VegAndes: the vegetation database for the Latin American highlands

Gwendolyn Peyre, Daniel Montesinos, Daniela Giraldo, Antonio Galán de Mera, Barbara Ruthsatz, Federico Luebert, Marcela Ontivero, Nicolas García, Miguel Álvarez, Rosa Isela Meneses, et al.

VEGETATION CLASSIFICATION AND SURVEY

International Association for Vegetation Science (IAVS)

Vegetation Classification and Survey 3: 287–296
doi: 10.3897/VCS.95750

LONG DATABASE REPORT NEOTROPICAL VEGETATION

VegAndes: the vegetation database for the Latin American highlands

Gwendolyn Peyre[1], Daniel Montesinos[2], Daniela Giraldo[1], Antonio Galán de Mera[3], Barbara Ruthsatz[4], Federico Luebert[5], Marcela Ontivero[6], Nicolas García[5], Miguel Álvarez[7], Rosa Isela Meneses[8], Pablo Lozano[9], Daniela León[10], Maximilian Weigend[7], Fabien Anthelme[11], Milagro Palma[12], Carolina Rodriguez[1]

1 *Departamento de Ingeniería Civil y Ambiental, Universidad de los Andes, Bogotá, Colombia*
2 *Freie Universität Berlin, Botanischer Garten und Botanisches Museum Berlin-Dahlem, Berlin, Germany*
3 *Ciencias Farmaceuticas y de la Salud, Universidad San Pablo-CEU, Madrid, Spain*
4 *Geobotanik, Universität Trier, Trier, Germany*
5 *Facultad de Ciencias Agronómicas y Departamento de Silvicultura y Conservación de la Naturaleza, Universidad de Chile, Santiago de Chile, Chile*
6 *Instituto Argentino de Investigaciones de las Zonas Áridas, Mendoza, Argentina*
7 *INRES, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany*
8 *Instituto de Arqueología y Antropología, Universidad Católica del Norte, Antofagasta, Chile*
9 *Herbario ECUAMZ, Universidad Estatal Amazónica, Puyo, Ecuador*
10 *Department of Botany, University of Tartu, Tartu, Estonia*
11 *AMAP, Univ Montpellier, IRD, CIRAD, CNRS, INRAE, Montpellier, France*
12 *Facultad de Ciencias Naturales, Exactas y Tecnología, Universidad de Panama, Panama City, Panama*

Corresponding author: Gwendolyn Peyre (gf.peyre@uniandes.edu.co)

Academic editor: Melisa A. Giorgis ♦ **Received** 28 September 2022 ♦ **Accepted** 17 November 2022 ♦ **Published** 30 December 2022

Abstract

In the era of Big Data, Latin American countries and biomes remain underrepresented. To remediate this issue, promoting repositories for biodiversity data focused on Latin America is a main priority. VegAndes -Dpt the vegetation database for the Latin American highlands (GIVD: SA-00-005), is a novel dataset for georeferenced and standardized information on vascular pants in the region. The database compiles 5,340 vegetation plots sampled above the montane treeline and below the permanent snowline in 11 Latin American countries and spanning over seven decades. VegAndes currently encompasses 5,804 taxon names, corresponding to 3,858 accepted names, as well as 136 syntaxon names. The database is nested within a scientific consortium of Latin American experts on highland vegetation and piloted from the University of the Andes (Colombia). Because the VegAndes data can support multi-scale studies in botany, ecology and biogeography, the database makes an essential contribution to biodiversity research and management perspectives in Latin America.

Taxonomic reference: TROPICOS (preferential source, www.tropicos.org/), World Flora Online (secondary source, www.worldfloraonline.org/).

Keywords

database, flora, highland, Mesoamerica, phytosociology, plot data, South America, vegetation

GIVD Fact Sheet: VegAndes

GIVD Database ID: SA-00-005			**Last update:** 2022-12-14
VegAndes		**Web address:**	
Database manager(s): Gwendolyn Peyre (lunariaster@gmail.com)			
Owner: Gwendolyn Peyre			
Scope: VegAndes - the vegetation database for the Latin American highlands, is a novel referent dataset that compiles vegetation information sampled at high elevation in Mesoamerican and South American mountains.			
Abstract: VegAndes includes vegetation plots sampled - or transformed - with the phytosociological method. It solely includes plots located in highland areas, either from the tropicalpine zone or transitional montane zone. VegAndes contains revised and precisely georeferenced information that is made gradually available to the public, so it can be employed for research and management projects targeting high mountain ecosystems in Latin America.			
Availability: according to a specific agreement		**Online upload:** no	**Online search:** no
Database format(s): MySQL, PostgreSQL		**Export format(s):** plain text file	
Plot type(s): normal plots		**Plot-size range (m²):** 0.5 to 400	
Non-overlapping plots: 5340	**Estimate of existing plots:** 5340	**Completeness:** 100%	**Status:** completed and continuing
Total no. of plot observations: 5340	**Number of sources (biblioreferences, data collectors):** 86		**Valid taxa:** 3858
Countries (%): AR: 13.3; BO: 5.1; CL: 5.4; CO: 32.7; CR: 0.12; EC: 19.2; GT: 0.3; MX: 1.8; PA: 0.1; PE: 16.2; VE: 5.8			
Formations: Forest: 2% = Terrestrial: 2% // Non Forest: 98% = Terrestrial: 98% (Arctic-alpin: 100%; Non arctic-alpin: 0%)			
Guilds: all vascular plants: 100%			
Environmental data (%): altitude: 97.8; slope aspect: 60.6; slope inclination: 61.5; surface cover other than plants (open soil, litter, bare rock etc.): 35.9			
Performance measure(s): presence/absence only: 5%; cover: 95%			
Geographic localisation: GPS coordinates (precision 25 m or less): 19%; point coordinates less precise than GPS, up to 1 km: 73.5%; small grid (not coarser than 10 km): 7.4%			
Sampling periods: 1950-1959: 0.6%; 1970-1979: 0.5%; 1980-1989: 12.2%; 1990-1999: 9.9%; 2000-2009: 33.3%; 2010-2019: 33.8%; unknown: 9.6%			
Information as of 2022-12-14; further details and future updates available from http://www.givd.info/ID/SA-00-005			

Introduction

Mountains only encompass 13% of the global land's surface, yet they disproportionally contribute to global biodiversity and provide key ecosystem services to billions (Körner et al. 2011). For instance, more than half of the human population relies on freshwater bodies with high mountain sources for water supply (Pomeroy et al. 2015). Mountains also harbor a great diversity of ecosystems and landscapes, which have been shaped by orogeny, geo-environmental factors, and land-use history. These landscapes are distributed along the altitudinal gradient, which can be divided into the following altitudinal belts (from low to high elevations): foothills, premontane, montane, subalpine, alpine, and subnival-nival. A pronounced change in vegetation physiognomy occurs at the timberline, typically located in the subalpine belt, where forest-dominated ecosystems give way to highland ecosystems, such as shrublands, grasslands and deserts. Unlike the naturally continuous forests below, mountain highlands often form archipelagos of continental biogeographic islands (Whittaker et al. 2017). This noteworthy spatio-environmental pattern can be found in small and isolated mountain ranges, such as the Australian Alps, but also in broad and continuous ranges, for instance the European Alps. Notable exceptions to the rule include the high plateaus of the Tibet and the Andean Altiplano, where most of the subalpine and alpine belts remain continuous and only the highest peaks prevail. Due to their sky island nature, highland floras and floristic assemblages are chiefly characterized by a high index of endemicity and extreme susceptibility to change (Sklenář et al. 2014).

With 550,000 km² of highlands, calculated as the surface above the potential treeline (Testolin et al. 2020), Latin America presents the second largest highland area after Asia and its 2,590,000 km². Latin American highlands stretch from Lat 20°N in Mexico to 55°S in the Chilean-Argentinian Patagonia, resulting into one of the longest sub-continuous mountain chains on earth. The 7,500 km Andean cordillera alone accounts for 99% of the Latin American highlands, spanning over the Northern, Central and Southern Andean biogeographic provinces (Anthelme and Peyre 2020). Additional highlands are found on smaller mountain ranges, whose uplift was either (i) related to Andean orogeny, such as the Cordillera de la Costa in Venezuela; (ii) or independent from it and due to isolated mountain building events, for example the Trans-Mexican Volcanic Belt. Therefore, Latin American mountains offer an outstanding overview of altitudinal belt compression and expansion from tropical to temperate climates, along a unique latitudinal gradient (Körner and Spehn 2019).

Scientific interest for Latin American mountain ecology and biogeography, has quickly increased over the last decades, mostly because they i) contribute to more than half of the 85,000 species pool that conform the Latin American vascular flora (Ulloa et al. 2017); and ii) coincide with several noteworthy biodiversity hotspots, such as the Tropical Andes, Mesoamerica, Tumbes-Chocó-Magdalena, and the Chilean Winter Rainfall-Valdivian Forests. Yet, broad-scale research on mountains, and highlands specifically, remains limited due to data availability (Sabatini et al. 2021). This said, floristic data contributions are steadily increasing and being made available in national repositories, such as the SiB Colombia (https://biodiversidad.co), as well as on global platforms, for instance (www.gbif.org) and TRY (https://www.try-db.org). In contrast, vegetation data remains scarce and difficult to access (but see Peyre et al. 2015). This type of data is very advantageous as it provides information on the distribution of both species and plant communities, and as a result, it should be a primary research focus for Latin American mountains (Bottin et al. 2020).

In this study, we present VegAndes, the vegetation database for the Latin American highlands, a novel dataset to address the dramatic data-gap in vegetation surveys in Latin American mountains and provide strong bases for future studies in plant ecology and biogeography. The VegAndes initiative emerges from consortium discussions involving expert botanists and vegetation scientists dedicated to highland research in Latin America. It builds on local and regional efforts, and compiles vegetation plot information that is methodologically and taxonomically standardized into a single repository housed at the University of the Andes in Bogotá (Colombia). Both the consortium and dataset are expected to grow with regards to data volume and outreach in the short to mid-term, with the publication of the VegAndes webpage.

Study area

The study area encompasses the highlands of all Latin American mountains, from Mexico to Chile and Argentina, with a focus on the Andes (Figure 1).

In the north, Mesoamerican highlands are distributed over the Mexican transition zone and Central American mountain ranges down to the Panama Isthmus (Lat 20°N–11°N). They are often isolated and encompass a variety of substrates and climates. For instance, Mexican mountains are mainly volcanic and are subject to strong temperature seasonality, whereas Panama and Costa Rica mountains are principally sedimentary and experience from overall low seasonality and humid climates (Kappelle and Horn 2016). The South American continent harbors the Andean cordillera and its three main biogeographic provinces, the Northern Andes (Lat 11°N–5°S), Central Andes (Lat 5°S–25°S) and Southern Andes (25°S–55°S), as well as small secondary ranges, such as the Amazonian volcanoes. The highlands of the Northern Andes belong to the Páramo region that stretches from Venezuela to northern Peru, as far as the Huancabamba depression biogeographical barrier (Weigend 2002). They are characterized by high overall humidity and low seasonality, leading to a great diversity of highland ecosystems (Luteyn et al. 1999; Rangel 2000). The Central Andes cover a vast area including the Peruvian cordilleras and the Altiplano, down to the summer/winter rain transition zone (Anthelme and Peyre 2020). The Puna region dominates the corresponding highlands and involves at least seasonally arid climates with distinct temperature seasonality. Finally, the Southern Andes run further south and include two distinct regions, known as the Mediterranean Andes, north of Temuco (Lat 25°S–38°S), and the Patagonian Andes, south of Temuco (Lat 38°S–55°S). Typically, the Mediterranean Andean highlands belong to the Prepuna or Cardonal region and are characterized by their highly seasonal climates with rainfall peaks during the austral winter (Rivas-Martínez et al. 2011). Lastly, the Patagonian Andes are chiefly fragmented and show a progressive shift from temperate highlands to austral boreal lowlands on the Darwin Cordillera (Martínez Carretero 1995).

On the altitudinal gradient, Latin American highlands typically encompass the subalpine (supratropical) belt above the timberline, the alpine (orotropical) belt, subnival belt (cryotropical) and nival belt (ice-athermic). However, each belt's surface and elevation range vary substantially with latitude and position within a mountain range. A recent study on alpine areas worldwide places the potential treeline roughly at a minimum of 1,000–1,500 m a.s.l. in the Patagonian Andes and a maximum of 3,500–4,000 m a.s.l. near the equator (Testolin et al. 2020). However, the accurate highland delimitation depends on the actual timberline, which is usually lower than the potential treeline and remains poorly documented in Latin America (e.g. Luebert and Pliscoff 2018; Peyre et al. 2021). The upper highland limit is either defined by the mountain top or glacier masses, which are currently located between approximately min. 1,500–2,000 m a.s.l. in the Patagonian Andes and 5,500–5,800 m a.s.l. in the Central Andes (Pfeffer et al. 2014; Dussaillant et al. 2019).

Methods

The database structure was programmed in SQL, following VegPáramo (Peyre et al. 2015). It revolves around a primary table containing the basic vegetation plot data, and five secondary tables that are connected to the primary table via the VegAndes plot ID (except for the taxon and syntaxon lists, detailed below) (Figure 2). Among the extensive array of data-sources consulted for Latin America, only those containing plot vegetation data expertly identified as highland or ecotonal vegetation by the data's authors were included into VegAndes. Plots were not discriminated between zonal and azonal vegetation as long as they fitted the at- or above-the-timberline rule, or equivalent in desertic areas. All plots were primary data except those proceeding from the multi-source VegPáramo dataset (www.vegparamo.com). In this case, plots were referenced as their secondary source and inquiring users should consult the VegPáramo webpage for further details on primary data-sources.

The **Primary table** contains the plot data, with unique values to represent species occurrence and/or cover

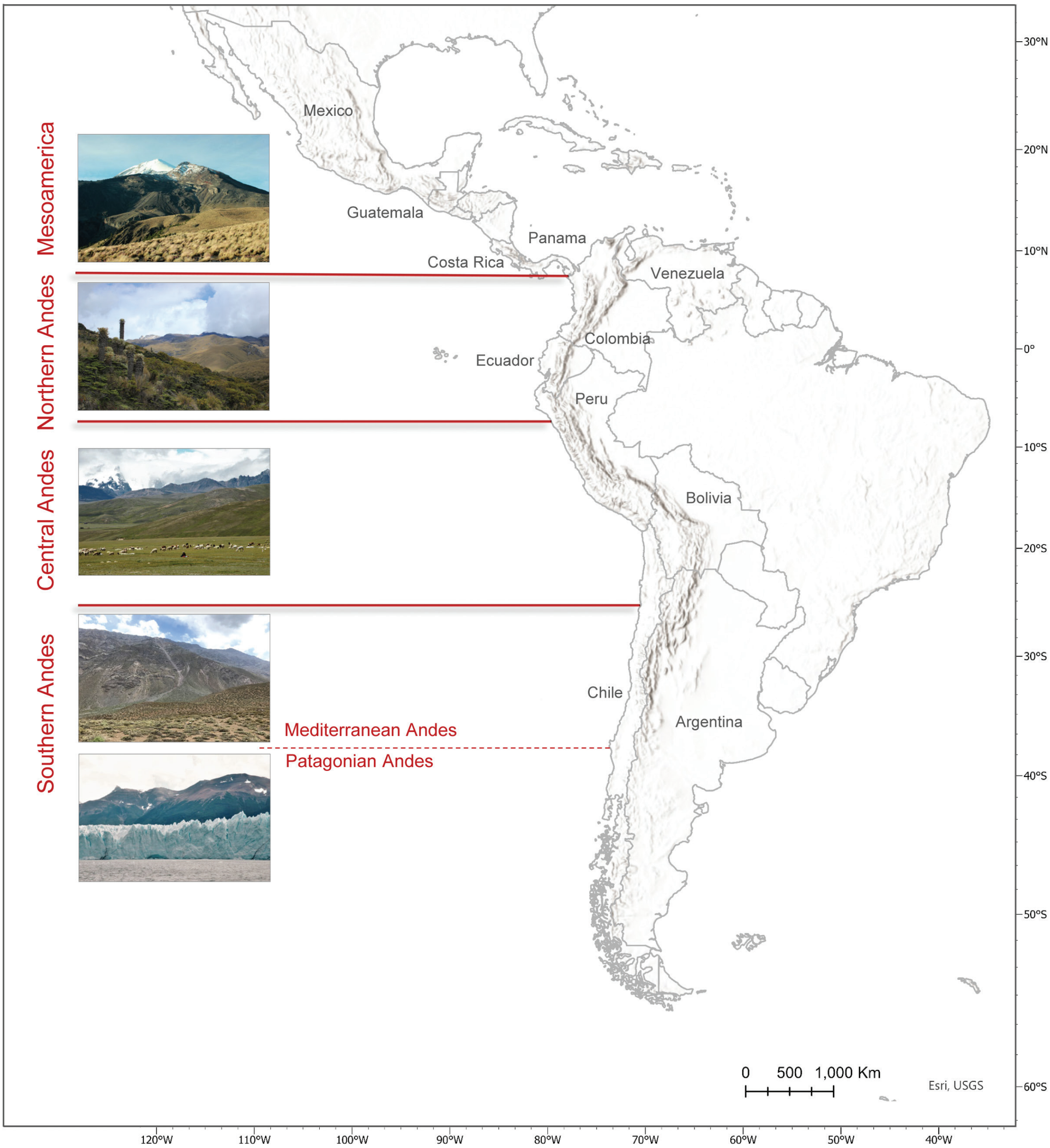

Figure 1. Topographic map of Latin America, with focus on mountain areas and their division into four main biogeographical provinces, Mesoamerica (photo: Pico de Orizaba, Mexico, credit G.P.), Northern Andes (photo: Nevado Santa Isabel, Colombia, credit G.P.), Central Andes (photo: Macusani, Peru, credit A.G.M.), and Southern Andes. The latter province can be divided into two main regions, the Mediterranean Andes (photo: Tacora, Chile, credit F.L.) and Patagonian Andes (photo: El Calafate, Argentina, credit A.G.M.). The map indicates names for those countries with highlands covered by the VegAndes vegetation data.

within plots. The species list uses the initial taxon names provided by the plot authors as well as a taxon ID assigned to each name. Non-vascular plants were excluded from the datasets. All non-identified taxa above the genus level were also discarded, and those at or below the genus level were included and assigned a codified taxon ID including the source name and unique number. Useful particles to assess the certainty of taxonomic determination such as cf., aff. and gr., were recorded. All names will subsequently be updated pending future taxonomic determinations. VegAndes allows for presence-absence data as well as several cover-scales: the Braun-Blanquet (1964) phytosociological scale, transformed versions of the Braun Blanquet scale, for instance Van der Maarel's (1979), and percentage cover (in %). The primary table can be converted into a work table containing accepted names only. To perform

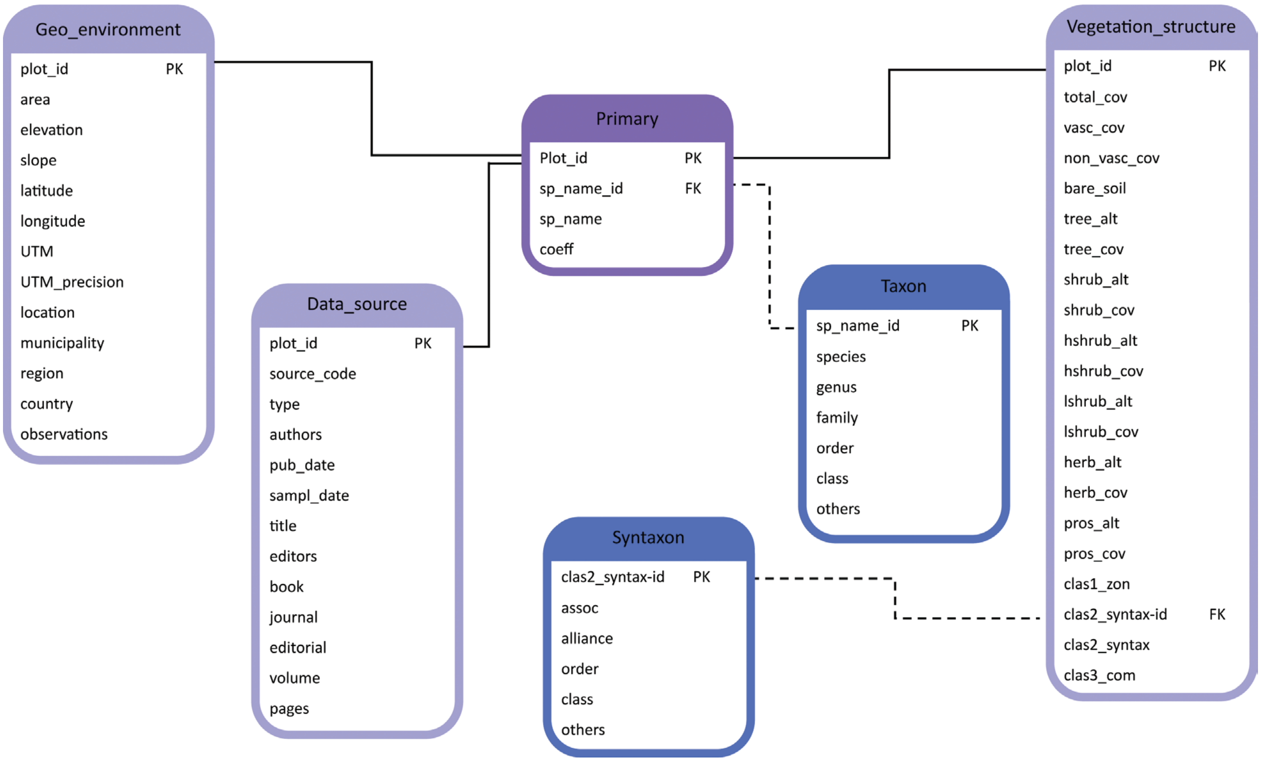

Figure 2. Schematic structure of the VegAndes database and its six tables, including the primary table (violet), three secondary tables with the plots' attributes (lilac) and two secondary tables with the taxon and syntaxon checklists (blue). PK: Primary Key, FK: Foreign Key.

the transfer, all recorded names need to be cross-checked with the Taxon list (see details below), so to remove unresolved names and merge rows with identical names. Whenever accepted names are duplicated in the same plot (for example after checking two sub-species of the same species in the same plot), their cover values are summed. However, if those values come as cover coefficients, then a simple procedure based on probability rules is used to either retain or increase the resulting coefficient.

The **Geo-environmental table** resumes the geographic and environmental information recorded for the VegAndes plots. It contains the plot surface area (m^2), elevation (m a.s.l), slope (°) and aspect. It also includes the plot geolocation, both as decimal coordinates and UTM grid cell (in Datum WGS84). Whenever a plot is precisely georeferenced, with a ≤ 10 m error as provided by modern GPS, the decimal coordinates are considered exact. Nevertheless, if the original plot georeferencing is imprecise (10 m–10 km), then its decimal coordinates refer to the centroid of the smallest resolution UTM grid cell available. Plots with very imprecise georeferencing, above 10 km, were not registered into VegAndes. Finally, the table includes the plot's precise location (often a toponym), municipality, region and country, as well as optional observations on terrain or additional features.

The **Vegetation-structure table** reports all attributes on plot physiognomy and plant community labels. It includes information on the total, vascular and non-vascular vegetation covers (in %), bare soil (in %), as well as height (in cm) and cover (in %) for the following vegetation strata (if applicable): trees, high and low shrubs, herbs and prostrate strata. Furthermore, this table accounts for three vegetation classification systems. First, it offers a simple classification into azonal (azo) and zonal (zo) vegetation, the latter with their respective altitudinal belt as subalpine (zo-sub), alpine (zo-alp) and subnival-nival (zo-niv). Second, the table records the syntaxon name that was initially given by the plot's authors and a syntaxon ID assigned to each name, which can be further checked by the Syntaxon list (see details below). Last, it shows the plant community informal name, combining structural qualifiers and the name of up to two dominant and/or diagnostic plant species, for example "Bunchgrassland with *Espeletia grandiflora* and *Eryngium humboldtii*".

The **Data-source table** contains the information related to the primary or secondary (when applicable) source of the vegetation plot data. Both published and unpublished sources are considered, but unpublished ones, such as private datasets and academic theses, were carefully and expertly revised beforehand. The table includes the source's full author list, publication date, sampling date, title, and journal/book complete reference. Currently, 53 data-sources are included in VegAndes (86 when counting VegPáramo's multi-source input), of which 11 are unpublished datasets (see Suppl. material 1 for a complete list of the data-sources).

Finally, VegAndes presents two additional secondary tables in the form of the **Taxon list** and **Syntaxon list** that are not connected to the Primary table by the VegAndes plot ID, but are instead meant to check name validity for taxa, via the taxon ID (Primary table), and syntaxa, through the syntaxon ID (Vegetation-structure table), respectively. The Taxon list was built based on the initial names contained in the Primary table. The global botanic platforms TROPICOS (www.tropicos.org) and World Flora Online (www.worldfloraonline.org) were used as preferential and secondary sources respectively to assess each name's validity. Whenever a taxon is classified as doubtful or unresolved by both platforms, the initial name is retained. Similarly, hybrid taxa are kept as such and considered as accepted names. The Taxon list also includes family names and higher clades to allow for easy data-searches in: (i) Angiosperms, including Magnoliids, Monocots and Eudicots; (ii) Gymnosperms; (iii) Lycophytes; (iv) Monilophytes. It is expected that the Taxon list will undergo annual synonymy revisions and grow as new data is added to VegAndes. The Syntaxon list is in its initial stages and compiles all existing information on syntaxonomy as displayed in the VegAndes data-sources. As a result, this table contains hierarchical syntaxon terms for classes, orders, alliances and associations. However, there has been no revision of the complete classification to date, so it is advised to handle this information as highly provisional. At present, 5,804 taxon names, corresponding to 3,858 accepted names, as well as 136 syntaxon names are recorded in the VegAndes lists.

Results and discussion

VegAndes is registered under Global Index of Vegetation-Plot Database (GIVD) – ID: SA-00-005 and it currently contains 5,340 vegetation plots. The temporal distribution of the data includes the 1950–2022 period, with 75% of the sampling taking place after the year 2000 (Suppl. material 2).

Distribution of the VegAndes plot data

The VegAndes plots are spread over the Latin American highlands from Lat. 20°N to 39°S in a fairly uniform manner, except for the Patagonian Andes region in the Southern Andes (Table 1, Figure 3A).

Most VegAndes data comes from the Northern Andes (almost 60% of the plot total) and especially from Colombia and Ecuador (together summing up to 52%). The Central Andes come second with almost 23% of the plot total, essentially from Peru (16%). Certain areas contribute disproportionately to the dataset, for instance the Nevados sector in Colombia (Lat. 4.5°N – Long. 75.5°W) with 472 plots in 10,000 km^2, and to lesser extent the Moquegua mountains in Peru (Lat. 16.5°S – Long. 70.5°W) with 385 plots in 10,000 km^2. Differences in sampling efforts are usually due to expert location and ease of access. For example, the guerrilla conflict in Colombia substantially limited fieldwork in certain mountain areas over the last decades, many of which have only become accessible after the peace treaty of 2016 (Negret et al. 2017). Conversely, other areas are now becoming difficult of access due to civil tensions or local conflicts, for instance related to illegal extractive practices. The remaining spatial data gaps in VegAndes should be addressed and filled whenever possible, especially in the Patagonian Andes, central Peru and Central American mountains.

From an altitudinal standpoint, the VegAndes plot data can serve as simple proxy and give approximations about the current timberline position and highland elevation range throughout Latin America (Figure 3B). The VegAndes plots show a hump-shaped pattern for the Andes, with high elevations concentrated at mid-latitudes and around the equator, and lower elevations towards the edges of the main cordillera around Lat. 10°N and 40°S. This pattern reflects alpine ecological theory, which states that altitudinal belts typically occur at higher elevation and expand towards the equator (Körner et al. 2011; Körner and Spehn 2019). Moreover, the plot data manifests the biogeographical transition at Lat. 5–7°S, with relatively low elevation plots in comparison to the high-elevation Central and Northern Andes (Weigend 2002). Similarly, a clear depression in plot elevation is observed at the biogeographical transition between the Northern Andes and Mesoamerica. According to the VegAndes data, the approximate timberline by proxy appears to be highest in the Altiplano region of the Central Andes (Lat. 20–30°S) and in the northernmost Mesoamerican mountains (Lat.

Table 1. Plot contributions (as absolute numbers and percentages) from each country and biogeographical province to the VegAndes dataset.

Countries	Mesoamerica	Northern Andes	Central Andes	Southern Andes	Total
Mexico	94				94 (1.8%)
Guatemala	18				18 (0.3%)
Costa Rica	1				1 (0.02%)
Panama	8				8 (0.1%)
Venezuela		312			312 (5.8%)
Colombia		1748			1748 (32.7%)
Ecuador		1023			1023 (19.2%)
Peru		57	810		867 (16.2%)
Bolivia			270		270 (5.1%)
Chile			121	169	290 (5.4%)
Argentina				709	709 (13.4%)
Total	121 (2.3%)	3140 (58.8%)	1201 (22.5%)	878 (16.4%)	

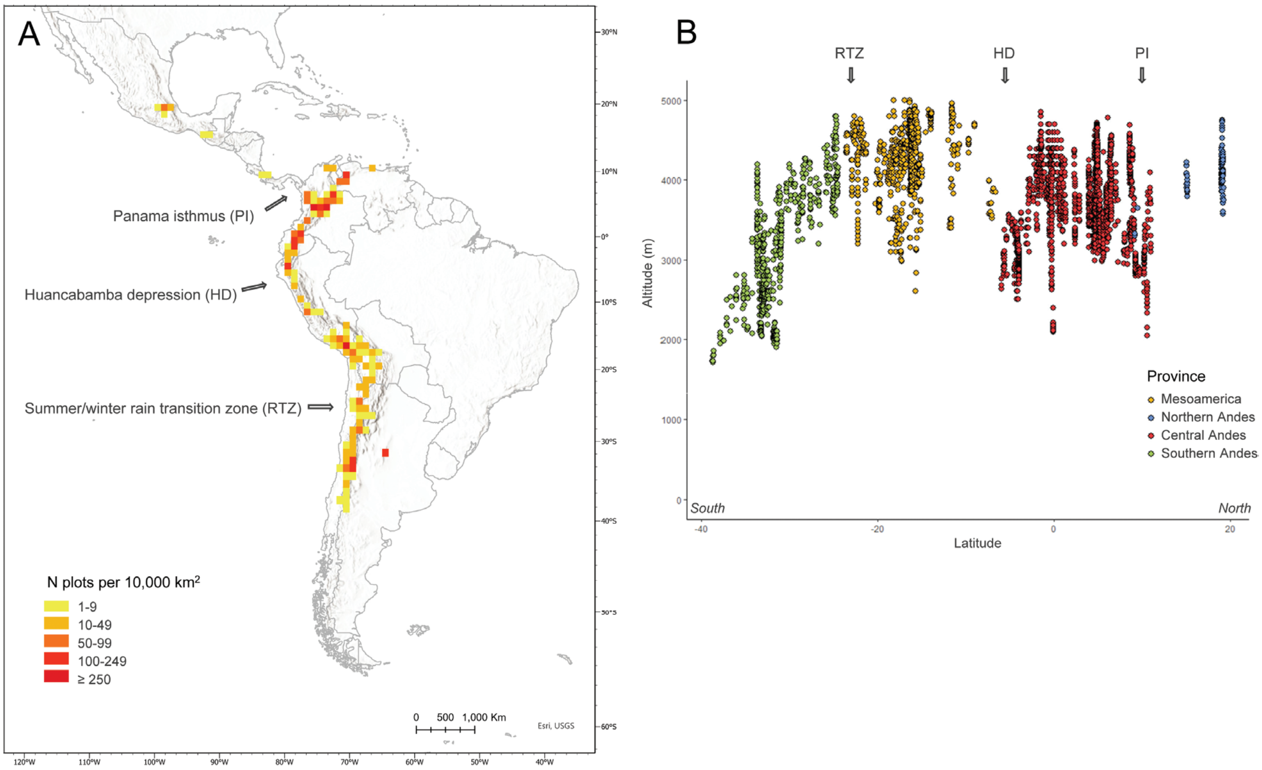

Figure 3. Distribution of the VegAndes plots in Latin America. (A) Plot density in 100 km × 100 km grid cells throughout the Latin American highlands, and position of the main biogeographical transitions. (B) Number of plots along the latitudinal and altitudinal gradient (colors represent the four main biogeographical provinces, from North to South: Mesoamerica, Northern Andes, Central Andes and Southern Andes).

20°N). In the Central Andes plots reached the highest elevations, near the 5,000 m isoline - presumably close to the local snowline. In comparison, most of the Northern Andes (Lat. 2°S–10°N) and, less so, part of the Mediterranean Andean region in the Southern Andes (Lat. 30–35°S), show some of the widest altitudinal ranges with typical highland vegetation reported over more than 2,000 m of elevation at the same latitude. Highland plots are found at dramatically decreasing elevations towards the Patagonian Andes region down to a timberline at ca. 1,500 m a.s.l. near Lat. 40°S, confirming previous findings on Southern Andean highlands (Martínez Carretero 1995; Luebert and Pliscoff 2018).

Floristic content of VegAndes

VegAndes contains 3,858 vascular plant species across 151 plant families, which in turn account for (i) 85% Angiosperms, including 67% Eudicots, 15% Monocots and 3% Magnoliids; (ii) 2% Gymnosperms; (iii) 2% Lycophytes; and (iv) 11% Monilophytes. According to the VegAndes data, dominant families in the Latin American highlands included *Asteraceae*, *Poaceae*, *Orchidaceae*, *Fabaceae* and *Cyperaceae*, summing up 40% of the total floristic inventory (Figure 4A). Comparing these results with World Flora Online input for the same family pool and at the global scale, we encountered a similar ranking for dominant families (although different proportions) with *Asteraceae* followed by *Orchidaceae*, *Fabaceae*, *Rubiaceae* and *Poaceae*. Therefore, the VegAndes provides an interesting subset of the world's flora, with an evident underlying grassland trend, as illustrated by the high relative preponderance of graminoid Monocots (*Poaceae* and *Cyperaceae*). The ranking found in family representation resembles other tropical mountain floras, such as the Papua New Guinea and Afromontane highlands (Sklenář et al. 2014). The VegAndes data also finds good representation of *Fabaceae* and *Lamiaceae* that are very important families in the subtropical and temperate highlands. At this stage, these results should be interpreted with caution since the data analysed might not represent the full spectrum of plant communities in the Latin American highlands and should be cross-checked with complementary floristic data types (e.g. Bottin et al. 2020).

A comparison of species richness patterns of the Latin American highlands based on VegAndes data shows complex patterns (Figure 4B). On average, the plot data summed up to about 20 families per 10,000 km², while certain areas doubled that number, principally in Ecuador (Lat. 4°S, Long. 79°W, and Lat. 0°N, Long. 78°W) and Colombia (Lat. 4°N, Long. 74°W). The family-rich areas could reflect very phytodiverse areas, but they could also show potential artefacts due to sampling effort or plot altitudinal range. For example, the spatial patterns observed in the Páramo region do not entirely match the areas identified as species rich vs. species poor in Peyre et al. (2019), who assessed richness variation at the local scale (plot level). Additional

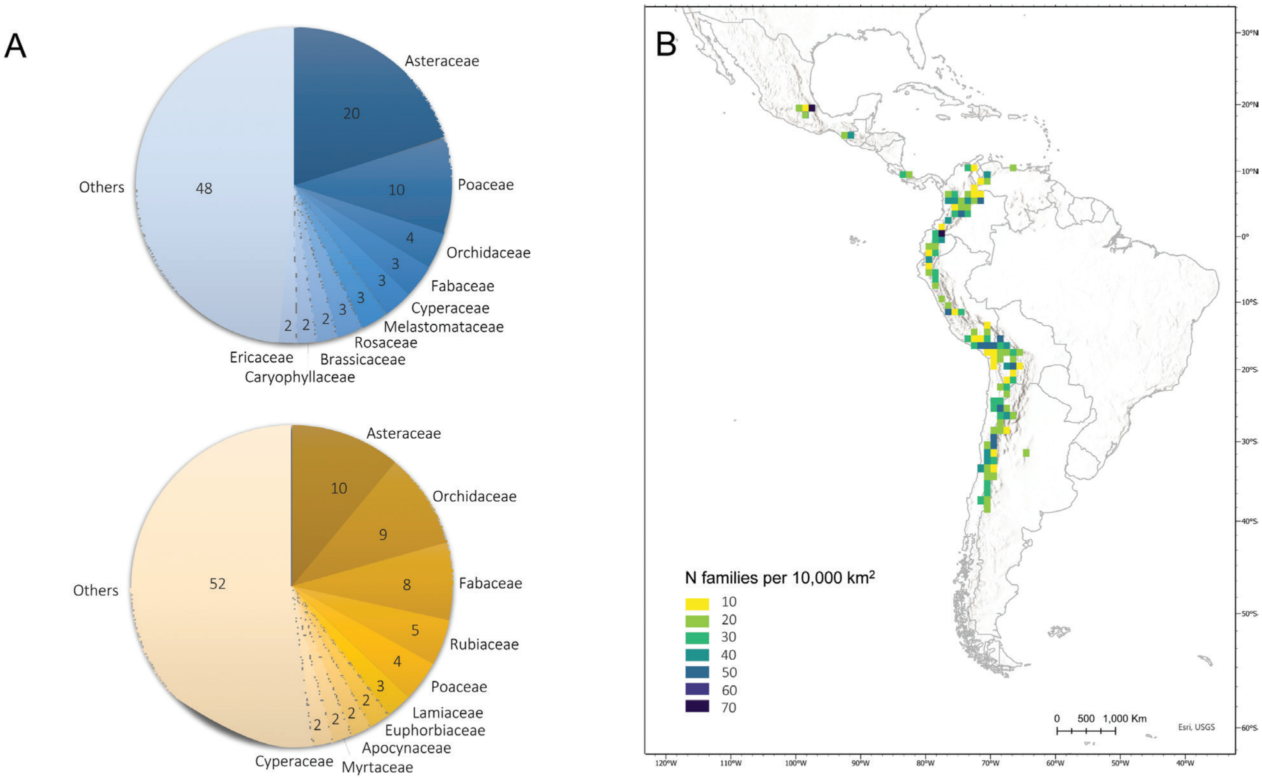

Figure 4. Distribution of the families of vascular plants in the VegAndes database. (A) Relative contribution of the dominant ten families to total species richness according to the VegAndes data (in blue) and the global vascular flora (species numbers obtained from the World Flora Online, www.worldfloraonline.org, accessed July 2022, in yellow). (B) Family richness in 100 × 100 km grid cells throughout the Latin American highlands.

data, paired with both socio-economic and ecological inputs should help shed light on these interesting patterns.

Future perspectives

With the emergence of global vegetation initiatives covering wide spatio-temporal frames, such as sPlot, the global vegetation database (Bruelheide et al. 2019), Latin America has more than ever been put on the map for its floristic richness and uniqueness (e.g. Testolin et al. 2021). Research boundaries should be pushed to improve our understanding of the fascinating Latin American highlands and their vegetation, and to do so, data limitations need to be overcome. VegAndes is a pioneer dataset that not only fills in data-gaps, but also offers an array of research and collaboration opportunities for regional vegetation scientists. At this early stage, VegAndes can already provide input for transcontinental studies and complement global initiatives with a better representation of Latin American mountains. At present, the data is fully available to the consortium partners and can be extended to external collaborators in view of agreed upon projects. On the short- to mid-term, the consortium is expected to grow, and with it, data contributions from less-sampled areas and across a larger timespan. It is our main priority to create a specific VegAndes webpage based on the University of the Andes server to allow easy data queries and downloads. Through its webpage, VegAndes will be made open-access and furthermore connected to other key botanical repositories, such as sPlot, VegPáramo (GIVD ID: SA-00-002, Peyre et al. 2015) and sudamerica (GIVD ID: SA-CL-001, Álvarez et al. 2012). The next project employing VegAndes will build on the existing syntaxonomic approximations in the region (e.g. Rangel 2000; Galán de Mera and Orellana 2006; Rivas-Martínez et al. 2011) and classify the complete extent of Latin American highlands into phytoregions. The promising consequent results will complement the current VegAndes syntaxon list and more importantly contribute to our understanding of the fascinating Latin American highland biogeography.

Author contributions

G.P. planned the research, all authors contributed the data. G.P. and D.G. compiled the data, G.P. built the database and led the writing, while all authors critically revised the manuscript.

Acknowledgements

We would like to thank the LMI-BioInca initiative for funding D.G and C.R. We are also grateful to Viviana Camelo, Sophia Lopez, and Nicolas Salcedo for their help with data handling.

References

Álvarez M, Möseler BM, San Martín C, Ramírez C, Amigo J (2012) CL-Dataveg–a database of Chilean grassland vegetation. Biodiversity and Ecology 4: 443. https://doi.org/10.7809/b-e.00230

Anthelme F, Peyre G (2020) Biogeography of South American highlands. In: Goldstein MI, DellaSala DA, DiPaolo DA (Eds) Earth Systems and Environmental Sciences, Encyclopedia of the world's biomes. Elsevier, Amsterdam, NL, 1–12. [eBook ISBN: 9780128160978] https://doi.org/10.1016/B978-0-12-409548-9.11811-1

Bottin M, Peyre G, Vargas C, Raz L, Richardson JE, Sanchez A (2020) Phytosociological data and herbarium collections show congruent large-scale patterns but differ in their local descriptions of community composition. Journal of Vegetation Science 31: 208–219. https://doi.org/10.1111/jvs.12825

Braun-Blanquet J (1964) Pflanzensoziologie, grundzuge der vegetations kunde. Springer Vienna, AT, 866 pp. https://doi.org/10.1007/978-3-7091-8110-2

Bruelheide H, Dengler J, Jiménez-Alfaro B, Purschke O, Hennekens SM, Chytrý M, Pillar VD, Jansen F, Kattge J, … Zverev A (2019) sPlot – A new tool for global vegetation analyses. Journal of Vegetation Science 30: 161–186. https://doi.org/10.1111/jvs.12710

Dussaillant I, Berthier E, Brun F, Masiokas M, Hugonnet R, Favier V, Rabatel A, Pitte P, Ruiz L (2019) Two decades of glacier mass loss along the Andes. Nature Geoscience 12: 802–808. https://doi.org/10.1038/s41561-019-0432-5

Galán de Mera A, Orellana JAV (2006) Aproximación al esquema sintaxonómico de la vegetación de la región del Caribe y América del Sur. Anales de Biología 28: 3–27.

Kappelle M, Horn SP (2016) The Páramo ecosystem of Costa Rica's highlands. In: Kappelle M (Ed.) Costa Rican ecosystems. The University of Chicago Press, Chicago, IL, US, 492–523. https://doi.org/10.7208/chicago/9780226121642.001.0001

Körner C, Spehn EM [Eds] (2019) Mountain biodiversity: a global assessment. Routledge, 350 pp. https://doi.org/10.4324/9780429342585

Körner C, Paulsen J, Spehn EM (2011) A definition of mountains and their bioclimatic belts for global comparisons of biodiversity data. Alpine Botany 121: 73–78. https://doi.org/10.1007/s00035-011-0094-4

Luebert F, Pliscoff P (2018) Sinopsis bioclimática y vegetacional de Chile. 2nd ed. Editorial Universitaria Universidad de Chile, Santiago, CL, 384 pp.

Luteyn JL, Churchill SP, Griffin D, Gradstein SR, Sipman HJ, Gavilanes AMR (1999) Páramos: a checklist of plant diversity, geographical distribution, and botanical literature. The New York Botanical Garden Press, NY, US, 278 pp.

Martínez Carretero E (1995) La puna argentina: delimitación general y división en distritos florísticos. Boletín de la Sociedad Argentina de Botánica 31: 27–40.

Negret PJ, Allan J, Braczkowski A, Maron M, Watson JE (2017) Need for conservation planning in postconflict Colombia. Conservation Biology 31: 499–500. https://doi.org/10.1111/cobi.12902

Peyre G, Balslev H, Martí D, Sklenář P, Ramsay P, Lozano P, Cuello N, Bussmann R, Cabrera O, Font X (2015) VegPáramo, a flora and vegetation database for the Andean páramo. Phytocoenologia 45: 195–201. https://doi.org/10.1127/phyto/2015/0045

Peyre G, Balslev H, Font X, Tello JS (2019) Fine-scale plant richness mapping of the Andean páramo according to macroclimate. Frontiers in Ecology and Evolution 7: 377. https://doi.org/10.3389/fevo.2019.00377

Peyre G, Osorio D, François R, Anthelme F (2021) Mapping the páramo land-cover in the Northern Andes. International Journal of Remote Sensing 42: 7777–7797. https://doi.org/10.1080/01431161.2021.1964709

Pfeffer WT, Arendt AA, Bliss A, Bolch T, Cogley JG, Gardner AS, Hagen JO, Hock R, Kaser G, … Randolph Consortium (2014) The Randolph Glacier Inventory: a globally complete inventory of glaciers. Journal of Glaciology 60: 537–552. https://doi.org/10.3189/2014JoG13J176

Pomeroy J, Bernhardt M, Marks D (2015) Research network to track alpine water. Nature 521: 32–32. https://doi.org/10.1038/521032c

Rangel JO (2000) Colombia: diversidad biótica III: La región de vida paramuna. Editorial Universidad Nacional de Colombia, Bogotá, CO, 902 pp.

Rivas-Martínez S, Navarro G, Penas A, Costa M (2011) Biogeographic map of South America. A preliminary survey. International Journal of Geobotanical Research 1: 21–40. https://doi.org/10.5616/ijgr110002

Sabatini FM, Lenoir J, Hattab T, Arnst EA, Chytrý M, Dengler J, De Ruffray P, Hennekens SM, Jandt U, … Wagner V (2021) sPlotOpen–An environmentally balanced, open-access, global dataset of vegetation plots. Global Ecology and Biogeography 30: 1740–1764. https://doi.org/10.1111/geb.13346

Sklenář P, Hedberg I, Cleef AM (2014) Island biogeography of tropical alpine floras. Journal of Biogeography 41: 287–297. https://doi.org/10.1111/jbi.12212

Testolin R, Attorre F, Jiménez-Alfaro B (2020) Global distribution and bioclimatic characterization of alpine biomes. Ecography 43: 779–788. https://doi.org/10.1111/ecog.05012

Testolin R, Attorre F, Borchardt P, Brand RF, Bruelheide H, Chytrý M, De Sanctis M, Dolezal J, Finckh M, … Jiménez-Alfaro B (2021) Global patterns and drivers of alpine plant species richness. Global Ecology and Biogeography 30: 1218–1231. https://doi.org/10.1111/geb.13297

Ulloa UC, Acevedo-Rodríguez P, Beck S, Belgrano MJ, Bernal R, Berry PE, Brako L, Celis M, Davidse G, … Jørgensen PM (2017) An integrated assessment of the vascular plant species of the Americas. Science 358: 1614–1617. https://doi.org/10.1126/science.aao0398

Van der Maarel E (1979) Transformation of cover-abundance values in phytosociology and its effects on community similarity. Vegetatio 39: 97–114. https://doi.org/10.1007/BF00052021

Weigend M (2002) Observations on the biogeography of the Amotape-Huancabamba zone in northern Peru. The Botanical Review 68: 38–54. https://doi.org/10.1663/0006-8101(2002)068[0038:OOTBOT]2.0.CO;2

Whittaker RJ, Fernández-Palacios JM, Matthews TJ, Borregaard MK, Triantis KA (2017) Island biogeography: Taking the long view of nature's laboratories. Science 357: eaam8326. https://doi.org/10.1126/science.aam8326

E-mail and ORCID

Gwendolyn Peyre (Corresponding author, gf.peyre@uniandes.edu.co), ORCID: https://orcid.org/0000-0002-1977-7181
Daniel Montesinos (dbmtperu@gmail.com), ORCID: https://orcid.org/0000-0002-4439-5089

Daniela Giraldo (da.giraldo11@uniandes.edu.co), ORCID: https://orcid.org/0000-0002-2983-1348
Antonio Galán de Mera (agalmer@ceu.es), ORCID: https://orcid.org/0000-0002-1652-5931
Barbara Ruthsatz (ruthsatz@uni-trier.de)
Federico Luebert (fluebert@u.uchile.cl), ORCID: https://orcid.org/0000-0003-2251-4056
Marcela Ontivero (montivero@mendoza-conicet.gob.ar), ORCID: https://orcid.org/0000-0003-1888-3422
Nicolás García (ngarcia@uchile.cl), ORCID: https://orcid.org/0000-0001-9003-1510
Miguel Álvarez (kamapu@posteo.de), ORCID: https://orcid.org/0000-0003-1500-1834
Rosa Isela Meneses (rosaiselameneses11@gmail.com), ORCID: https://orcid.org/0000-0001-8779-2757
Pablo Lozano (pablo_lozanoc@yahoo.com), ORCID: https://orcid.org/0000-0003-0857-8141
Daniela León (daniela.leon@ut.ee), ORCID: https://orcid.org/0000-0002-7295-8256
Maximilian Weigend (mwei@uni-bonn.de), ORCID: https://orcid.org/0000-0003-0813-6650
Fabien Anthelme (fabien.anthelme@ird.fr), ORCID: https://orcid.org/0000-0001-6249-995X
Milagro Palma (milagro.albalina@up.ac.pa), ORCID: https://orcid.org/0000-0002-0380-750X
Carolina Rodríguez (dc.rodriguez13@uniandes.edu.co), ORCID: https://orcid.org/0000-0001-5373-7880

Supplementary material

Supplementary material 1
Summary of the 53 data sources included in the VegAndes database and their plot content (*.pdf)
Link: https://doi.org/10.3897/VCS.95750.suppl1

Supplementary material 2
Temporal distribution of the vegetation plots contained in VegAndes (5,340 plots) based on their source (publication year or dataset description). Plot numbers are expressed as independent values (gray) as well as cumulative values (orange) per decade (*.pdf)
Link: https://doi.org/10.3897/VCS.95750.suppl2