Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Parcelas-CL: a community-sourced vegetation dataset for multi-scale biodiversity monitoring in Chile

José Miguel Cerda-Paredes

jose.cerda@dataobservatory.net

Data Observatory Foundation

Alejandro Miranda

Universidad de La Frontera

Adison Altamirano

Universidad de la Frontera

Pablo Ignacio Becerra

Pontificia Universidad Católica de Chile

Cynnamon Dobbs

Universidad de Los Andes, Chile

Mauricio Galleguillos

Pontificia Universidad Católica de Chile

Marcelo D. Miranda

Pontificia Universidad Católica de Chile

Narkis S. Morales

Bioeconomy Science Institute

Juan F. Ovalle

Universidad de Chile

Alejandro Venegas-González

Universidad de O'Higgins

Carlos Zamorano-Elgueta

Universidad de Aysén

Dylan Craven

Universidad Mayor

Javier Lopatin

Adolfo Ibáñez University

Short Report

Keywords: Vegetation plots, Spatial ecology, Data integration, FAIR, CARE, Geographic biases

Posted Date: June 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9986019/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Parcelas-CL: a community-sourced vegetation dataset for multi-scale biodiversity monitoring in Chile

José M. Cerda-Paredes[1,2*], Alejandro Miranda[3,4], Adison Altamirano[5], Pablo I. Becerra[6], Cynnamon Dobbs[7], Mauricio Galleguillos[4,6,8], Marcelo D. Miranda[9], Narkis S. Morales[10], Juan F. Ovalle[11], Alejandro Venegas-González[12], Carlos Zamorano-Elgueta[13], Dylan Craven[1,14], Javier Lopatin[1,2,4]

[1]Data Observatory, Santiago, Chile.

[2]Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Santiago, Chile.

[3]Laboratorio de Ecología del Paisaje y Conservación, Departamento de Ciencias Forestales, Universidad de La Frontera

[4]Center for Climate and Resilience Research, Universidad de Chile, Santiago, Chile.

[5]Universidad de La Frontera, Facultad de Ciencias Agropecuarias y Medioambiente. Laboratorio de Ecología del Paisaje y Conservación, Temuco, Chile.

[6]Departamento de Ecosistemas y Medio Ambiente, Facultad de Agronomía y Sistemas Naturales, Pontificia Universidad Católica de Chile, Santiago, Chile.

[7]Centro de Estudios Territoriales, Universidad de Los Andes, Santiago, Chile.

[8]Departamento de Ingeniería Hidráulica y Medio Ambiente, Pontificia Universidad Católica de Chile, Santiago Chile

[9]Facultad de Agronomía y Sistemas Naturales, Centro de Cambio Global UC, Pontificia Universidad Católica de Chile, Santiago, Chile.

[10]Bioeconomy Science Institute, Lincoln, New Zealand.

[11]Departamento de Silvicultura y Conservación de la Naturaleza, Universidad de Chile, Santiago, Chile.

[12]Instituto de Ciencias Agroalimentarias, Animales y Ambientales (ICA3), Universidad de O'Higgins, San

Fernando, Chile.

[13]Departamento de Ciencias Naturales y Tecnología, Universidad de Aysén, Coyhaique, Chile.

[14]GEMA Center for Genomics, Ecology & Environment, Universidad Mayor, Santiago, Chile.

[*]Corresponding Author: José Miguel Cerda-Paredes; jose.cerda@dataobservatory.net/+56977646141

ORCID

José M. Cerda-Paredes (0000-0002-0087-5494)

Alejandro Miranda (0000-0002-8211-0179)

Pablo I. Becerra (0000-0002-2205-4478)

Cynnamon Dobbs (0000-0001-9845-6660)

Mauricio Galleguillos (0000-0003-4617-0980)

Marcelo D. Miranda (0000-0002-8386-4888)

Narkis S. Morales (0000-0002-3942-1916)

Juan F. Ovalle (0000-0002-8487-3487)

Alejandro Venegas-González (0000-0003-4568-4533)

Carlos Zamorano-Elgueta (0000-0002-5588-3248)

Dylan Craven (0000-0003-3940-833X)

Javier Lopatin (0000-0002-5540-7428)

ABSTRACT

CONTEXT

Geographic biases and fragmented data limit the integration of ecological data across spatial scales, particularly in the Global South. In Mediterranean ecosystems, vegetation data are dispersed across independent studies, constraining landscape-level biodiversity analyses.

OBJECTIVES

We aimed to (1) describe the structure, sources, coverage and standardization of the Parcelas-CL database, including its main limitations and sources of uncertainty; (2) evaluate its spatial representativeness and potential for landscape-scale analyses.

METHODS

We compiled 1,485 georeferenced vegetation plots across south-central Chile and named Parcelas-CL (*parcelas* means *plots* in Spanish). We harmonized data through taxonomic standardization, spatial alignment, and classification of ecological units based on vegetation formation. We evaluated sampling coverage and

representativeness across geographic regions, ecological units, and species richness patterns.

RESULTS

Parcelas-CL spans a broad latitudinal range but shows uneven spatial distribution, with a sampling bias in central Chile between 1976 and 2026. Most plots include abundance data (75%), while others are limited to presence-absence records. Sampling efforts are concentrated on four well-represented vegetation formations, whereas other formations are underrepresented. Parcelas-CL includes 676 woody plant species, with species richness patterns strongly associated with sampling intensity. Species accumulation curves indicate substantial coverage of the regional species pool, but do not reach an asymptote, suggesting further sampling is needed.

CONCLUSIONS

This harmonized dataset enables cross-scale biodiversity analyses and supports integration with spatial analyses. However, uneven sampling highlights the need to expand coverage across underrepresented regions and vegetation formations. Strengthening data standardization and collaborative data sharing will improve biodiversity monitoring.

KEYWORDS

Vegetation plots; Spatial ecology; Data integration; FAIR; CARE; Geographic biases

INTRODUCTION

Despite increasing global efforts to compile ecological data, geographic biases persist, with the Global South remaining underrepresented in the production and representation of scientific knowledge (Beck et al., 2014; Wiegleb & Bruns, 2025). Biodiversity data are disproportionately concentrated in ten countries, with ~80% of GBIF records, largely in the Global North (Hughes et al., 2021). Compiling ecological data is difficult, as it is often dispersed across independent studies, institutions, and monitoring projects (Farley et al., 2018), and is frequently collected using heterogeneous sampling designs, stored in multiple –often proprietary– digital formats, and non-English languages, primarily Spanish (Heberling et al., 2021). This is a global challenge, but in some regions (e.g., South America), data mobilization efforts remain limited. Together, these factors lead to fragmentated ecological information, which limit interoperability and comparative analyses across sites and scales, ultimately hindering the development of integrative approaches in landscape ecology (Griffith et al., 2026).

Drylands host multiple biodiversity hotspots (Myers et al. 2000) characterized by high endemism and pronounced spatial heterogeneity (Maestre et al., 2021). Sclerophyllous forests are one of the most common vegetation types in this eco-region, occupying large portions of the landscape and hosting a significant fraction of its biodiversity (Berdugo et al., 2022). However, available vegetation data from this region remain dispersed and are often not harmonized or publicly accessible, which limits their use in large-scale analyses and coordinated monitoring initiatives (Gonzalez et al., 2026). Currently, the Chilean hotspot faces unprecedented threats, including a decades-long megadrought that triggered a synchronous widespread browning of the vegetation in 2019 (Miranda et al., 2020; Venegas-González et al., 2023; Fuentes et al., 2025). Synthesizing existing data thus represents a crucial step for enabling spatially explicit analyses of species diversity and composition and vegetation structure, and for improving biodiversity monitoring and assessment. .

A key challenge in landscape ecology is integrating fine-scale field observations with broader spatial representations of biodiversity. Although plot-based vegetation data are among the most reliable sources of information on species composition and community structure (Chytrý et al., 2015; Bruelheide et al., 2019), their utility at landscape scales depends on data availability, spatial coverage, consistency, and interoperability. Combining georeferenced plots with standardized taxonomic information and an explicit characterization of sampling resolution is therefore essential for enabling cross-scale analyses (Sabatini et al., 2022; van Tiel et al., 2024).

Recent advances in spatial technologies, such as remote sensing and geospatial modeling, have created new opportunities to analyze biodiversity patterns over large spatial extents (e.g., mapping of vegetation structure, estimating species diversity, and analyzing spatial heterogeneity). However, these approaches require spatially explicit field datasets (i.e., taxonomically standardized, georeferenced, and subject to quality control procedures, including taxonomic validation and the detection of inconsistencies in species records and plot-level attributes) for calibration, validation and interpretation. The lack of taxonomically and spatially harmonized vegetation plot databases in many regions remains a major obstacle (Chytrý et al., 2015; Griffith et al., 2026) to advancing these integrative approaches (i.e., Wallacean shortfall; Hortal et al., 2014). Translate ecological data into standardized biodiversity indicators should aim to achieve technical interoperability in accordance with the FAIR principles, while also adopting a governance approach based on the CARE principles, ensuring ethical management and the

collective benefit of those who safeguard ecological knowledge in the region (Gudde et al., 2026; Jantzen & Vriend, 2026).

Publicly available and curated vegetation data from this region are still limited (but see Peyre et al. 2022; Seipel et al., 2007–2023), highlighting the need to improve not only data availability but also standardization and quality control for large-scale ecological analyses. Here, we present a harmonized database of vegetation plots for Mediterranean ecosystems in central Chile, with a particular focus on sclerophyllous forests. The database comprises 1,485 georeferenced plots and includes information on species composition, plot area, and geographic coordinates. This provides a consistent basis for spatially explicit analyses of plant communities. Our objectives were to: (1) describe the structure, sources, coverage and standardization of the dataset, including its main limitations and sources of uncertainty; (2) evaluate its spatial representativeness and potential for landscape-scale analyses; and (3) highlight its relevance as a resource for biodiversity monitoring and integration with spatial analytical approaches.

METHODS

Data compilation and scope

This study is based on a scientific community-sourced collaboration initiative launched in 2022. The initiative aimed to integrate vegetation plot data from sclerophyllous forests by inviting researchers to contribute vegetation plot data using a standardized template that included species composition, abundance, geographic coordinates, and a site name. We compiled a total of 1,264 vegetation plots from 12 contributors, concentrated mostly in south-central Chile, and 221 plots from sPlot (Bruelheide et al., 2019) (Fig. 1). While the first phase of this initiative is mainly focused on Mediterranean ecosystems, we plan on including data from other ecosystems across Chile in subsequent phases.

Data structure and variables

In addition to species composition, Parcelas-CL includes plot-level metadata that describes the sampling method and spatial location. Core variables include an unique plot observation identifier, the year of sampling, the data contributor, plot size (m^2), and geographic coordinates. To ensure spatial consistency across datasets, all spatial data were standardized to a common coordinate reference system (WGS84/UTM Zone 19S). All plots included in Parcelas-CL contained valid spatial coordinates, ensuring their suitability for spatially explicit analyses.

Data harmonization

We harmonized the contributed data to ensure consistency and interoperability while preserving the original information provided by contributors. Taxonomic standardization involved unifying species names into a consistent format, correcting typographical errors, and ensuring uniform naming conventions across datasets. We standardized species names using the Taxonomic Name Resolution Service (TNRS; Boyle et al., 2013) implemented in R. We performed name matching and validation against authoritative taxonomic databases, specifically the World Flora Online (WFO) and the World Checklist of Vascular Plants (WCVP), reducing

inconsistencies related to synonymy and spelling across datasets. We kept the original species name assigned by each researcher, the accepted species name, and taxonomic rank (e.g., genus, species, subspecies, etc.). Because the contributed datasets included multiple abundance metrics, we harmonized species abundances using a two-level structure. First, we defined a binary variable for each plot to indicate whether an abundance parameter was present (0 = Only presence/abundance; 1 = species' abundance data available). Second, we used an additional column to specify how abundance was quantified (e.g., number of individuals or percentage cover). This allows users to easily subset the data according to their analytical needs. We retained plot size because it is well known to influence species diversity estimates. However, not all contributed datasets included information on plot size. In these cases, we recorded the plot size as *NA* to explicitly indicate the missing information. We did not rescale or standardize plot area because preserving the original sampling resolution is essential to maintaining data integrity and supporting scale-dependent analyses.

To further standardize the ecological context of Parcelas-CL, we extracted the vegetation formations defined by Luebert and Pliscoff (2017) for each plot. These formations are based on vegetation physiognomy, dominant species, and bioclimatic conditions. They are also widely used as proxies for ecosystem types in Chile, although alternative classification frameworks exist (e.g., global ecosystem typologies (Keith et al., 2022), and biogeographic ecoregions (Dinerstein et al., 2017)). In this study, we used vegetation formations to define ecological units representing ecosystems, enabling consistent plot classification across Parcelas-CL.

Data Availability

Parcelas-CL is publicly available, should be referenced by citing this paper, and is located in a Zenodo repository (http://doi.org/10.5281/zenodo.20602096). We encourage users to report mistakes, make suggestions and contribute data.

DESCRIPTION OF THE DATA

Sampling Effort

The compiled database comprises 1,485 vegetation plots distributed across south-central Chile, spanning a broad latitudinal (from 30.25°S to 54.82°S) range that encompasses the majority of Mediterranean ecosystems. The sampling effort is spatially heterogeneous, with a strong concentration of plots in the Valparaíso, Metropolitana, and Maule regions (Fig. 1). We identified a clustered plot distribution with higher densities in accessible and historically studied areas, while other regions remain comparatively underrepresented (Fig. 1).

Plot area varied between 78.5 and 10000 m^2, with a median of 400 m^2 (± 38.7). We note a higher frequency of plots with an area of 100 m^2, and between 400 and 710 m^2 (Fig. 1). However, 10.8% of the plots did not include a spatial resolution (i.e., plot size = *NA*). Our database includes plots surveyed between 1976 and 2026 (Fig. 2a). Sampling effort is temporally uneven, with peaks in 1999, 2010, 2013, 2014, 2019 and 2022, each exceeding 100 plots, whereas most other years include substantially fewer observations (Fig. 2a).

In addition to spatial coverage, Parcelas-CL reflects the diversity of methods used to monitor woody vegetation

in the study region. 75% of the plots include abundance information, with vegetation cover being the most common reported measure (42.1% of plots; N = 625), followed by the number of individuals (32.9%; N = 489) (Fig. 2b). 25% of plots in the database only include species presence/absence information (N = 371) (Fig. 2b).

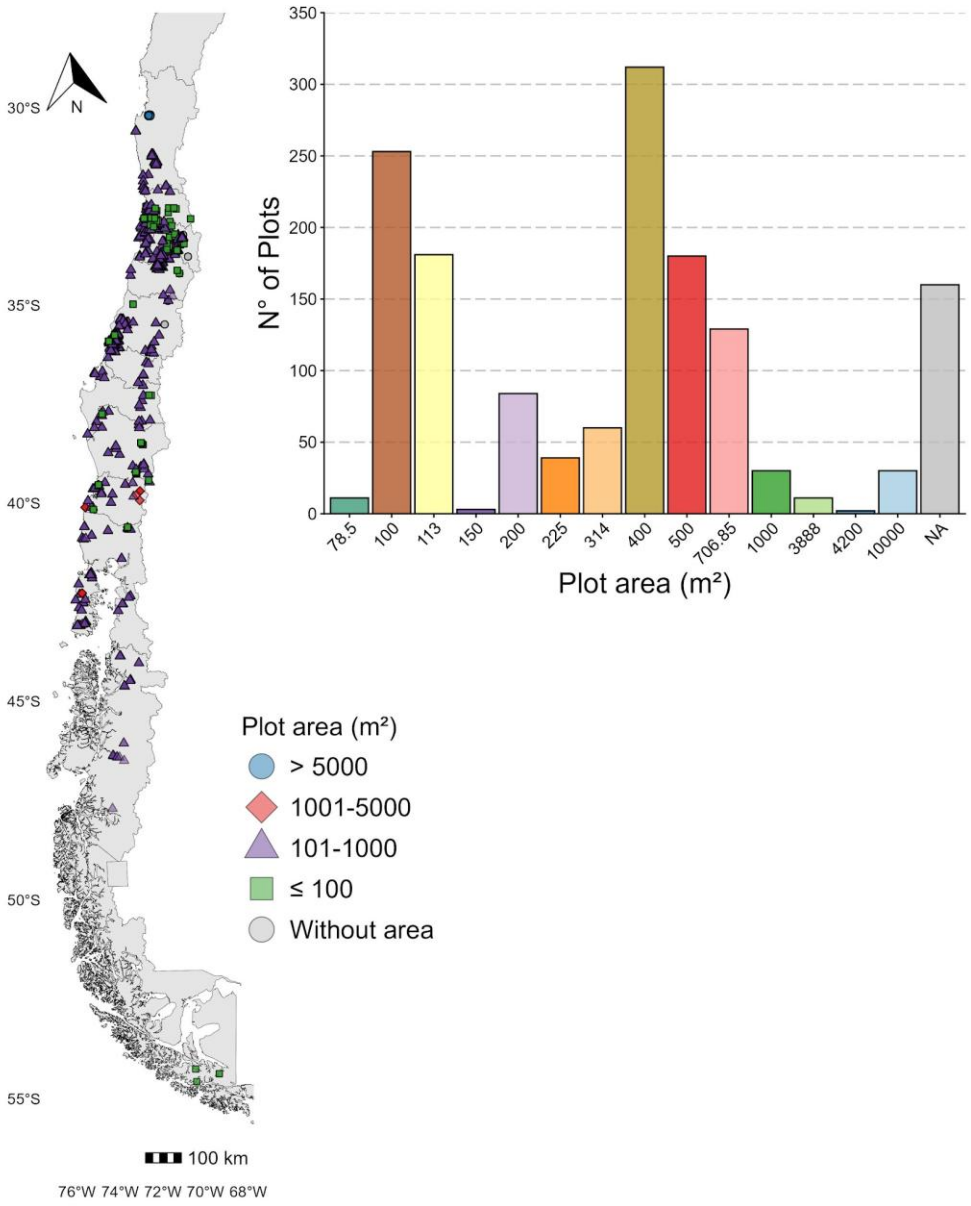


Fig. 1. Spatial distribution and sampling resolution of vegetation plots across Mediterranean-type ecosystems in Chile. Circles represent individual plots (n = 1,485), and color is the plot area (m^2). The bar chart shows the

frequency distribution of plot sizes.

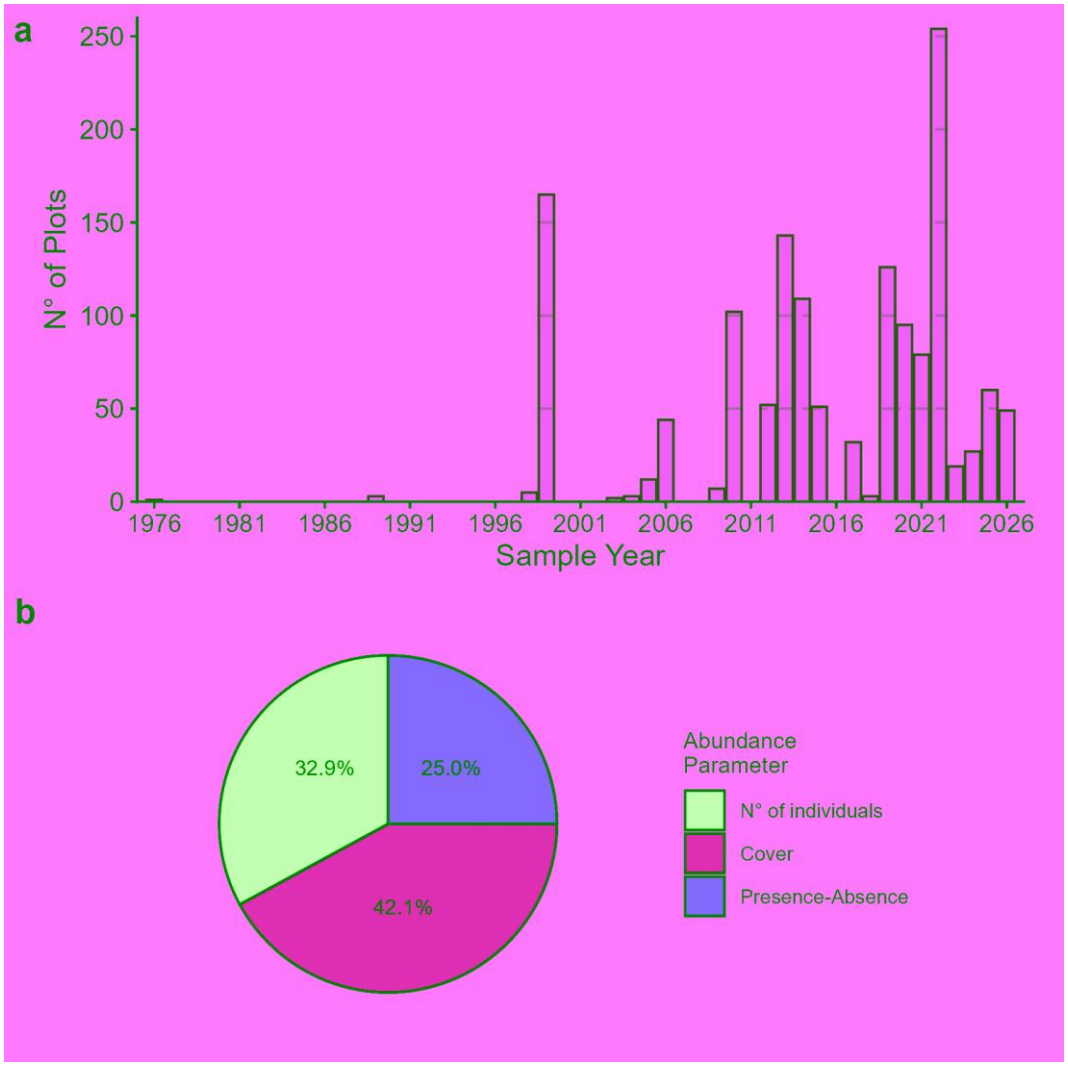


Fig. 2. a Frequency of plots by sampled year. **b** Distribution of types of abundance data across vegetation plots included in the database across south-central Chile. Categories include presence–absence, number of individuals, percentage cover, and crown diameter.

Representation across ecological units

Plots are asymmetrically distributed across vegetation types across south-central Chile. A large proportion of plots is concentrated in sclerophyllous forest and shrubland formations (Fig. 3), reflecting their ecological importance, the historical focus of research efforts in these ecosystems. Other vegetation formations (e.g., lauriphyllous forest or deciduous arborescent shrubland) are represented by fewer plots, indicating gaps in coverage across certain ecological contexts (Fig. 3).

Overall, Parcelas-CL captures a broad range of ecological conditions, 71.43% of the vegetation formations within the Mediterranean region (i.e., thorny, high-Andean dwarf and desert scrub, sclerophyllous shrubland, and thorny, deciduous and sclerophyllous forest; Fig. 3). However, sampling intensity varies across vegetation formations; certain Mediterranean vegetation formations are absent in Parcelas-CL, such as desert dwarf scrub, high-elevation

dwarf shrubland and high-elevation grassland. This uneven representation highlights vegetation formations where additional sampling would improve the ecological completeness of Parcelas-CL.

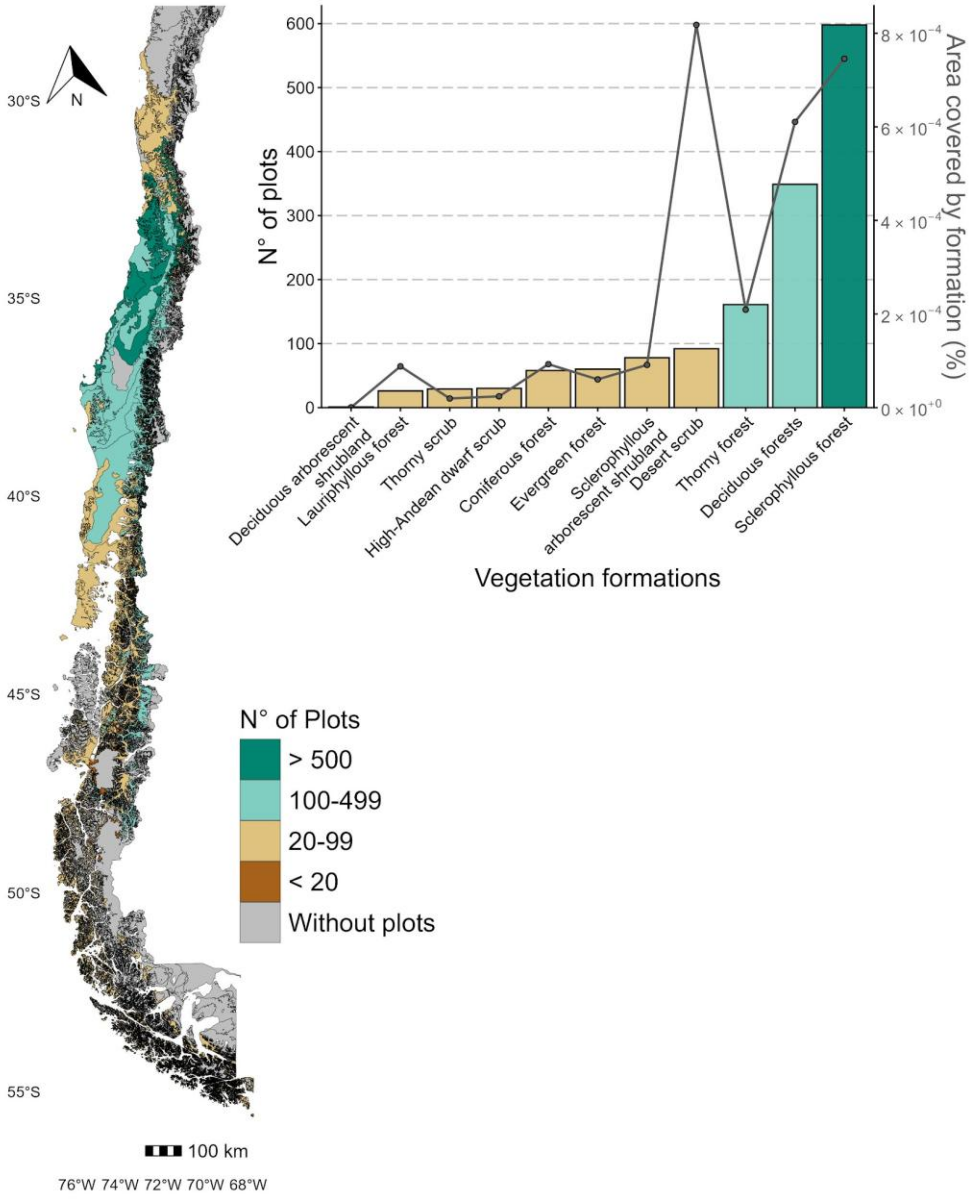


Fig. 3. Spatial distribution of vegetation formations across Mediterranean ecosystems in Chile, following Luebert and Pliscoff (2017) represented in the vegetation plot dataset. Colors represent ranges of the number of plots per vegetation formation. The barplot shows the distribution of plot frequency (left y-axis) and points representing percentage of total vegetation cover in south-central Chile(right y-axis).

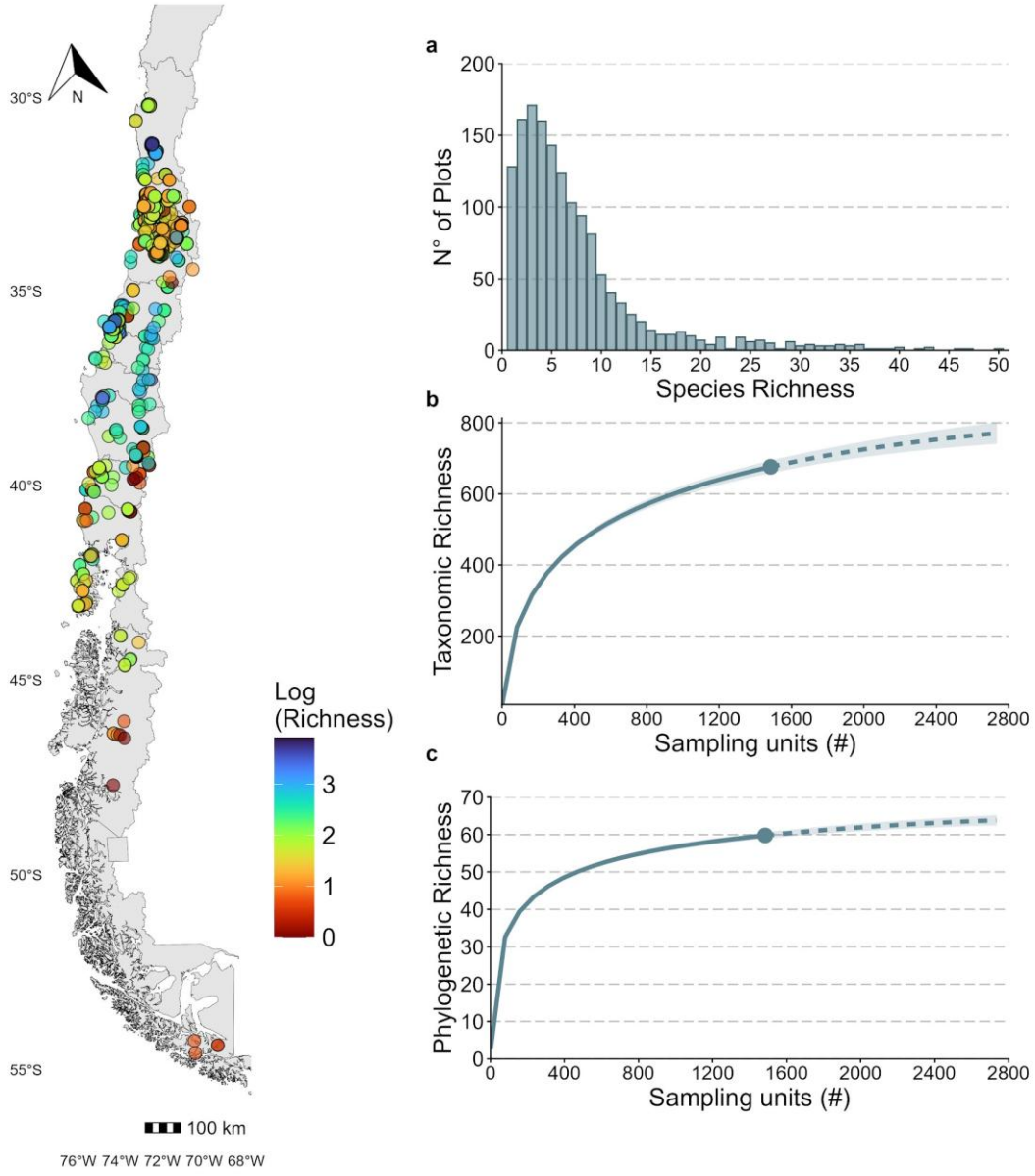
Species representation and sampling completeness

The final database includes 676 plant species recorded across all plots. Species richness ranges from 0 to 50 species per plot, with a median of 5 species (±0.14). Most observations contain a relatively low to moderate number of species (0 - 10), whereas fewer plots exhibit high species richness (Fig. 4). Sampling effort curves

based on Hill numbers (q = 0) indicate that Parcelas-CL captures a substantial portion of the regional species pool. Species richness increases rapidly initially and gradually saturates (Fig. 4). However, the absence of a clear asymptote suggests that additional sampling in underrepresented vegetation formations such as coniferous, evergreen and lauriphyllous formations would likely increase species diversity in Parcelas-CL representation. A similar pattern is observed for phylogenetic diversity, with a rapid initial increase followed by a gradual leveling-off. Nevertheless, the curve approaches saturation more quickly than taxonomic richness, suggesting that current sampling captures a substantial fraction of the phylogenetic diversity, even though additional sampling would still contribute to resolving finer-scale evolutionary structure (Fig. 4).

Spatial patterns of species richness indicate that higher species richness is generally associated with the central area of the Mediterranean of Chile (i.e., around Santiago), where sampling density is highest. This overlap between sampling effort and species richness underscores the importance of considering sampling intensity when

222 interpreting biodiversity patterns derived from compiled datasets (Roswell et al., 2021).



223

224 **Fig. 4.** Spatial distribution and richness patterns of vegetation plots across south-central Chile. Circles represent
225 individual plots (n = 1,485), with color indicating species richness per plot. **a** Frequency of plots based on their
226 species richness. **b** Sampling effort curves for taxonomic and **c** phylogenetic diversity. **b and c** show species and
227 lineage diversity for Hill number 0, which gives similar weight to abundant and rare species or lineages. Solid
228 lines represent interpolation (observed accumulation), while dashed lines represent rarefaction.

229

230 **CONCLUSIONS AND FUTURE DIRECTIONS**

231 Currently, Parcelas-CL provides an extensive, spatially explicit database of vegetation plots in Mediterranean
232 ecosystems in south-central Chile. It significantly improves access to standardized vegetation plot data in this
233 underrepresented region, increasing the number of available plots from 117 records in sPlot to 1,485 plots in

Parcelas-CL. Thus, it helps reduce geographic data gaps and supports the development of integrative, cross-scale approaches in landscape ecology. By integrating heterogeneous data sources, Parcelas-CL establishes a solid foundation for local to landscape-scale biodiversity analysis and monitoring.

However, Parcelas-CL reveals uneven sampling efforts across the study region and vegetation formations. Beyond thorny, deciduous and sclerophyllous forests, relatively few vegetation formations are represented (e.g., four formations with no plots, and seven formations with fewer than 100 plots). These patterns highlight areas that should be prioritized for improving ecological knowledge and reducing bias in analyses across environmental gradients.

Looking forward, future efforts should focus on expanding spatial coverage to all of Chile, increasing representation of undersampled vegetation formations, and strengthening data standardization protocols (i.e., promoting the use of biodiversity data standards such as Darwin Event Core (Wieczorek et al., 2012) or Humboldt Extension (Guralnick et al., 2018)). Collaborative data-sharing initiatives should be encouraged, in order to improve the completeness, interoperability, and long-term usability of vegetation datasets in Chile and Latin America. Furthermore, the evolution of these data infrastructures must balance technical interoperability (e.g., FAIR) with ethical governance (e.g., CARE), ensuring that the efforts to make ecological data more broadly accessible serves not only global or regional scientific progress but also the collective benefit and sovereignty of local knowledge holders (Wilkinson et al., 2016; Carroll et a., 2020).

In this context, interoperability with complementary regional databases, such as Rasgos-CL (Alfaro et al., 2023), offers a pathway to integrating ecological data infrastructures and enabling cross-scale analyses of multi-faceted biodiversity patterns, highlighting the need for coordinated, open, and scalable data systems to support biodiversity monitoring and informed conservation and policy decisions amid ongoing environmental change.

ACKNOWLEDGMENTS

We thank Agencia Nacional de Investigación y Desarrollo de Chile (ANID; FONDECYT Iniciación 11241088 and FSEQ210022), and the Data Observatory Foundation.

REFERENCES

Alfaro E, Pérez-Tello V, Acevedo M, Ovalle J, Segovia R, Craven D (2023) Rasgos-CL: a functional trait database of Chilean woody plants. Glob Ecol Biogeogr 32(12):2072–2084. https://doi.org/10.1111/geb.13755

Beck J, Böller M, Erhardt A, Schwanghart W (2014) Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. Ecol Inform 19:10–15. https://doi.org/10.1016/j.ecoinf.2013.11.002

Berdugo M, Vidiella B, Solé RV, Maestre FT (2022) Ecological mechanisms underlying aridity thresholds in global drylands. Funct Ecol 36(1):4–23. https://doi.org/10.1111%2F1365-2435.13962

Boyle B, Hopkins N, Lu Z, Raygoza Garay JA, Mozzherin D, Rees T et al (2013). The taxonomic name resolution service: an online tool for automated standardization of plant names. BMC Bioinform 14(1):16.

https://doi.org/10.1186/1471-2105-14-16

Bruelheide H, Dengler J, Jiménez-Alfaro B, Purschke O, Hennekens SM, Chytrý M et al (2019) sPlot–A new tool for global vegetation analyses. J Veg Sci 30(2):161–186. https://doi.org/10.1111/jvs.12710

Carroll SR, Garba I, Figueroa-Rodríguez OL, Holbrook J, Lovett R, Materechera S et al (2020). The CARE principles for indigenous data governance. Data Sci J 19(43):1–12. https://doi.org/10.5334/dsj-2020-043

Chytrý M, Hennekens SM, Jiménez-Alfaro B, Knollová I, Dengler J, Jansen F et al (2016). European Vegetation Archive (EVA): an integrated database of European vegetation plots. App Veg Sci 19(1):173–180. https://doi.org/10.1111/avsc.12191

Dinerstein E, Olson D, Joshi A, Vynne C, Burgess ND, Wikramanayake E et al (2017) An ecoregion-based approach to protecting half the terrestrial realm. Biosci 67(6):534–545. https://doi.org/10.1093/biosci/bix014

Farley SS, Dawson A, Goring SJ, Williams JW (2018) Situating ecology as a big-data science: Current advances, challenges, and solutions. Biosci 68(8):563–576. https://doi.org/10.1093/biosci/biy068

Fuentes I, Lopatin J, Galleguillos M, McPhee J (2025. Vegetation browning as an indicator of drought impact and ecosystem resilience. Sci Remote Sens 11:100219. https://doi.org/10.1016/j.srs.2025.100219

Gonzalez A, August T, Bailey S, Bobiwash K, Boersch-Supan PH, Burgess ND et al (2026) From data to decisions: Toward a biodiversity monitoring standards framework. Proc Natl Acad Sci USA 123(10):e2519347123. https://doi.org/10.1073/pnas.2519347123

Griffith J, Lord JM, Catchen MD, Arce-Plata MI, Blanchet FG, Chandramohan M et al (2026). BON in a box: an open and collaborative platform for biodiversity monitoring, indicator calculation, and reporting. Biosci 76(4):345–358. https://doi.org/10.1093/biosci/biaf189

Gudde RM, Krug RM, Sica YV, Nelson HP, Françoise F, Gómez-Suárez M et al (2026). Implementing the FAIR and CARE Principles Simultaneously: Emerging Insights from IPBES. Data Sci J 25(1):3. https://doi.org/10.5334/dsj-2026-003

Guralnick R, Walls R, Jetz W (2018). Humboldt Core–toward a standardized capture of biological inventories for biodiversity monitoring, modeling and assessment. Ecography 41(5):713–725. https://doi.org/10.1111/ecog.02942

Heberling JM, Miller JT, Noesgaard D, Weingart SB, Schigel D (2021). Data integration enables global biodiversity synthesis. Proc Natl Acad Sci USA 118(6):e2018093118. https://doi.org/10.1073/pnas.2018093118

Hortal J, De Bello F, Diniz-Filho JAF, Lewinsohn TM, Lobo JM, Ladle RJ (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. Annu Rev Ecol Evol Syst 46(1):523–549. https://doi.org/10.1146/annurev-ecolsys-112414-054400

Hughes AC, Orr MC, Ma K, Costello MJ, Waller J, Provoost P et al (2021). Sampling biases shape our view of

the natural world. Ecography 44(9):1259–1269. https://doi.org/10.1111/ecog.05926

Jantzen CC, Stefan JV (2026). Putting FAIR into practice for ecologists: How to make ecological data more reusable. Ecol Inform 95:103712. https://doi.org/10.1016/j.ecoinf.2026.103712

Keith DA, Ferrer-Paris JR, Nicholson E, Bishop MJ, Polidoro BA, Ramirez-Llodra E et al (2022). A function-based typology for Earth's ecosystems. Nature 610(7932):513–518. https://doi.org/10.1038/s41586-022-05318-4

Luebert F, Pliscoff P (2022). The vegetation of Chile and the EcoVeg approach in the context of the International Vegetation Classification project. Veg Classif Surv 3:15–28. https://doi.org/10.3897/VCS.67893

Maestre FT, Benito BM, Berdugo M, Concostrina-Zubiri L, Delgado-Baquerizo M, Eldridge DJ et al (2021). Biogeography of global drylands. New Phytol 231(2):540–558. https://doi.org/10.1111%2Fnph.17395

Miranda A, Lara A, Altamirano A, Di Bella C, González ME, Camarero JJ (2020). Forest browning trends in response to drought in a highly threatened mediterranean landscape of South America. Ecol Indic 115:106401. https://doi.org/10.1016/j.ecolind.2020.106401

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J (2000). Biodiversity hotspots for conservation priorities. Nature 403(6772):853–858. https://doi.org/10.1038/35002501

Peyre G, Montesinos D, Giraldo D, Galán de Mera A, Ruthsatz B, Luebert F et al (2022). VegAndes: the vegetation database for the Latin American highlands. Veg Classif Surv 3:287–296. https://doi.org/10.3897/VCS.95750

Roswell M, Dushoff J, Winfree R (2021). A conceptual guide to measuring species diversity. Oikos 130(3):321–338. https://doi.org/10.1111/oik.07202

Sabatini FM, Jiménez-Alfaro B, Jandt U, Chytrý M, Field R, Kessler M et al (2022). Global patterns of vascular plant alpha diversity. Nat Commun 13(1):4683. https://doi.org/10.1038/s41467-022-32063-z

Seipel T, Haider S, MIREN consortium. (2007–2023). MIREN survey of plant species in mountains (Version v3) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.15749536

Van Tiel N, Fopp F, Brun P, Van Den Hoogen J, Karger DN, Casadei CM et al (2024). Regional uniqueness of tree species composition and response to forest loss and climate change. Nat Commun 15(1):4375. https://doi.org/10.1038/s41467-024-48276-3

Venegas-González A, Muñoz AA, Carpintero-Gibson S, González-Reyes A, Schneider I, Gipolou-Zuñiga T et al (2023). Sclerophyllous forest tree growth under the influence of a historic megadrought in the Mediterranean ecoregion of Chile. Ecosystems 26(2):344–361. https://doi.org/10.1007/s10021-022-00760-x

Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R et al (2012). Darwin Core: an evolving community-developed biodiversity data standard. PloS One 7(1):e29715.

https://doi.org/10.1371/journal.pone.0029715

Wiegleb V, Bruns A (2025). Whose knowledge counts? Unpacking the uneven geographies and politics of knowledge co-production in IPBES. Hum Ecol 53(1):73–86. https://doi.org/10.1007/s10745-025-00578-w

Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A et al (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3(1):1–9. https://doi.org/10.1038/sdata.2016.18