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

Samantha Munroe[1,2], Greg Guerin[1,2], Tom Saleeba[1,2], Irene Martín-Forés[1,2], Bernardo Blanco-Martin[3], Ben Sparrow[1,2], and Andrew Tokmakoff[1,2]

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

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

[3]Department of Agriculture and Fisheries, Queensland Government, Australia

Correspondence:

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

Email: samantha.munroe@adelaide.edu.au

Funding Information:

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

Abstract

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

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

Introduction

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

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

TERN’s Ecosystem Surveillance platform performs standardised field surveys and sampling across a national plot network. This platform collects crucial data for effective monitoring of Australia’s ecosystems, including vegetation structure and composition, soil characteristics, and soil and plant samples (Sparrow et al. 2020b). TERN has established over 750 plots across every major terrestrial environment in Australia (Figure 1), recoding >5,000 unique species and collecting >65,000 plant and soil samples for analysis (Table 1).

The *ausplotsR* package was designed for the R statistical computing environment (R Core Development Team 2020) to provide free and simple access to the TERN Ecosystem Surveillance plot survey database (Guerin et al. 2020). The package has a straightforward workflow to enable a range of vegetation analyses. First, easy-to-understand functions extract raw survey data for all plots within the network from a regularly updated database. Second, raw data can be incorporated into downstream functions that calculate a variety of metrics, such as species cover, fractional cover, and basal area. These functions distinguish *ausplotsR* from other vegetation databases because they provide a fast and reliable way to calculate some of the most common metrics in terrestrial research. *ausplotsR* also provides information on TERN's substantial soil and plant sample library, such as voucher numbers and the date and location samples were collected. Data are formatted to support compatibility with global (e.g. Kattge et al. 2020) and Australian plant trait datasets (e.g. Falster et al. in review) and can be integrated with additional TERN data products. The goal of *ausplotsR* is to facilitate quality ecosystem research and effective land management across Australia through delivery and pre-processing of field data. Instant access to continental-scale plant community data in the R environment provides a valuable resource to vegetation scientists and modellers for testing ecological ideas, tools, and methods.

TERN plot survey methods

TERN plots are surveyed using the 'AusPlots Rangelands' method, composed of modules to collect vegetation and soil data (Table 2; also see Sparrow et al. 2020b). Here we provide a short overview of the modules used to collect data available through *ausplotsR*. One-hectare plots are established in a homogenous area of terrestrial vegetation. Vegetation structure and composition are measured using the point-intercept module. Transects (10 x 100 m) are laid out in a grid pattern spaced 20 m apart. Species identity, growth form, height, and systematic absences (e.g. bare ground) are recorded at 1 m points along transects, resulting in 1010 survey points. Vouchers of each species are collected and sent to herbaria for identification.

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

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

The ausplotsR package

The *ausplotsR* package (CRAN: https://CRAN.R-project.org/package=ausplotsR; latest development version and patches: https://github.com/ternaustralia/ausplotsR) provides access to most data modules. *ausplotsR* sources its data via a dedicated server stack running inside a Virtual Machine in the Australian Research Data Common’s (ARDC’s) NECTAR research cloud. The stack contains a number of services which: (1) expose the published AusPlots data for external use, (2) collect usage statistics, (3) Reverse-Proxy the API for scalability and provides threat protection, and (4) authenticate/authorise users for fine-grained data access control. This open source stack is available at: https://github.com/ternandsparrow/swarm-rest. *ausplotsR* functions enable users to quickly read and prepare survey data for ecological analysis. Next, we review key *ausplotsR* functions to extract, visualise, and analyse plot data (Table 3).

Data extraction

Data are extracted using the function *get_ausplots* (Table 4). By default, the *get_ausplots* function extracts a list of data tables for the point-intercept (*$veg.PI*) and vegetation voucher (*$veg.vouchers*) modules for all surveys. It also provides a site information table (*$site.info*) which describes survey details and environmental features. Using these data tables, TERN’s specimen collection can be cross-referenced against site information and vegetation metrics.

Arguments of ‘*get_ausplots’* allow users to select data for individual plots (*my.Plot_IDs*), rectangular spatial coordinates (*bounding_box*), plant families (*family_search*), or species (see *Data Format*).

```
# for selected plots:
my.data <- get_ausplots(my.Plot_IDs=c("SATFLB0004",
"QDAMGD0022"))

# plots within a geographic area:
my.data <- get_ausplots(bounding_box= c(120, 140, -30, -10)

# plots where "Myrtaceae" were recorded:
Myrtaceae <- get_ausplots(family_search="Myrtaceae")
```

Users can request data not extracted by default, like basal wedge data (*basal.wedge=TRUE*) or soil characteristics (*soil_character=TRUE*). Once extracted, data tables can be incorporated into various downstream processing functions that calculate vegetation indices (Figure 2).

Data format

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* (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 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 visits.

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

Alternatively, species identifications can be taken from the *standardised_name* field. *standardised_name* values are based on *herbarium_determination* values standardised to match the most widely accepted synonym according to ‘World Flora Online’ (www.worldfloraonline.org). This ensures scientific names will not differ between States/Territories and increases consistency with global databases. It also excludes non-standard entries like ‘dead’ identifications. The remaining taxonomic fields are derived from the *standardised_name*.

Plot data extractions can be filtered by *herbarium_detemination* or *standardised_name*:

```
#plot and voucher records where "Eucalyptus moderata" was
identified:

Eucalyptus_moderata <-
get_ausplots(herbarium_determination_search="Eucalyptus
moderata")
```

Data processing and calculation of vegetation indices

1. Community composition matrices

The *species_table* function takes the data table of individual point-intercept hits (*$veg.PI*) and returns species occurrence matrices. *species_table* can calculate species presence/absence, percent cover, frequency (based on occurrences on different transects), or Importance Value Index. Users can select the preferred species name option from *standardised_name* , *herbarium_determination*, or *genus_species*. For example, to compare species level cover:

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

Details on how vegetation indices are calculated are in the help manual.

2. Vegetation cover:

Several functions in *ausplotsR* calculate different aspects of vegetation (and ground) cover based on raw point-intercept input data (*$veg.PI*). *growth_form_table* generates occurrence matrices for plant growth forms based on presence/absence, cover, or species richness:

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

Other functions calculate fractional cover (i.e. proportional cover of living vegetation, dead litter and bare ground; *fractional_cover*; Figure 1), cover of individual growth forms subset by height (*single_cover_value*), and the cross-sectional area of tree trunks measured through the basal wedge module (*basal_area*). The relative cover of species, plant growth forms, and fractional cover can be illustrated per plot using the function *ausplots_visual* (Figure 3).

3. Optimising species accumulation

The *optim_species* function applies different biodiversity metrics as optimisers to select a subset of plots that maximise species accumulation. The framework underlying this function is the ‘Maximal coverage problem’, applied in conservation biology to design reserves (Church et al. 1996). The function identifies plots that will ‘protect’ the maximum number of species in a limited number of sites. The biodiversity metrics included are species richness (biodiversity hotspots), range rarity richness (high biodiversity and uniqueness; Guerin and Lowe 2015), corrected weighted endemism (areas with range-restricted endemic species; Crisp et al. 2001), Shannon-Wiener diversity index and the Simpson diversity index (which include species relative abundances), and the Simpson dissimilarity (maximises species turnover; Baselga & Leprieur 2015). Users specify the number of plots to select, and each optimiser selects the subset of plots that accumulates the largest number of species. To do this, biodiversity metrics are applied only to species occurrences within the dataset. *optim_species* can also visualise the optimisations (Figure 4). The input data is a species versus sites matrix; thus, it can be used with the data generated with *species_table* or with any other vegetation dataset.

Data Licence and authorship guidelines

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

Applications and uptake

ausplotsR makes it easy for scientists to quickly access and integrate plot data into their analyses. Data is imported directly into R, avoiding the need to gain permission to access data or store data in numerous csv files. Convenient data formatting supports seamless integration with a range of ecological, statistical, and graphical R packages with repeatable, self-contained script workflow (e.g., Manion et al. 2017; Oksanen et al. 2017), making it a useful exploratory dataset for vegetation scientists. For example, *ausplotsR* and TERN plot data have been used to determine what factors affect the biosynthetic domain composition of secondary metabolites encoded by soil bacteria (Lemetre et al. 2017), and to validate remotely sensed estimates of forest cover in dryland biomes (Bastin et al. 2017). Other examples of data applications are described in Sparrow et al. (2020b). Because *ausplotsR* is embedded in R's software environment, TERN plot data can easily be enriched with additional data (e.g. climatic or altitude).

In recent years, the package has seen significant uptake from users. At the time of writing, we have served over ten thousand requests for data to over 350 users since early formats of the package were released in 2018 (Figure 5). These users have downloaded 5,000,000 sites of data and 1.2 billion total records. We expect *ausplotsR* will continue to enable ecological research on Australian ecosystems and enhance opportunities for vegetation modelling internationally.

Importantly, the 'AusPlots Rangelands' method and TERN data model (i.e. the way data items are connected and modelled, https://linkeddata.tern.org.au) could be adapted by users to create their own data management system. Similarly, *ausplotsR* code can be extracted to calculate vegetation indices for unaffiliated datasets. The function *optim_species*, can already incorporate non-TERN data.

Concluding remarks

For over a decade, TERN has provided comprehensive and research-ready data on Australia's ecosystems. With *ausplotsR*, it is easier than ever for scientists to access Australia's only database of standardised terrestrial ecosystem measurements. *ausplotsR* provides unfettered access to a broad and high-quality dataset combining information on different data streams and metrics in a user-friendly format. Data not provided through *ausplotsR*, such as photo-panoramas and physical samples, can be requested from TERN (www.tern.org.au). TERN continues to expand its plot network across Australia and enhance the database with new information and resources. *ausplotsR* will be updated as data becomes available. New and improved tools will also be added to the package in subsequent versions.

Acknowledgements

We acknowledge the TERN Ecosystem Surveillance field team and volunteers who work tirelessly to collect TERN plot data. We also acknowledge the support of TERN by the National Collaborative Research Infrastructure Strategy.

Author Contributions:

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 functions and documentation; B.S. collected data; S.M. drafted the paper. All authors contributed to the paper and developed the package and dataset.

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study

References

Baselga, A. & Leprieur, F. 2015. Comparing methods to separate components of beta diversity. *Methods in Ecology and Evolution* 6: 1069-1079.

Bastin, J.-F., Berrahmouni, N., Grainger, A., Maniatis, D., Mollicone, D., Moore, R., . . . Castro, R. 2017. The extent of forest in dryland biomes. *Science* 356: 635-638.

Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S.M., . . . Jansen, F. 2018. Global trait–environment relationships of plant communities. *Nature Ecology & Evolution* 2: 1906-1917.

Church, R.L., Stoms, D.M. & Davis, F.W. 1996. Reserve selection as a maximal covering location problem. *Biological Conservation* 76: 105-112.

Chytrý, M., Chiarucci, A., Pärtel, M., Pillar, V.D., Bakker, J.P., Mucina, L., . . . White, P.S. 2019. Progress in vegetation science: Trends over the past three decades and new horizons. *Journal of Vegetation Science* 30: 1-4.

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

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

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

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

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

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

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

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

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

Lemetre, C., Maniko, J., Charlop-Powers, Z., Sparrow, B., Lowe, A.J. & Brady, S.F. 2017. Bacterial natural product biosynthetic domain composition in soil correlates with changes in latitude on a continent-wide scale. *Proceedings of the National Academy of Sciences* 114: 11615-11620.

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

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

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

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

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

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

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

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

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

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

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

Tables

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

Data Type	**Count**
Number of plots surveyed	667
Number of plots revisited	106
Number of plant voucher specimens collected	41141
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 available	17,082

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 ausplotsR?**
Plot layout	Accurate layout using DGPS; installation of permanent markers.	Accurate relocation; remote sensing validation	Yes
Vegetation			
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
Soils and Landscapes			
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

Table 3. Descriptions of *ausplotsR* functions

Function	**Description**
ausplots_visual	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 *get_ausplots* and other pre-processing functions)
basal_area	Calculates basal area (or number of basal wedge hits) for each plot, using the raw basal wedge data returned from *get_ausplots*
fractional_cover	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 *get_ausplots*)
get_ausplots	This function is the starting point for accessing data through the ausplotsR package. It extracts the data
growth_form_table	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 *get_ausplots* function
optim_species	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
plot_opt	This function plots different species accumulation curves obtained through different optimisers in the optim_*species* function
single_cover_value	Calculates a single vegetation cover value per site based on the plot-based point-intercept data generated by get_*ausplots*. Cover can be subsetted to vegetation that has a specified minimum or maximum height and/or by plant growth forms
species_list	Nice species lists based on the vegetation voucher module of Ausplots generated using the *get_ausplots* function
species_table	This function takes a data frame of individual raw point-intercept hits generated using the *get_ausplots* function, and generates species occurrence matrices as desired based on presence/absence, cover, frequency or IVI

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

** Extracted by default*

Figures

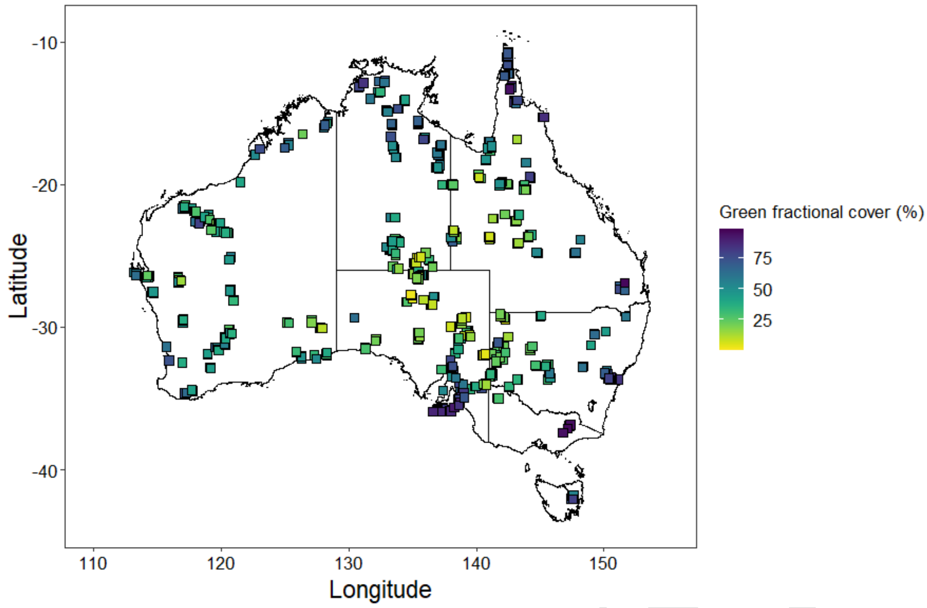

Figure 1. TERN AusPlots monitoring plot locations as generated within the package coded by percent fractional green cover.

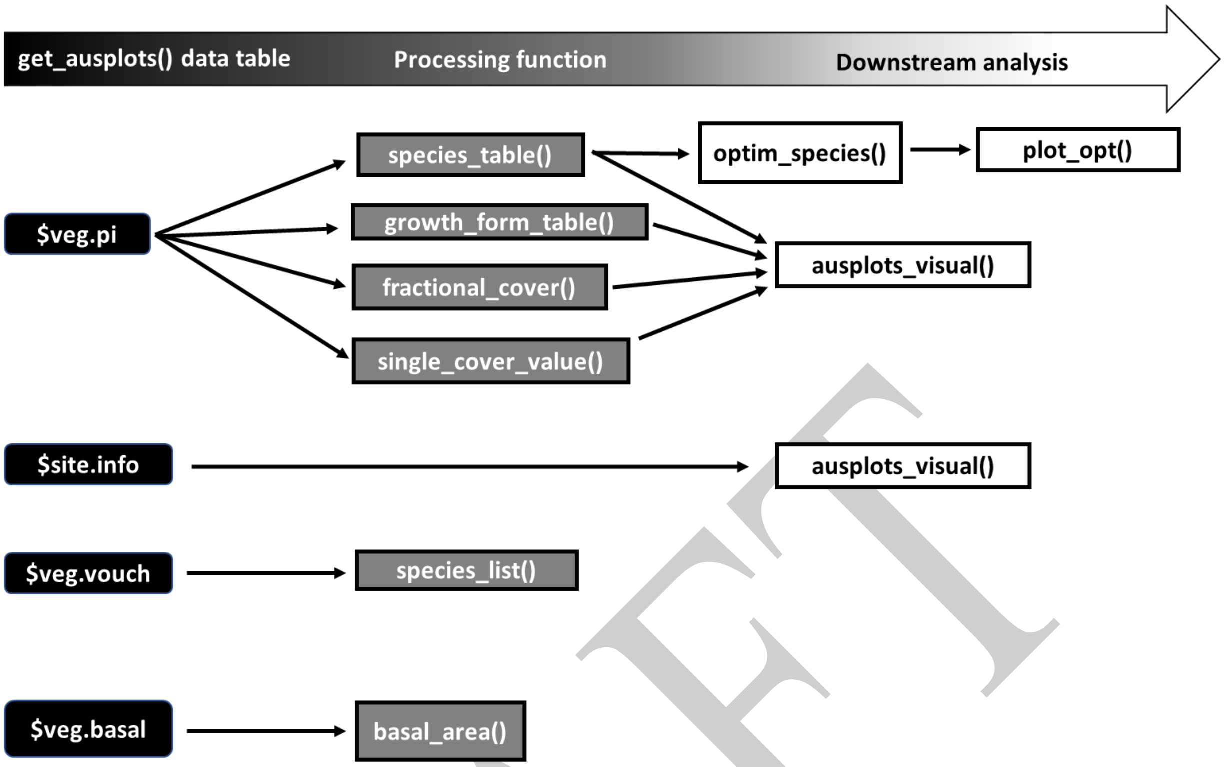

Figure 2. Workflow of the *ausplotsR* package, demonstrating how raw data tables generated from *get_ausplots()* can be incorporated in processing and downstream functions

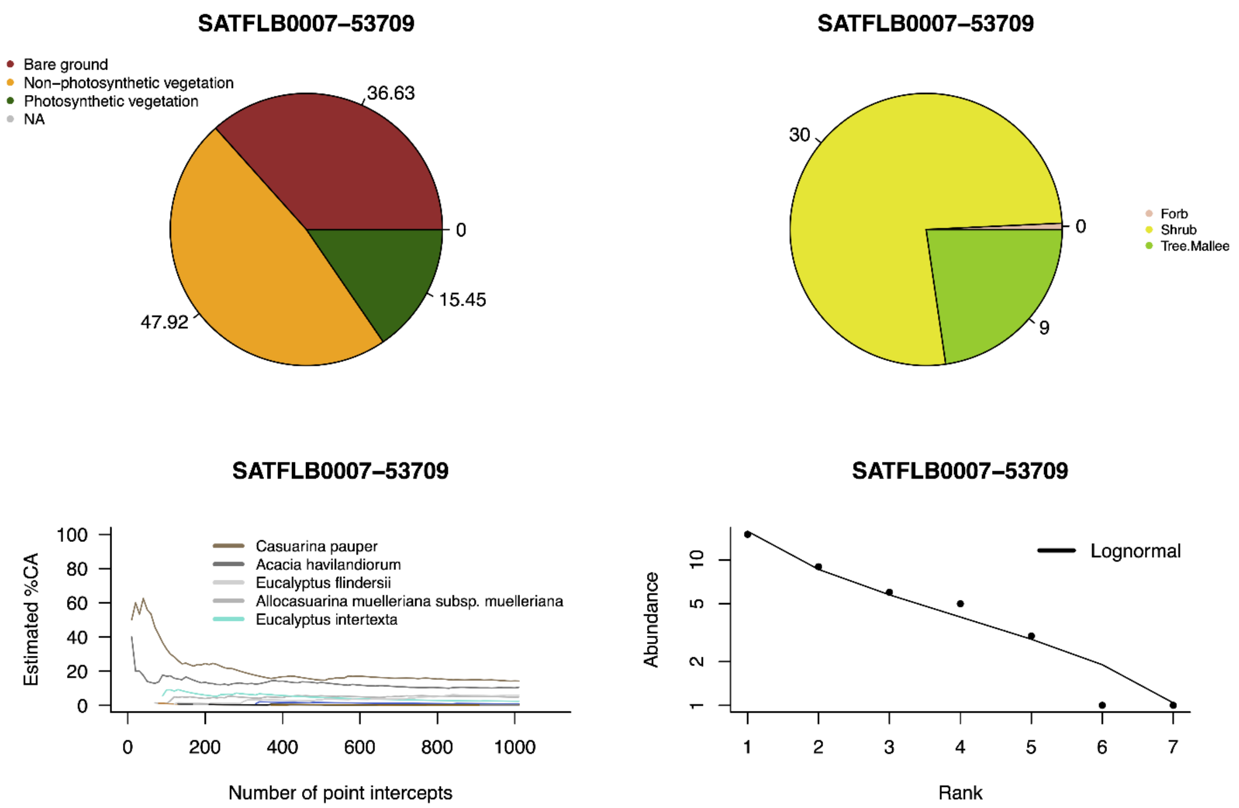

Figure 3. Example of data presentation panels for a single plot visit using *ausplots_visual*(): a) fractional vegetation cover; b) the relative abundance of plant growth forms; c) cumulative estimates of percent cover by species as point-intercept hits are taken across the plot (Guerin et al. 2017);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

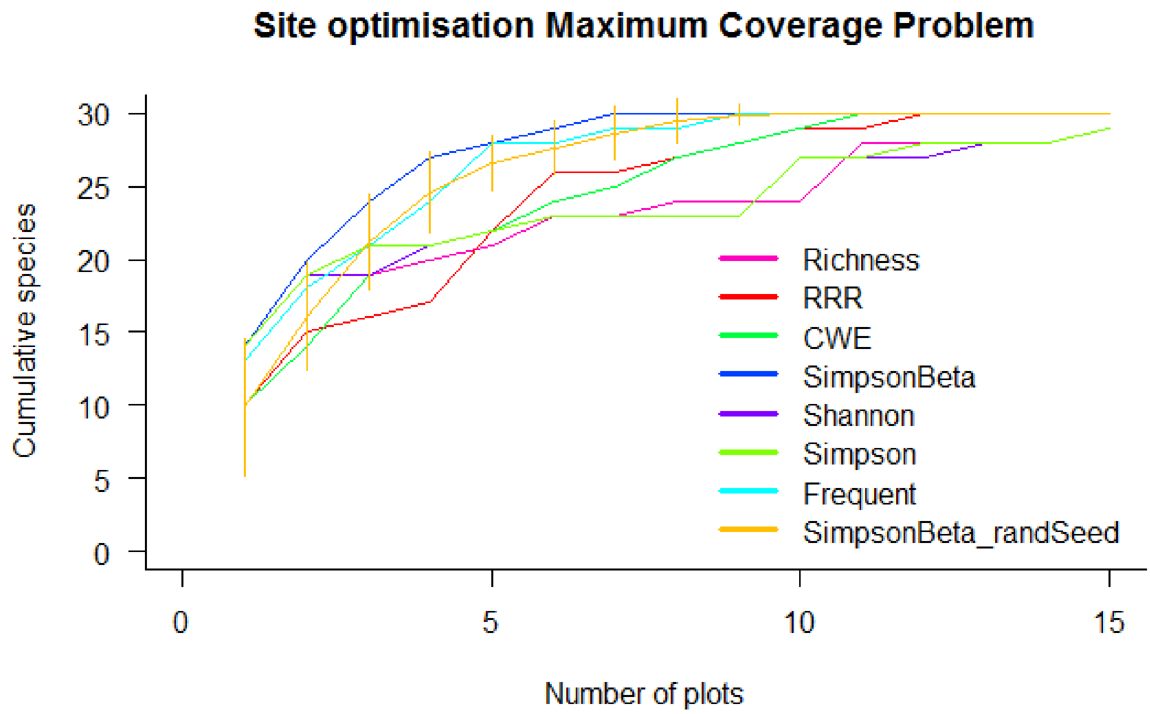

Figure 4. Species accumulation plot comparing different biodiversity metrics included in *optim_species*. The image has been created using the *dune* dataset from the *vegan* package as an input, selecting 15 plots, and 60 iterations to calculate the most frequently selected ones based on random starts.

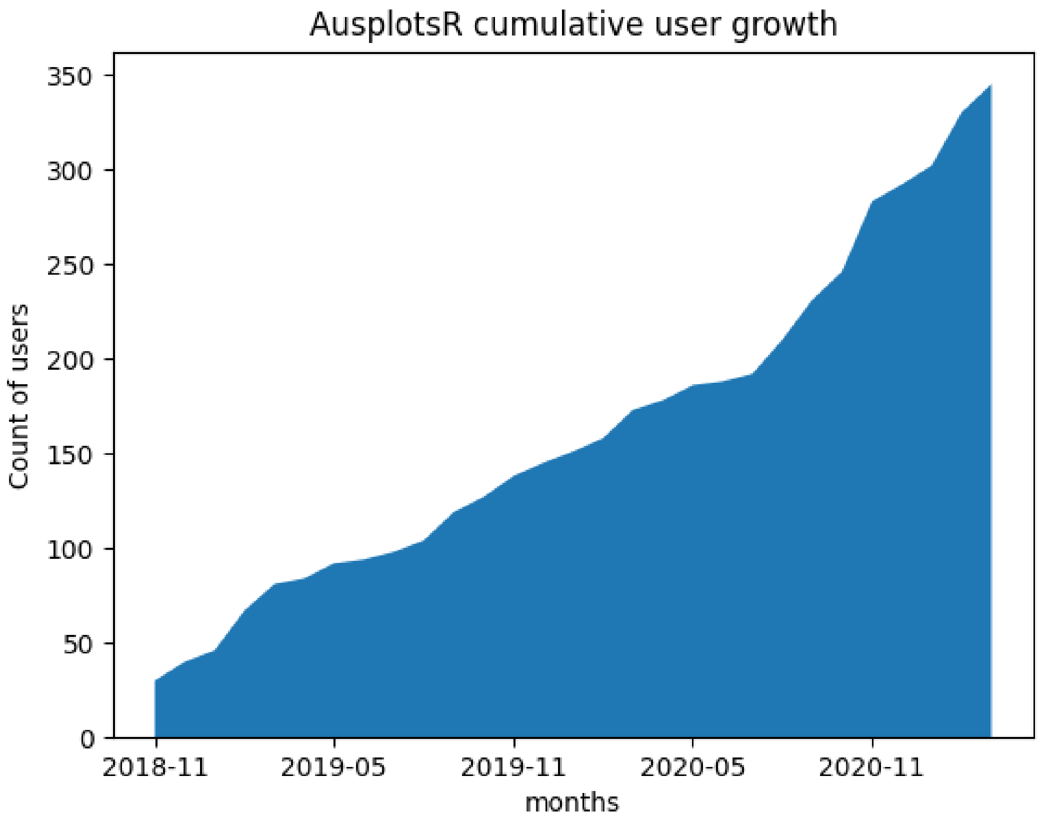

Figure 5. Cumulative number of unique users who extracted TERN Ecosystem Surveillance data via *ausplotsR* over two years.