies will
have the same photosynthetic pathway. Moreover, not all species can be identified to
the same level of detail (some only have family, or genus for example) so first we must
reduce the herbarium determination down to genus_species, where applicable.
cover_matrix<- read.csv("cover_matrix_Pre2018.csv", header = TRUE, sep = ",")
#First, we remove unneeded taxonomic detail
cover_matrix<-cover_matrix[ , -which(names(cover_matrix) %in% c("Var.1566"))]
names(cover_matrix)<-sub("X..chrysanthemoides.monilifera",
"Chrysanthemoides.monilifera", names(cover_matrix))
names(cover_matrix)<-sub("X.lachnagrostis.aemula", "Lachnagrostis.aemula",
names(cover_matrix))
names(cover_matrix)<-sub("X.leptorhynchos.squamatus.subsp..alpinus",
"Leptorhynchos.squamatus.subsp..alpinus", names(cover_matrix))
names(cover_matrix)<-sub("X.poa.costiniana", "Poa.costiniana", names(cover_matrix))
names(cover_matrix)<-sub("X.rhagodia.candolleana.subsp..candolleana",
"Rhagodia.candolleana.subsp..candolleana", names(cover_matrix))
names(cover_matrix)<-sub("X.sclerolaena.diacantha", "Sclerolaena.diacantha",
names(cover_matrix))
names(cover_matrix)<-gsub('[.]', '_', names(cover_matrix))#first convert all periods to _
names(cover_matrix)<-sub("_var.*", "", names(cover_matrix))#remove all info in
names that is not genus species
names(cover_matrix)<-sub("_subsp.*", "", names(cover_matrix))
names(cover_matrix)<-sub("_[[:digit:]].*", "", names(cover_matrix))
names(cover_matrix)<-sub("_f_.*", "", names(cover_matrix))
names(cover_matrix)<-sub("_f__.*", "", names(cover_matrix))
names(cover_matrix)<-sub("_x_.*", "", names(cover_matrix))
names(cover_matrix)<-sub("_benth_.*", "", names(cover_matrix))
names(cover_matrix)<-sub("_aff__.", "_", names(cover_matrix))
names(cover_matrix)<-sub("_aff_.", "_", names(cover_matrix))
names(cover_matrix)<-sub("__dc_.*", "", names(cover_matrix))
names(cover_matrix)<-sub("_steud_.*", "", names(cover_matrix))

#Final corrections to determinations (typos)
names(cover_matrix)<-sub("Triodia_basedowii_aff__t__lanigera",
"Triodia_basedowii", names(cover_matrix))
names(cover_matrix)<-sub("Eucalyptus_cyanopylla", "Eucalyptus_cyanophylla",
names(cover_matrix))
names(cover_matrix)<-sub("Euphordia_dummondii__", "Euphordia_dummondii",
names(cover_matrix))
names(cover_matrix)<-sub("Medicago__lacianata", "Medicago_lacianata",
names(cover_matrix))
names(cover_matrix)<-sub("Buchania", "Buchanania", names(cover_matrix))
names(cover_matrix)<-sub("Wahlenbergeria", "Wahlenbergia", names(cover_matrix))
```

```
#names(cover_matrix)<-sub("Desmodiump____trichostachyum",
"Desmodiump_trichostachyum", names(cover_matrix))
names(cover_matrix)<-sub("_s_.*", "", names(cover_matrix))
names(cover_matrix)<-sub("_ssp.*", "", names(cover_matrix))

names(cover_matrix)<-sub("Allocasuarina_muelleriana_l_a",
"Allocasuarina_muelleriana", names(cover_matrix))
names(cover_matrix)<-sub("Arthropodium_sp", "Arthropodium_sp_",
names(cover_matrix))
names(cover_matrix)<-gsub("_sp__", "_sp_",names(cover_matrix))
names(cover_matrix)<-sub("Atalaya_variifolia_hemiglauca", "Atalaya_hemiglauca",
names(cover_matrix))
names(cover_matrix)<-sub("Blennospora__drumondii", "Blennospora_drumondii",
names(cover_matrix))
names(cover_matrix)<-sub("Craspedia", "Craspedia_variabilis", names(cover_matrix))
names(cover_matrix)<-sub("Cyperaceae", "Cyperaceae_sp_", names(cover_matrix))
names(cover_matrix)<-sub("Geranium_sp", "Geranium_sp_", names(cover_matrix))
names(cover_matrix)<-sub("Liliaceae", "Liliaceae_sp_", names(cover_matrix))
names(cover_matrix)<-sub("Rubiaceae", "Rubiaceae_sp_", names(cover_matrix))
names(cover_matrix)<-sub("Tephrosia_sp_macarthur_river", "Tephrosia_sp_",
names(cover_matrix))
names(cover_matrix)<-sub("Tephrosia_sp_willowra__g_m_chippendale",
"Tephrosia_sp_", names(cover_matrix))

write.csv(cover_matrix, file="correct_cover_matrix.csv")

#Convert those entries that only have a genera to genera_sp so that they can be
identified later
names(cover_matrix)<-gsub("_sp_.*","_sp_",names(cover_matrix))
colnames(cover_matrix) #double check new names

#Now we need to combine genus_species that may have multiple entries (columns)
because we have now reduced the entries down to genus_species
#In other words, subspecies that were recorded as distinct entries, their % cover values
can now be combined in each plot
indx <- sapply(cover_matrix, is.numeric) #Check which columns are numeric and save
as a vector
nm1 <- which(indx) #Check which among the integer columns are duplicated
indx2 <- duplicated(names(nm1))|duplicated(names(nm1),fromLast=TRUE) #index
duplicated names
nm2 <- nm1[indx2]
indx3 <- duplicated(names(nm2))
cover_matrix[nm2[!indx3]] <- Map(function(x,y) rowSums(x[y]),
                        list(cover_matrix),split(nm2, names(nm2)))
datN <- cover_matrix[ -nm2[indx3]]
write.csv(datN, file="Cover_Clean_Pre2018.csv")

#Step_2: Assigning species a photosynthetic pathway
```

```
#First, import the cover matrix
cover_matrix<- read.csv("Cover_Clean_Pre2018.csv", header = TRUE, sep = ",")
#Next, the cover matrix needs to be "melted" in order to change the structure of the data frame
md <- melt(cover_matrix, id=(c("Site_Visit")))
head(md)
#Change column names
names(md)[2]<-paste("Species.name")
names(md)[3]<-paste("Cover")

#Next, we want to create a column for just the genera
md$genus <- word(md$Species.name,1, sep = fixed('_'))
unique(md$genus)

#Now combine the TERN species list with our list of species and their photosynthetic pathways
#Begin by uploading the PPdatabase
PPdatabase <- read.csv("PP_database_SIA_final_assignments.csv")
head(PPdatabase)
#Then, finally, we can merge the data frames
Pathway.assignment <- merge(md, PPdatabase, by.x='Species.name', all.y=TRUE)
Pathway.assignment<-Pathway.assignment[ , -which(names(Pathway.assignment) %in% c("X"))]
head(Pathway.assignment)
write.csv(Pathway.assignment, file="Pathway.assignment_Pre2018.csv")
Pathway.assignment<- read.csv("Pathway.assignment_Pre2018.csv", header = TRUE, sep = ",")
#note that after the first merge, we identified species in the TERN database not in the PP database, so we have to go back and add them with PP pathways
#we need to search other sources for potential options

#Removing any sites that are not part of the 20
PAssignment_new <- subset(Pathway.assignment, Site_Visit %in% c("NTADAC0001-53518", "NTAGFU0031-53678", "NTAGFU0040-53687", "NTAGFU0017-53664", "NTAGFU0010-53657", "NTAGFU0008-53655", "NTAGFU0001-53648", "NTABRT0004-53619", "NTAFIN0019-53639", "NTAFIN0022-53642", "SAASTP0004-53722", "SAASTP0001-53719", "SATFLB0005-58656", "SATFLB0008-53752", "SATFLB0010-53714", "SATFLB0014-53702", "SATFLB0012-58677", "SATFLB0015-58676", "SATKAN0002-53689", "SATKAN0001-53688"))
write.csv(PAssignment_new, file="PAssignment_20Sites.csv")

Pathway.assignment2<- read.csv("PAssignment_20Sites.csv", header = TRUE, sep = ",")

#Check for any NAs, these are species in the TERN database but not account for in the PPdatabasr
NA_pathways<- PAssignment_new[is.na(PAssignment_new$Site_Visit),]
```

```
head(NA_pathways)
missing_names<-unique(NA_pathways$Species.name)
View(missing_names)
write.csv(missing_names, file="missing_names.csv")
#alter list as needed

#Step 3: Calculate C3 and C4 cover
#We now have a data frame that provides the PP for every species recorded by the TERN point-intercept method. We are well on our way to running all sorts of models and analysis, but first, we need to:
#1. work out the proportional cover of C3/C4 at each plot
#2. add in all needed relevant site data, such as Lat and Long, and other info about the plot location that might be useful
#Total C4 abundance
#first, we can use the ddply package to create summary stats for each site
sum_stats<-ddply(Pathway.assignment2, c("Site_Visit", "Photosynthetic.pathway"), summarise, Total_Cover = sum(Cover, na.rm = T))
#this gives the total cover of each of the different photosynthetic pathways at each plot
#this function can be used to get all sorts of summary stats
head(sum_stats)

#next, we want to subset only C3 and C4 species
C3_C4 <- subset(Pathway.assignment2, Pathway.assignment2$Photosynthetic.pathway %in% c('C3','C4'))
head(C3_C4) #so we have eliminated the other plants with different or unknown PP

#next, we calculate the total cover of C3 and C4 species at each plot
sum_stats_C3_C4<-ddply(C3_C4, c("Site_Visit", "Photosynthetic.pathway"), summarise, Total_Cover = sum(Cover, na.rm = T))
head(sum_stats_C3_C4)
#next we calculate the relative C3/C4 cover (% C4) = C4 cover/C4 + C3 cover
sum_stats_C3_C4<-na.omit(sum_stats_C3_C4)
cover_summary<-
sum_stats_C3_C4%>%group_by(Site_Visit)%>%summarise(c3_c4_cover = sum(Total_Cover),
relative_C4_cover=Total_Cover[which(Photosynthetic.pathway=="C4")]/sum(Total_Cover))
head(cover_summary)

write.csv(cover_summary, file="cover_summaryunfinished.csv")

cover_summary2<- read.csv("cover_summaryunfinished.csv", header = TRUE, sep = ",")

#now we need to add relevent metadata, such as location
cover_summary2$longitude <-
my.ausplots.data$site.info$longitude[match(cover_summary2$Site_Visit,
```

```
my.ausplots.data$site.info$site_unique)] #and lat and long information to the
spreadsheet
cover_summary2$latitude <-
my.ausplots.data$site.info$latitude[match(cover_summary2$Site_Visit,
my.ausplots.data$site.info$site_unique)]
cover_summary2$bioregion_name <-
my.ausplots.data$site.info$bioregion_name[match(cover_summary2$Site_Visit,
my.ausplots.data$site.info$site_unique)]
cover_summary2$soil_observation_type <-
my.ausplots.data$site.info$soil_observation_type[match(cover_summary2$Site_Visit,
my.ausplots.data$site.info$site_unique)]
View(cover_summary2)
write.csv(cover_summary2, file="cover_summary.csv")
cover_summary<- read.csv("cover_summary.csv", header = TRUE, sep = ",")
```

APPENDIX D. CARBON ISOTOPIC COMPOSITION AND STANDARD DEVIATION

Table D.1. Bulk SOM and *n*-alkane carbon isotopic ($\delta^{13}C$) values and standard deviation for each run of the data on the EA-IRMS.

Plot name	Run 1 SOM $\delta^{13}C$ (‰)	σ	Run 2 SOM $\delta^{13}C$ (‰)	Run 3 SOM $\delta^{13}C$ (‰)	σ	Run 1 *n*-alkane $\delta^{13}C$ (‰)	Run 2 *n*-alkane $\delta^{13}C$ (‰)	σ
NTABRT0004	-21.55	0.16	-21.69	-21.46	0.09	-32.43	-32.03	0.28
NTADAC0001	-142.65	0.16	-26.81	-26.75	0.09	-32.05	-32.22	0.12
NTAFIN0019	-20.25	0.16	-20.27	-20.40	0.09	-30.71	-30.46	0.18
NTAFIN0022	-22.96	0.16	-23.14	-23.03	0.09	-31.86	-32.12	0.19
NTAGFU0001	-18.56	0.16	-18.10	-17.84	0.09	-26.69	-26.59	0.07
NTAGFU0008	-15.91	0.16	-15.55	-15.42	0.09	-23.20	-22.39	0.57
NTAGFU0010	-21.79	0.16	-22.42	-23.24	0.09	-22.53	-22.06	0.33
NTAGFU0017	-20.94	0.16	-20.66	-20.72	0.09	-32.20	-32.31	0.08
NTAGFU0031	-24.61	0.16	-24.46	-24.22	0.09	Contaminated		
NTAGFU0040	-23.37	0.16	-22.06	-21.86	0.09	-29.94	-29.52	0.30
SAASTP0001	-23.53	0.16	-23.14	-23.15	0.09	-31.31	-31.45	0.10
SAASTP0004	-19.25	0.16	-18.28	-17.49	0.09	-30.11	-30.10	0.01
SATFLB0005	No beam		-25.85	-25.75	0.09	-33.05	-33.01	0.03
SATFLB0008	-26.54	0.16	-23.89	-22.70	0.09	-29.90	-29.94	0.02
SATFLB0010	-24.38	0.16	-23.57	-24.37	0.09	-33.35	-33.32	0.03
SATFLB0012	-27.59	0.16	-27.06	-26.98	0.09	-34.00	-33.99	0.01
SATFLB0014	-25.76	0.16	-25.54	-25.55	0.09	-32.49	-32.49	0
SATFLB0015	-28.76	0.16	-28.04	-27.92	0.09	-35.21	-35.41	0.14
SATKAN0001	-28.67	0.16	-27.37	-27.51	0.09	-34.76	-34.62	0.10
SATKAN0002	-27.94	0.16	N/A	-27.48	0.09	-34.13	-34.22	0.07

APPENDIX E. PROPORTIONAL COVER AND CLIMATE DATA

Table E.1. Proportional grass, tree and chenopod cover derived from the vegetation survey data. Mean annual temperature (MAT) and seasonal water availability (SWA) were averaged across the 1970 to 2018 records from the Australian Gridded Climate Data set from the Australian Government Bureau of Meteorology (2018).

Plot name	Proportional grass cover (%)	Proportional tree cover (%)	Proportional chenopod cover (%)	Seasonal water availability (%)	Mean annual temperature (°C)
NTABRT0004	37.6	0.0	1.8	92.9	22.8
NTADAC0001	63.5	19.8	0.0	100	26.6
NTAFIN0019	72.4	18.8	3.0	85.9	21.9
NTAFIN0022	31.7	0.0	0.7	85.3	22.0
NTAGFU0001	67.6	11.0	0.5	98.9	25.6
NTAGFU0008	72.6	4.9	0.3	98.9	25.4
NTAGFU0010	68.0	31.4	0.0	100	25.8
NTAGFU0017	48.2	4.9	0.0	100	25.6
NTAGFU0040	32.1	23.7	0.0	100	27.7
SAASTP0001	27.4	0.0	40.2	83.1	22.0
SAASTP0004	7.1	0.4	24.6	83.5	22.4
SATFLB0005	0.6	21.6	0.0	48.1	17.5
SATFLB0008	47.5	0.0	0.0	34.4	14.5
SATFLB0010	0.5	58.8	19.2	39.5	16.9
SATFLB0012	4.6	11.6	0.0	25.9	15.8
SATFLB0014	1.7	33.5	0.0	31.1	14.5
SATFLB0015	4.7	46.7	0.0	18.1	14.1
SATKAN0001	3.9	45.0	0.0	15.6	14.5
SATKAN0002	0.5	53.2	0.0	19.0	14.0