

1 **ausplotsR: An R package for rapid extraction and analysis of vegetation**
2 **and soil data collected by Australia's Terrestrial Ecosystem Research**
3 **Network**

4
5 Samantha Munroe^{1,2}, Greg Guerin^{1,2}, Tom Saleeba^{1,2}, Irene Martín-Forés^{1,2}, Bernardo
6 Blanco-Martin³, Ben Sparrow^{1,2}, and Andrew Tokmakoff^{1,2}

7
8 ¹ School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005,
9 Australia

10 ² Terrestrial Ecosystem Research Network (TERN), University of Adelaide, Adelaide, South
11 Australia 5005, Australia

12 ³ Department of Agriculture and Fisheries, Queensland Government, Australia

13
14 Correspondence:

15 Samantha Munroe, Terrestrial Ecosystem Research Network (TERN), University of
16 Adelaide, Adelaide, South Australia 5005, Australia

17 Email: samantha.munroe@adelaide.edu.au

18
19 **Funding Information:**

20 TERN is supported by the Australian Government through the National Collaborative
21 Research Infrastructure Strategy.

22
23
24
25

26 **Abstract**

27 The Terrestrial Ecosystem Research Network (TERN), Australia’s national land ecosystem
28 monitoring program, measures critical environmental attributes from local to continental
29 scale and generates quality data for research and land management. Since 2011, TERN has
30 performed standardised field surveys and sampling across a national plot network. At each
31 plot, TERN records vegetation structure, composition and diversity, soil characteristics, and
32 collects plant and soil samples for analysis. At the time of submission, TERN has established
33 over 750 plots and performed over 1000 plot surveys across Australia. Here we present
34 *ausplotsR*, an R package for the R statistical computing environment that provides a user-
35 friendly interface to rapidly import, visualise, and analyse TERN plot survey data. Easy-to-
36 use functions extract the data and compile data tables that can be incorporated into a variety
37 of statistical analysis, most notably multivariate applications requiring plant community data
38 with standardised relative abundances. *ausplotsR* includes functions to calculate useful
39 vegetation metrics, such as species presence/absence, cover, and basal area. The package also
40 provides information on TERN’s extensive soil and plant sample collection. We expect
41 *ausplotsR* will help facilitate and advance ecological research and management throughout
42 Australia and provide useful data for vegetation modellers globally.

43
44 **Keywords:** Australia, big data, biodiversity, ecosystem monitoring, environmental
45 modelling, plot survey, R statistical environment, Terrestrial Ecosystem Research Network
46 (TERN)

47
48
49
50
51

52 **Introduction**

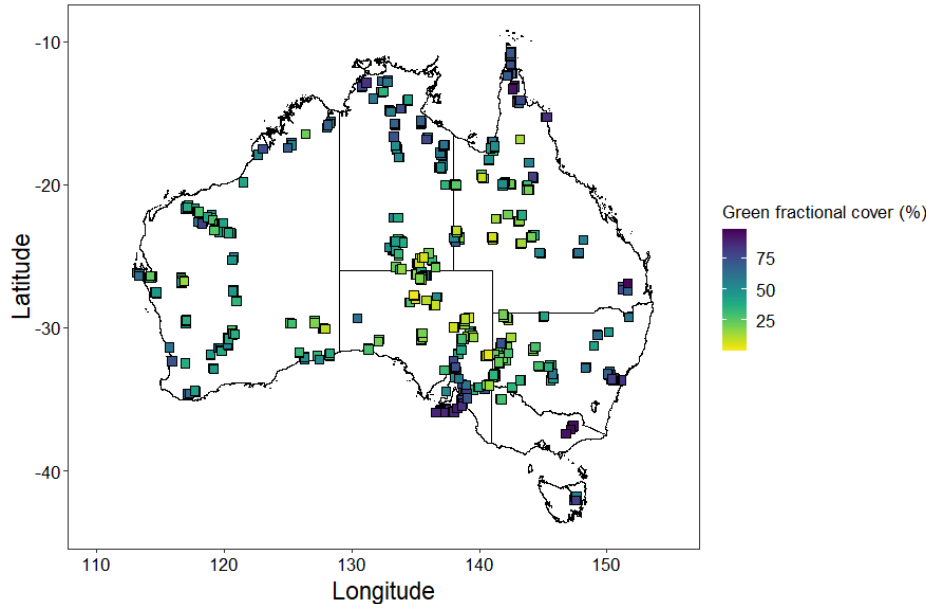
53 Understanding the distribution and composition of terrestrial plant communities requires
54 large amounts of reliable and diverse data. Examining important issues such as the geography
55 of plant functions (Šimová et al. 2018), invasive species (Pyšek et al. 2020), or the
56 maintenance of ecosystem services (Kubiszewski et al. 2020), requires wide-spread
57 ecological community datasets (Kao et al. 2012; Kissling et al. 2018). The rapid increase of
58 continental and global analyses in vegetation science (e.g., Velazco et al. 2017; Bruelheide et
59 al. 2018; Jiménez-Alfaro et al. 2018) has also triggered the development of new tools and
60 software that facilitate prompt data access and analysis (e.g., Maitner et al. 2018; Kattge et al.
61 2020). Thus, the challenge lies not only in constructing comprehensive datasets, but also in
62 designing user-friendly data delivery systems that provide open access to standardised and
63 often complex databases (Chytrý et al. 2019).

64
65
66 The Terrestrial Ecosystem Research Network (TERN), Australia's ecosystem observatory,
67 measures important terrestrial ecosystem attributes over time from local to continental scale
68 at hundreds of sites across the country. TERN provides freely-accessible data to empower
69 scientists to detect and understand patterns and changes in terrestrial ecosystems. TERN is
70 comprised of data collection platforms that gather complementary data at scales ranging from
71 remote sensing to micro-meteorological observations (Sparrow et al. 2020a).

72
73 TERN's Ecosystem Surveillance platform performs standardised field surveys and sampling
74 across a national plot network. This platform collects crucial data for effective monitoring of
75 Australia's ecosystems, including vegetation structure and composition, soil characteristics,
76 and soil and plant samples for analysis (Sparrow et al. 2020b). TERN has established over
77 750 plots and conducted over 1,000 surveys across every major terrestrial environment in

78 Australia (Figure 1), recoding >5,000 unique species and collecting >65,000 plant and soil
79 samples (Table 1).

80



81

82 Figure 1. TERN AusPlots monitoring plot locations as generated within the package coded by
83 percent fractional green cover.

84

85 Table 1. TERN survey and sample records available through *ausplotsR*

Data Type	Count
Number of plots surveyed	752
Number of plots revisited	106
Number of plant voucher specimens collected	42505
Number of leaf tissue samples available	52065
Number of plant species collected	5245
Number of point-intercepts	872761
Number of soil bulk density samples	1675
Number of soil metabarcoding samples available	8091
Total number of soil samples	17,082

86

87 The *ausplotsR* package was designed for the R statistical computing environment (R Core

88 Development Team 2020) to provide free and simple access to the TERN Ecosystem

89 Surveillance plot survey database (Guerin et al. 2020). The package has a straightforward

90 workflow to enable a range of vegetation analyses. First, easy-to-understand functions extract

91 raw survey data for all plots within the network from a regularly updated database. Second,

92 raw data can be incorporated into downstream functions that calculate a variety of metrics,
93 such as species cover, fractional cover, and basal area. *ausplotsR* also provides information
94 on TERN's substantial soil and plant sample library, such as voucher numbers and the date
95 and location samples were collected. Data are formatted to support compatibility with global
96 (e.g. Kattge et al. 2020) and Australian plant trait datasets (e.g. Falster et al. in review) and
97 can be integrated with additional TERN data products. The goal of *ausplotsR* is to facilitate
98 quality ecosystem research and effective land management across Australia through delivery
99 and pre-processing of field data. Instant access to continental-scale plant community data in
100 the R environment provides a valuable resource to vegetation scientists and modellers for
101 testing ecological ideas, tools, and methods.

102
103

104 **TERN plot survey methods**

105 TERN plots are surveyed using the 'AusPlots Rangelands' method, composed of modules to
106 collect vegetation and soil data (Table 2; also see Sparrow et al. 2020b). Here we provide a
107 short overview of the modules used to collect data available through *ausplotsR*.

108 Table 2. Modules in the AusPlots rangelands monitoring method and the data types available from the package *ausplotsR* (Sparrow et al. 2020b)

Module	Protocol	Application	Data available in <i>ausplotsR</i> ?
Plot layout	Accurate layout using DGPS; installation of permanent markers.	Accurate relocation; remote sensing validation	Yes
<u>Vegetation</u>			
Photo-panoramas	Collection of 360° photographs from three points	Computer vision analysis, point clouds and measures of basal area	No
Vouchering	Collection of vascular plant species	Taxonomy; spatial/temporal analysis of presence—absence	Yes
Tissue samples	Collection of single tissue samples from vascular plants	Genetic/isotopic analysis	Yes
Point-intercept	Collection of species, height, phenology, growth-form, senescence at 1010 points	Change in relative abundance, cover and structure; remote sensing validation	Yes
Basal area	Collection by species using basal wedge at nine points	Convertible to biomass	Yes
Structural summary	Recording of three dominant species in each of three strata (upper, mid, ground)	Community descriptions	Yes
Leaf Area Index	Collection of at least 50 evenly spaced readings with the LiCor LAI 2200 LAI meter	Ecophysiological modelling; remote sensing validation	No
<u>Soils and Landscapes</u>			
Plot description	Record location, substrate, microtopography, erosion/disturbance	Assessment of characteristics/impact of disturbance	Yes
Soil pit characterisation	Collection of soil samples/data at 10 cm increments or identifiable horizons to 1 m	Characterisation and classification. Correlate with vegetation	Yes
Sub-site characterisation	Collection of nine samples in differing microhabitats at 0-10, 10-20 and 20-30 cm	Soil variability across plot	Yes
Bulk density	Collection of three measures at the soil pit at 0-10, 10-20 and 20-30 cm	Conversion to volumetric measures	Yes
Soil metagenomics	Collection of nine samples	Identify biota	Yes

110 One-hectare plots are established in a homogenous area of terrestrial vegetation. Vegetation
111 structure and composition are measured using the point-intercept module. Transects (10 x
112 100 m) are laid out in a grid pattern spaced 20 m apart. Species identity, growth form, height,
113 and systematic absences (e.g. bare ground) are recorded at 1 m points along transects,
114 resulting in 1010 survey points. Vouchers of each species are collected and sent to herbaria
115 for identification.

116
117
118 Soil modules collect information on a range of physical and chemical soil characteristics. A 1
119 m deep pit is dug in the southwest corner of the plot, which enables the description of the
120 upper soil profile. TERN measures soil pH, bulk density, electrical conductivity, texture,
121 colour, and structure. Soil samples are also collected at nine sub-sites across the plot to assess
122 microhabitat variability and enable metagenomic analysis of environmental DNA. Plant and
123 soil samples are tracked using alphanumeric barcode labels and stored for later analysis.

124
125 AusPlots data is collected via the AusScribe app, a custom Android/iOS app that is designed
126 for operation on tablets (Tokmakoff et al. 2016). As observers complete data entry for a plot,
127 they upload collected data from the AusScribe app to a backend system which makes the data
128 available for curation by TERN staff. Once curated, visits are marked as published and
129 automatically made available via *ausplotsR*. This data pipeline ensures new data is made
130 rapidly available to users.

131

132 **The *ausplotsR* package**

133 The *ausplotsR* package (CRAN: <https://CRAN.R-project.org/package=ausplotsR>; latest
134 development version and patches: <https://github.com/ternaustralia/ausplotsR>) provides access
135 to most data modules collected. *ausplotsR* sources its data via a dedicated server stack that is
136 running inside a Virtual Machine in the Australian Research Data Common's (ARDC's)

137 NECTAR research cloud. The stack contains a number of services which together: (1) expose
138 the published AusPlots data for external use, (2) collect usage statistics, (3) Reverse-Proxy
139 the API for scalability and provides threat protection, and (4) authenticate/authorise users for
140 fine-grained data access control. This open source stack is available at:
141 <https://github.com/ternandsparrow/swarm-rest>. *ausplotsR* functions enable users to quickly
142 read and prepare survey data for ecological analysis. Next, we review some key *ausplotsR*
143 functions and workflow to extract, visualise, and analyse TERN plot data (Table 3).

DRAFT

144 Table 3. Descriptions of *ausplotsR* functions

Function	Description
<i>ausplots_visual</i>	Generates a set of graphical displays representing TERN AusPlots data based on geographic locations and vegetation attributes from the plot-based point intercepts (as generated by <i>get_ausplots</i> and other pre-processing functions)
<i>basal_area</i>	Calculates basal area (or number of basal wedge hits) for each plot, using the raw basal wedge data returned from <i>get_ausplots</i>
<i>fractional_cover</i>	Calculates fractional cover (i.e., the proportional cover of green vegetation, dead vegetation and bare substrate) based on plot-based point-intercept data (as generated by <i>get_ausplots</i>)
<i>get_ausplots</i>	This function is the starting point for accessing data through the <i>ausplotsR</i> package. It extracts the data
<i>growth_form_table</i>	Generates occurrence matrices for plant growth forms in plots as desired based on presence/absence, percent cover or species richness (i.e., the number of species assigned to a growth form). The input is a data frame of raw point intercept data generated using the <i>get_ausplots</i> function
<i>optim_species</i>	This function applies different optimisation methods to select a subset of plots that maximise species accumulation. The function operates under the ‘Maximum covering problem’ framework
<i>plot_opt</i>	This function plots different species accumulation curves obtained through different optimisers in the <i>optim_species</i> function
<i>single_cover_value</i>	Calculates a single vegetation cover value per site based on the plot-based point-intercept data generated by <i>get_ausplots</i> . Cover can be subsetting to vegetation that has a specified minimum or maximum height and/or by plant growth forms
<i>species_list</i>	Nice species lists based on the vegetation voucher module of Ausplots generated using the <i>get_ausplots</i> function
<i>species_table</i>	This function takes a data frame of individual raw point-intercept hits generated using the <i>get_ausplots</i> function, and generates species occurrence matrices as desired based on presence/absence, cover, frequency or IVI

146 *Data extraction*

147 Data are extracted using the function *get_ausplots* (Table 4). By default, the *get_ausplots*
148 function extracts a list of data tables for the point-intercept (*\$veg.PI*) and vegetation voucher
149 (*\$veg.vouchers*) modules for all surveys. It also provides a site information table (*\$site.info*)
150 which describes survey details and environmental features. As demonstrated below,
151 arguments of '*get_ausplots*' allow users to select data for individual plots (*my.Plot_IDs*),
152 rectangular spatial coordinates (*bounding_box*), plant families (*family_search*), or species
153 (see *Data Format*).

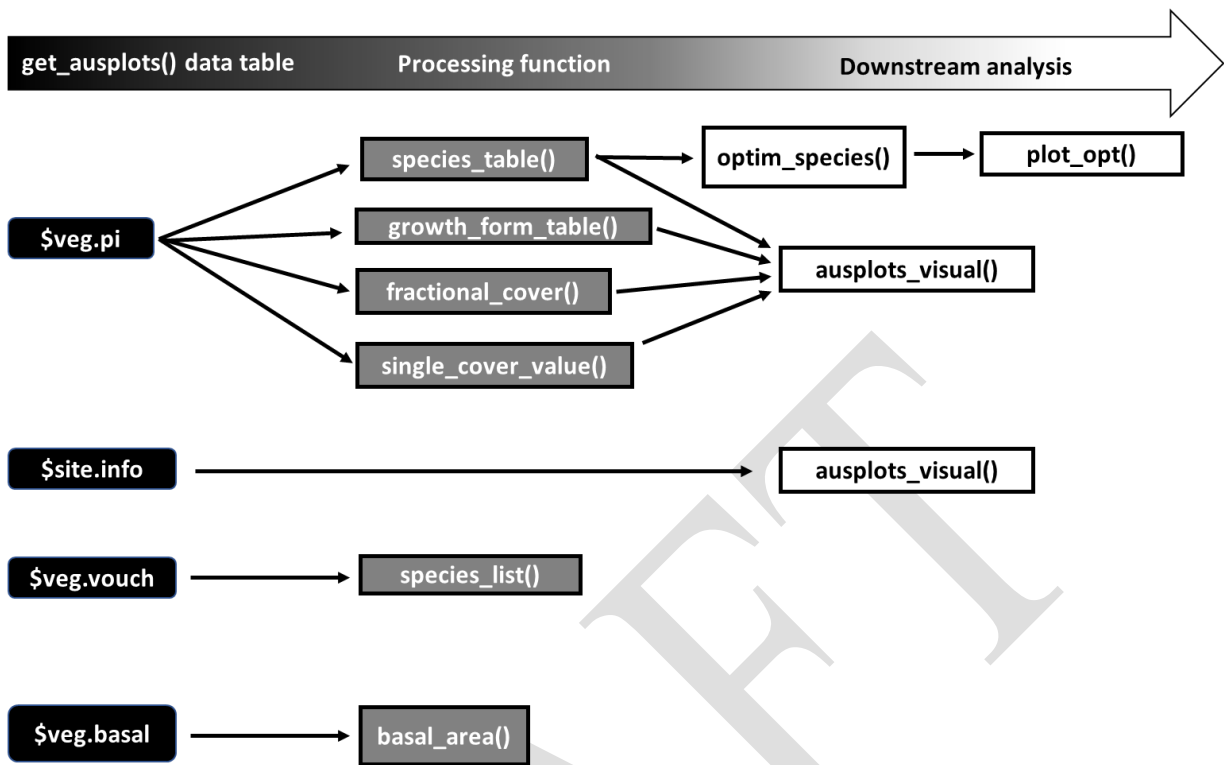
```
154  
155     # for selected plots:  
156     my.data <- get_ausplots(my.Plot_IDs=c("SATFLB0004", "QDAMGD0022"))  
157  
158     # plots within a specific geographic area:  
159     my.data <- get_ausplots(bounding_box= c(120, 140, -30, -10) # data are in longlat  
160  
161     # plots where the family "Myrtaceae" was recorded:  
162     Myrtaceae <- get_ausplots(family_search="Myrtaceae")  
163  
164
```

165 Users can also filter searches by survey module and request data not extracted by default,
166 such as basal wedge data (*basal.wedge=TRUE*) or soil characteristics
167 (*soil_character=TRUE*). Once extracted, data tables can be incorporated into various
168 downstream processing functions that calculate vegetation indices, such as species cover or
169 basal area (Figure 2).

170 Table 4. Data tables generated by *get_ausplots()*

Module	get_ausplots() data table	Description
*Plot layout	\$site.info	Data frame with basic site information including location
Vegetation		
*Point-intercept	\$veg.PI	Data frame with individual point-intercept data
*Vascular plant vouchering and tissue samples	\$veg.vouch	Data frame with rows for each voucher and information on species determinations and silica-dried tissue samples
Basal area	\$veg.basal	Data frame with compiled raw basal wedge hit data
Soil and Landscapes		
Structural summary	\$struct.summ	Data frame with vegetation structural summaries for each plot
Sub-site characterisation and soil metagenomics	\$soil.sub	Data frame with details of soil subsites within each plot including sample barcode identification
Bulk density	\$soil.bulk	Data frame with raw bulk density data from each plot
Soil pit characterisation	\$soil.char	Data frame with soil characterisation data from the 1 m pit at the SW corner of each plot
Metadata		
Variable dictionary	\$metadata.dictionary	Lists and describes each variable and corresponding values in each data frame
*Data citation	\$citation	Auto-generated citation for the data extracted

171 * *Extracted by default*



173
174

175 Figure 2. Workflow of the *ausplotsR* package, demonstrating how raw data tables generated
176 from *get_ausplots()* can be incorporated in processing and downstream functions

177
178
179

Data format

180
181
182
183
184
185
186
187
188
189

All *ausplotsR* data tables list data by survey. Unique plot surveys are identified by their *site_location_name* (plot ID used to distinguish each unique plot) and *site_location_visit_id* (unique numeric value that delineates each survey visit). *site_location_name* is an alphanumeric value that indicates State/Territory (e.g., Western Australia, W.A.; South Australia, S.A., etc.) and bioregion (Thackway & Cresswell 1995), as well as a sequential number based on the number of plots in that bioregion. For example, the *site_location_name* ‘SAAFLB0008’ indicates the plot is in South Australia (SAA), in the Flinders Lofty Block (FLB) bioregion, and it was the eighth plot in that bioregion. *site_location_name* and *site_location_visit_id* are concatenated into the *site_unique* field that identifies unique plot visits.

190 The point-intercept (*\$veg.PI*) and vegetation voucher (*\$veg.vouchers*) data tables include
191 columns with taxonomic information: *family*, *genus*, *specific_epithet*, *genus_species* (*genus*
192 and *specific_epithet* combined), and *infraspecific_epithet* (e.g., subspecies, variety).
193 *ausplotsR* provides two species name options: *herbarium_determination* and
194 *standardised_name*. *herbarium_determination* contains species identifications to the lowest
195 possible taxonomic rank provided by herbaria. However, nomenclature sometimes differs
196 between States (although this problem is rare). *herbarium_determination* values also include
197 vegetation identifications for incomplete, dead, or generic specimens (e.g. "Dead
198 Tree/Shrub", "Annual Grass").

199
200 Alternatively, species identifications can be taken from the *standardised_name* field.
201 *standardised_name* values are based on the *herbarium_determination* values standardised to
202 match the most widely accepted scientific name synonym according to 'World Flora Online'
203 (www.worldfloraonline.org). This helps ensure scientific names will not differ between
204 States/Territories and increases consistency with global databases. It also excludes non-
205 standard entries such as 'dead' identifications. The remaining taxonomic fields are derived
206 from the *standardised_name*. Plot data extractions can be initially filtered by
207 *herbarium_determination* or *standardised_name*:

```
208  
209     # plot and voucher records where "Eucalyptus moderata" was identified:  
210     Eucalyptus_moderata <- get_ausplots(herbarium_determination_search="Eucalyptus  
211     moderata")  
212  
213
```

214 *Data processing and calculation of vegetation indices*

215

216

217 *1. Community composition matrices*

218 The *species_table* function takes the data table of individual point-intercept hits (*\$veg.PI*) and
219 returns species occurrence matrices. *species_table* can calculate species presence/absence,
220 percent cover, frequency (based on occurrences on different transects), or Importance Value

221 Index. Users can select the preferred species name option from *standardised_name* (the
222 default), *herbarium_determination*, or *genus_species*. To ensure nomenclatural consistency
223 when comparing plots, we recommend using *standardised_name* (`species_name="SN"`) or
224 *genus_species* (`species_name="GS"`). For example, to compare species level cover:

```
225  
226     species_table(my.data$veg.PI, m_kind="percent_cover", species_name="GS")
```

227
228
229 *2. Vegetation cover:*

230 Several functions in *ausplotsR* calculate different aspects of vegetation (and ground) cover
231 based on raw point-intercept input data extracted using *get_ausplots* (`$veg.PI`).

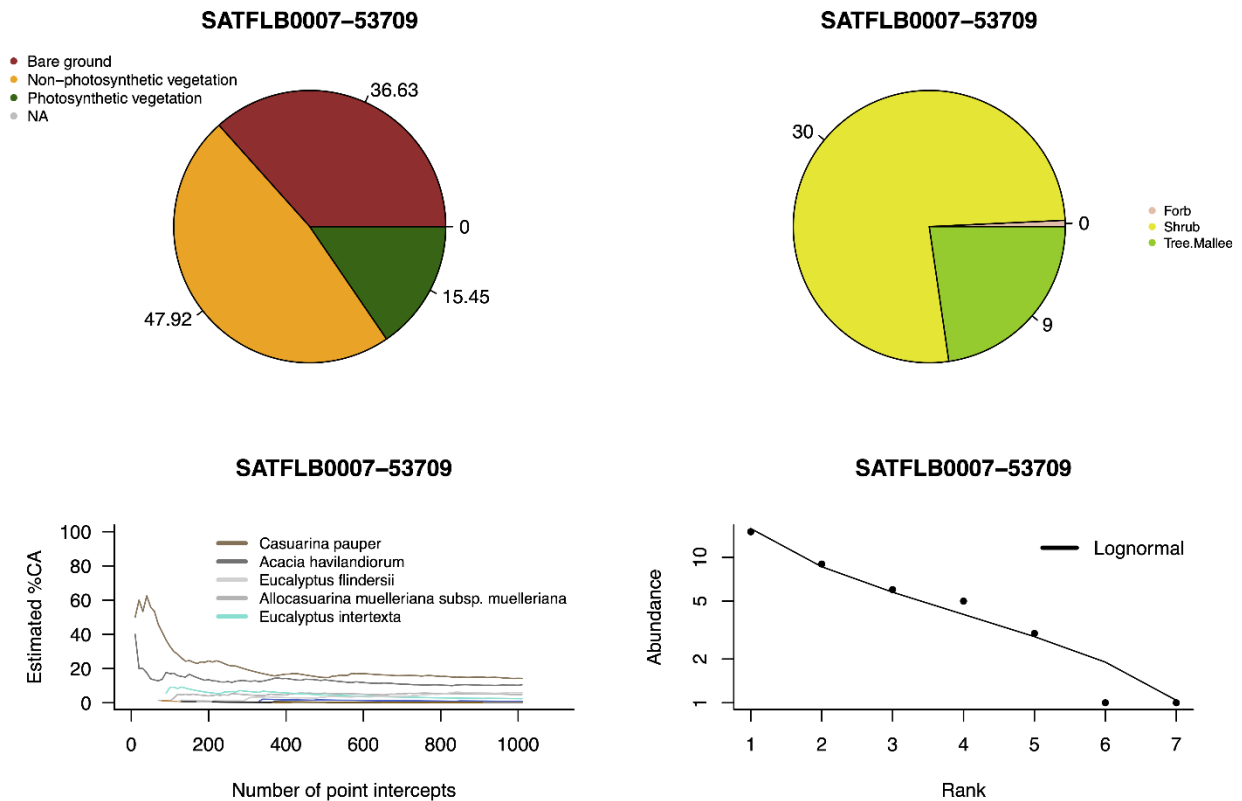
232 *growth_form_table* generates occurrence matrices for plant growth forms based on
233 presence/absence, cover, or species richness:

```
234  
235     growth_form_table(my.veg.PI, m_kind="percent_cover")
```

236

237 Other functions calculate fractional cover (i.e. proportional cover of living vegetation, dead
238 litter and bare ground; *fractional_cover*; Figure 1), cover of individual growth forms subset
239 by height such as the cover of trees over 5 m tall (*single_cover_value*), and the cross-
240 sectional area of tree trunks as measured through the basal wedge module (*basal_area*). The
241 relative cover of species, plant growth forms, and fractional cover of living and dead
242 vegetation as calculated using these functions is rapidly illustrated per plot using the function
243 *ausplots_visual* (Figure 3).

244



246

247 Figure 3. Example of data presentation panels for a single plot visit using *ausplots_visual()*:
 248 a) fractional vegetation cover; b) the relative abundance of plant growth forms; c) cumulative
 249 estimates of percent cover by species as point-intercept hits are taken across the plot (Guerin
 250 et al. 2017); d) Whittaker plot (Whittaker 1965) of species relative abundance fitted with a
 251 lognormal SAD curve.

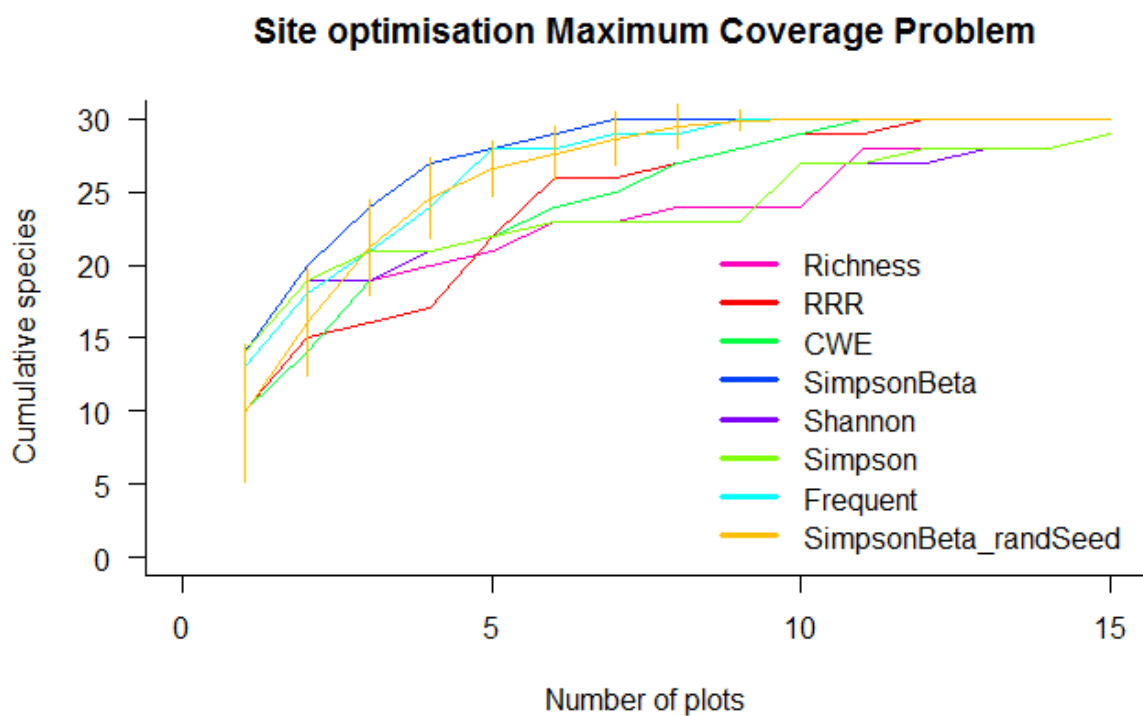
252

253

254 3. Optimising species accumulation

255 The *optim_species* function applies different biodiversity metrics as optimisers to select a
 256 subset of plots that maximise species accumulation. The framework underlying this function
 257 is the 'Maximal coverage problem', applied in conservation biology to design natural
 258 reserves (Church et al. 1996). The function seeks to identify plots that will 'protect' the
 259 maximum number of species in a limited number of sites. The biodiversity metrics included
 260 are species richness (biodiversity hotspots), range rarity richness (high biodiversity and
 261 uniqueness; Guerin and Lowe 2015), corrected weighted endemism (areas with range-

262 restricted endemic species; Crisp et al. 2001), Shannon-Wiener diversity index and the
263 Simpson diversity index (which include species relative abundances), and the Simpson
264 dissimilarity (maximises species turnover; Baselga & Leprieur 2015). The user specifies the
265 number of plots to select, and each optimiser selects the subset of plots that accumulates the
266 largest number of species. The function also includes the option of visually comparing the
267 optimisations (Figure 4). The input data is a species versus sites matrix; thus, it can be used
268 with the data generated with *species_table* from *ausplotsR* or with any other vegetation
269 dataset.
270



271

272 Figure 4. Species accumulation plot comparing different biodiversity metrics included in
273 *optim_species*. The image has been created using the *dune* dataset from the *vegan* package as
274 an input, selecting 15 plots, and 60 iterations to calculate the most frequently selected ones
275 based on random starts.
276

277 **Data Licence and authorship guidelines**

278 Data provided by TERN through *ausplotsR* are distributed by a CC-BY Creative Commons
279 license (<https://creativecommons.org/licenses/by/4.0/>). Publications using data provided by
280 TERN retrieved using *ausplotsR* should cite both the data and the package. The suggested
281 citation for the plot data is automatically generated by *ausplotsR* when you extract data. The
282 citation for the *ausplotsR* package can be printed using the command:

```
283  
284     citation('ausplotsR')  
285  
286  
287
```

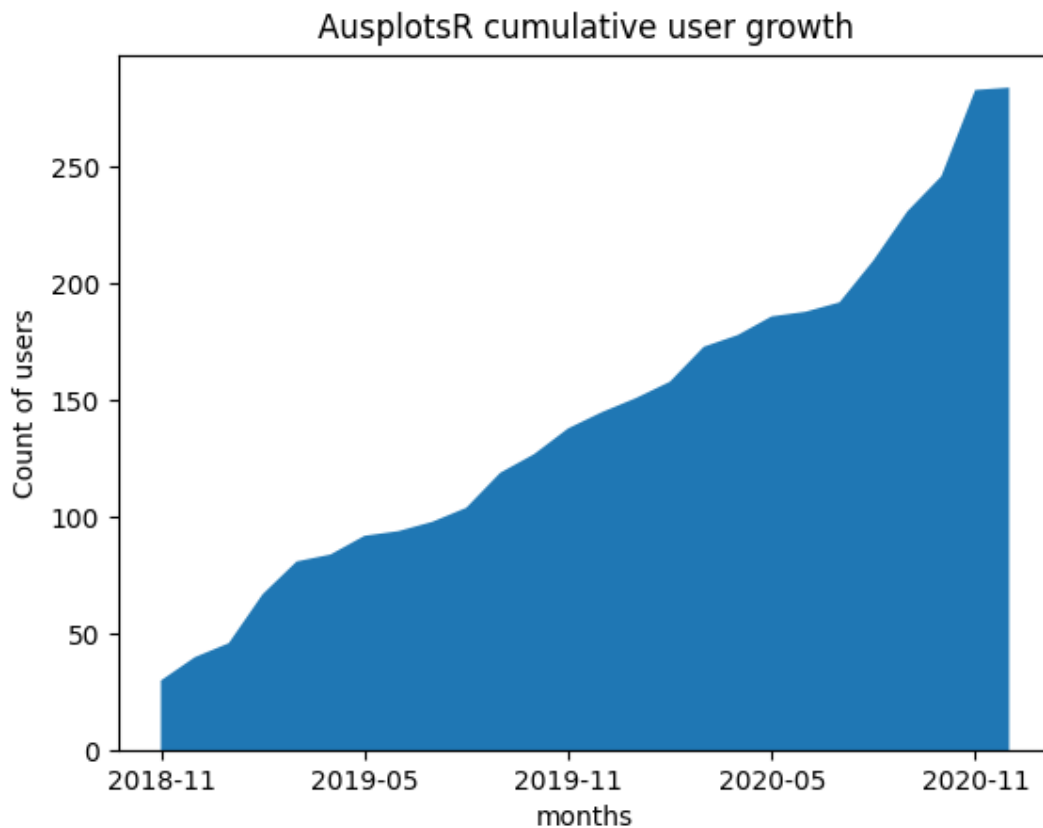
288 **Applications and uptake**

289 *ausplotsR* makes it easy for scientists to quickly access and integrate plot data into their
290 analyses. Data is imported directly into R, avoiding the need to store data in numerous csv
291 files. Convenient data formatting supports seamless integration with a range of ecological,
292 statistical and graphical R packages with repeatable, self-contained script workflow (e.g.,
293 Manion et al. 2017; Oksanen et al. 2017), making it a useful exploratory dataset for
294 vegetation scientists worldwide. For example, *ausplotsR* and TERN plot data have been used
295 to determine what factors affect the biosynthetic domain composition of secondary
296 metabolites encoded by soil bacteria (Lemetre et al. 2017), and to validate remotely sensed
297 estimates of forest cover in dryland biomes globally (Bastin et al. 2017). Because *ausplotsR*
298 is embedded in R's software environment, the TERN plot data extracted by the package can
299 easily be enriched with additional data (e.g. climatic, altitude or additional plot data). This
300 data integration can allow for further insights into the user's research questions.

301

302 In recent years, the package has seen significant uptake from users. At the time of writing, we
303 have served over three thousand requests for data to over 260 users since early formats of the
304 package were released in 2018 (Figure 5). To date, these users have downloaded 890,000

305 sites worth of data and 800 million total records, including all raw vegetation data. We expect
306 *ausplotsR* will continue to enable ecological research on Australian ecosystems and enhance
307 opportunities for vegetation modelling and aggregation of data internationally.



308

309 Figure 5. Cumulative number of unique users who extracted TERN Ecosystem Surveillance
310 data via *ausplotsR* over two years.

311

312

313 **Concluding remarks**

314 For over a decade, TERN has provided comprehensive and research-ready data on Australia's

315 ecosystems. With *ausplotsR*, it is easier than ever for scientists to access Australia's only

316 database of standardised terrestrial ecosystem measurements. Data not provided through

317 *ausplotsR*, such as photo-panoramas, can be requested from TERN (www.tern.org.au). TERN

318 continues to expand its plot network across Australia and enhance the database with new

319 information and resources. `ausplots` R will be updated as data becomes available. New and
320 improved tools will also be added to the package in subsequent versions.

321

322 **Acknowledgements**

323 We acknowledge the TERN Ecosystem Surveillance field team and volunteers who work
324 tirelessly to collect TERN plot data.

325

326 **Author Contributions:**

327 A.T. and G.G. conceived the package; A.T., G.G., T.S., S.M., I.M.F. and B.B.M. wrote R
328 functions and documentation; B.S. collected data; S.M. drafted the paper. All authors
329 contributed to the paper and developed the package and dataset.

330

331 **References**

- 332 Baselga, A. & Leprieur, F. 2015. Comparing methods to separate components of beta
333 diversity. *Methods in Ecology and Evolution* 6: 1069-1079.
- 334 Bastin, J.-F., Berrahmouni, N., Grainger, A., Maniatis, D., Mollicone, D., Moore, R., . . .
335 Castro, R. 2017. The extent of forest in dryland biomes. *Science* 356: 635-638.
- 336 Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S.M., . .
337 . Jansen, F. 2018. Global trait–environment relationships of plant communities.
338 *Nature Ecology & Evolution* 2: 1906-1917.
- 339 Church, R.L., Stoms, D.M. & Davis, F.W. 1996. Reserve selection as a maximal covering
340 location problem. *Biological Conservation* 76: 105-112.
- 341 Chytrý, M., Chiarucci, A., Pärtel, M., Pillar, V.D., Bakker, J.P., Mucina, L., . . . White, P.S.
342 2019. Progress in vegetation science: Trends over the past three decades and new
343 horizons. *Journal of Vegetation Science* 30: 1-4.

344 Crisp, M.D., Laffan, S., Linder, H.P. & Monro, A. 2001. Endemism in the Australian flora.
345 *Journal of Biogeography* 28: 183-198.

346 Falster, D., Gallagher, R., Wenk, E., Wright, I., Indiarso, D., Baxter, C., . . . Groom, P.K. in
347 review. Austraits – a curated plant trait database for the Australian flora.

348 Guerin, G.R. & Lowe, A.J. 2015. ‘Sum of inverse range-sizes’ (SIR), a biodiversity metric
349 with many names and interpretations. *Biodiversity and conservation* 24: 2877-2882.

350 Guerin, G.R., Saleeba, T., Munroe, S., Blanco-Martin, B., Martín-Forés, I. & Tokmakoff, A.
351 2020. ausplotsR: TERN AusPlots analysis package. R package version 1.2.

352 Jiménez-Alfaro, B., Suárez-Seoane, S., Chytrý, M., Hennekens, S.M., Willner, W., Hájek,
353 M., . . . Tsiripidis, I. 2018. Modelling the distribution and compositional variation of
354 plant communities at the continental scale. *Diversity and Distributions* 24: 978-990.

355 Kao, R.H., Gibson, C.M., Gallery, R.E., Meier, C.L., Barnett, D.T., Docherty, K.M., . . .
356 Schimel, D. 2012. NEON terrestrial field observations: designing continental-scale,
357 standardized sampling. *Ecosphere* 3: art115.

358 Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., . . . Wirth, C. 2020.
359 TRY plant trait database – enhanced coverage and open access. *Global Change*
360 *Biology* 26: 119-188.

361 Kissling, W.D., Ahumada, J.A., Bowser, A., Fernandez, M., Fernández, N., García, E.A., . . .
362 Hardisty, A.R. 2018. Building essential biodiversity variables (EBVs) of species
363 distribution and abundance at a global scale. *Biological Reviews* 93: 600-625.

364 Kubiszewski, I., Costanza, R., Anderson, S. & Sutton, P. 2020. The future value of ecosystem
365 services: Global scenarios and national implications. In: *Environmental Assessments*.
366 Edward Elgar Publishing.

367 Lemetre, C., Maniko, J., Charlop-Powers, Z., Sparrow, B., Lowe, A.J. & Brady, S.F. 2017.
368 Bacterial natural product biosynthetic domain composition in soil correlates with

369 changes in latitude on a continent-wide scale. *Proceedings of the National Academy*
370 *of Sciences* 114: 11615-11620.

371 Maitner, B.S., Boyle, B., Casler, N., Condit, R., Donoghue II, J., Durán, S.M., . . . Enquist,
372 B.J. 2018. The bien r package: A tool to access the Botanical Information and
373 Ecology Network (BIEN) database. *Methods in Ecology and Evolution* 9: 373-379.

374 Manion, G., Lisk, M., Ferrier, S., Nieto-Lugilde, D., Mokany, K. & Fitzpatrick, M.C. 2017.
375 gdm: Generalized Dissimilarity Modeling. R package version 1.3.

376 Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., . . . Wagner,
377 H. 2017. vegan: Community Ecology Package. R package version 2.6.

378 Pyšek, P., Hulme, P.E., Simberloff, D., Bacher, S., Blackburn, T.M., Carlton, J.T., . . .
379 Genovesi, P. 2020. Scientists' warning on invasive alien species. *Biological Reviews*
380 95: 1511-1534.

381 R Core Development Team 2020. R: A language and environment for statistical computing.
382 In. R Foundation for Statistical Computing, Vienna, Austria.

383 Šímová, I., Violle, C., Svenning, J.-C., Kattge, J., Engemann, K., Sandel, B., . . . Enquist, B.J.
384 2018. Spatial patterns and climate relationships of major plant traits in the New World
385 differ between woody and herbaceous species. *Journal of Biogeography* 45: 895-916.

386 Sparrow, B.D., Edwards, W., Munroe, S.E.M., Wardle, G.M., Guerin, G.R., Bastin, J.-F., . . .
387 Lowe, A.J. 2020a. Effective ecosystem monitoring requires a multi-scaled approach.
388 *Biological Reviews* 95: 1706-1719.

389 Sparrow, B.D., Foulkes, J.N., Wardle, G.M., Leitch, E.J., Caddy-Retalic, S., van Leeuwen,
390 S.J., . . . Lowe, A.J. 2020b. A vegetation and soil survey method for surveillance
391 monitoring of rangeland environments. *Frontiers in Ecology and Evolution* 8.

392 Thackway, R. & Cresswell, I. 1995 (Eds). An Interim Biogeographic Regionalisation for
393 Australia: a framework for establishing the national system of reserves, Version 4.0.
394 Australian Nature Conservation Agency, Canberra.

395 Tokmakoff, A., Sparrow, B., Turner, D. & Lowe, A. 2016. AusPlots Rangelands field data
396 collection and publication: Infrastructure for ecological monitoring. *Future*
397 *Generation Computer Systems* 56: 537-549.

398 Velazco, S.J.E., Galvão, F., Villalobos, F. & De Marco Júnior, P. 2017. Using worldwide
399 edaphic data to model plant species niches: An assessment at a continental extent.
400 *PLOS ONE* 12: e0186025.

401
402

DRAFT