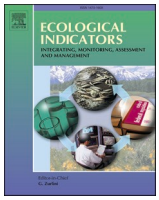
Ecological Indicators 183 (2026) 114669

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind



A global map of groundwater-dependent vegetation in the Mediterranean biome

Léonard El-Hokayem [a,c,*], Gabriella Damasceno [b,c], Francesco M. Sabatini [d], Helge Bruelheide [b,c], Gianmaria Bonari [e,f], Ciara Dwyer [g], Fernando Gonçalves [h], Borja Jiménez-Alfaro [i,j], Jonathan Millett [k], Josep Peñuelas [l,m], Jens-Christian Svenning [n], Jan Altman [o,p], Han Y.H. Chen [q,r], Tetiana Dziuba [s], Mohamed A. El-Sheikh [t], Behlül Güler [u], Daniel Hending [v], Bruno Hérault [w], Mohamed Z. Hatim [x,y], Florian Jansen [z], Carsten Meyer [c], Sharif A. Mukul [aa], Remigiusz Pielech [ab], Flávio Rodrigues [ac], Hua-Feng Wang [ad], Christopher Conrad [a,c]

[a] Institute of Geosciences and Geography, Martin Luther University Halle Wittenberg, Halle (Saale) 06120, Germany
[b] Institute of Biology / Geobotany and Botanical Garden, Martin Luther University Halle Wittenberg, Halle (Saale) 06108, Germany
[c] German Centre for Integrative Biodiversity Research (iDiv) Halle Jena Leipzig, Leipzig 04103, Germany
[d] BIOME Lab, Department of Biological, Geological and Environmental Sciences (BiGeA), Alma Mater Studiorum University of Bologna, Bologna 40126, Italy
[e] Department of Life Sciences, University of Siena, Siena 53100, Italy
[f] NBFC, National Biodiversity Future Center, 61 90133 Palermo, Italy
[g] Centre for Environmental and Climate Science, Lund University, 223 62 Lund, Sweden
[h] Department of Evolutionary Biology and Environmental Studies, University of Zurich, 8057 Zurich, Switzerland
[i] Biodiversity Research Institute (IMIB), University of Oviedo-CSIC-Principality of Asturias, Mieres, Spain
[j] Department of Organismal and Systems Biology, University of Oviedo, 33071 Oviedo, Spain
[k] Geography & Environment, Loughborough University, LE11 3TU Loughborough, United Kingdom
[l] CSIC, Global Ecology Unit CREAF-CSIC-UAB, Bellaterra, 08193, Catalonia, Spain.
[m] CREAF, Cerdanyola del Vallès, 08193, Catalonia, Spain
[n] Center for Ecological Dynamics in a Novel Biosphere (ECONOVO), Department of Biology, Aarhus University, Ny Munkegade 114, DK-8000 Aarhus C, Denmark
[o] Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, 165 00, Prague 6, Suchdol, Czech Republic
[p] Institute of Botany of the Czech Academy of Sciences, 252 43 Průhonice, Czech Republic
[q] Lakehead University faculty of natural resources management, P7B 5E1 Thunder Bay, Ontario, Canada
[r] Institute for Global Change Biology, and School for Environment and Sustainability, University of Michigan, Ann Arbor, MI 48109, USA
[s] M.G. Kholodny Institute of Botany, National Academy of Sciences of Ukraine, 01601 Kyiv, Ukraine
[t] Botany and Microbiology Department, College of Science, King Saud University, Riyadh 11451, Saudi Arabia
[u] Biology Education, Dokuz Eylul University, 35380 Izmir, Turkey
[v] Department of Biology, University of Oxford, OX1 3SZ Oxford, United Kingdom
[w] CIRAD, Forêts et Sociétés, F-34398 Montpellier, France
[x] Plant Ecology and Nature Conservation Group, Environmental Sciences Department, Wageningen University and Research, 6700, AA, Wageningen, Netherlands
[y] Botany and Microbiology Department, Faculty of Science, Tanta University, 31527 Tanta, Egypt
[z] Faculty of Agriculture, Civil and Environmental Sciences, University of Rostock, 18059 Rostock, Germany
[aa] Department of Environment and Development Studies, United International University, 1212 Dhaka, Bangladesh
[ab] Institute of Botany, Jagiellonian University in Kraków, 30-387 Kraków, Poland
[ac] Programa de Pós-graduação em Ecologia e Conservação (PPGEC/UNEMAT), Universidade do Estado de Mato Grosso, 78690-000 Nova Xavantina, Brazil
[ad] College of Tropical Agriculture and Forestry, Hainan University, 570228 Haikou, China

ARTICLE INFO

Keywords:
Biodiversity
Groundwater-dependent vegetation
Mediterranean
Phreatophytes

ABSTRACT

Groundwater-dependent vegetation (GDV) plays a vital role in maintaining biodiversity and ecosystem services in the Mediterranean biome but is increasingly threatened by climate and land use change. Large-scale GDV mapping in arid regions is critical for effective conservation and water resource management but remains challenging due to limited ground-truth data and the lack of high-resolution remote sensing-based spatial

* Corresponding author at: Institute of Geosciences and Geography, Martin Luther University Halle Wittenberg, Halle (Saale) 06120, Germany.
E-mail address: leonard.el-hokayem@geo.uni-halle.de (L. El-Hokayem).

https://doi.org/10.1016/j.ecolind.2026.114669
Received 12 May 2025; Received in revised form 19 November 2025; Accepted 27 January 2026
Available online 4 February 2026

Machine learning
Remote sensing
Terrestrial ecosystem

models. In this study, we mapped GDV across the world's five Mediterranean climate regions using a multi-level approach combining species-occurrence, vegetation-plot, and remote-sensing data. At the species level, we compiled a global list of phreatophyte and groundwater-associated species, which we used to extract occurrence records from GBIF. At the community level, we classified vegetation-plot data from the sPlot database based on phreatophyte presence and coverage, resulting in a unique ground-truth species-community dataset. At the biome level, we trained Random Forest models with eleven predictor variables in order to map GDV distribution at 30 m resolution for the period 2018–2023. We identified 482,000 km^2 of GDV, covering 28% of the study area and predicted GDV hotspots for the Western Iberian Peninsula, Southern France, California, Chile, and the West Coast of Australia. The largest absolute area of GDV was found in the Mediterranean Basin (306,000 km^2), while the highest relative coverage was found in California (42%) and Chile (40%). Notably, only about a quarter of Mediterranean GDV lies within protected areas. Key environmental predictors include soil sand content, dry season vegetation vitality, and elevation. By integrating species, community, and remote-sensing data, our high-resolution GDV map provides a crucial basis for monitoring ecosystem response to global change, conservation planning, sustainable groundwater management, and risk assessment in the drought-stressed Mediterranean biome.

1. Introduction

The rapid decline in groundwater levels due to climate and land use change is reducing groundwater availability for ecosystems worldwide, particularly in dry regions (Jasechko et al., 2024). As groundwater becomes scarcer, ecosystems that rely on this resource, such as springs, rivers, wetlands, lagoons, and terrestrial phreatophytic vegetation (Eamus et al., 2016; Barron et al., 2014; Kløve et al., 2011) are under increasing anthropogenic pressures, especially during droughts (Kløve et al., 2014). Protecting these groundwater-dependent ecosystems is crucial, as they support biodiversity and provide ecosystem services at a regional scale. Quantifying and managing these services, including climate regulation, water filtration and habitat provision (Saccò et al., 2023; Howard et al., 2023), requires harmonised large-extent fine-resolution maps of groundwater-dependent ecosystems (Link et al., 2023). This is particularly important for drought-stressed regions, such as the five regions of the world characterised by a Mediterranean climate, with highly seasonal patterns of hot, dry summers and mild, wet winters. These regions comprise the Mediterranean biome - a biodiversity hotspot, where 30–50% of the vegetation may rely on groundwater (El-Hokayem et al., 2023b; Evaristo and McDonnell, 2017; Milano et al., 2013; Soares and Lima, 2022).

Groundwater-dependent vegetation (GDV) comprises plant communities whose structure or functioning are influenced by groundwater availability. Within these communities, plants with access to groundwater, known as phreatophytes (Meinzer, 1927) often have an adaptive advantage during dry periods (Canadell et al., 1996) and serve as key indicators of groundwater influence (Box et al., 2022; Duran-Llacer et al., 2022). While groundwater dependence can also manifest through community composition or ecosystem processes such as soil moisture buffering and productivity stability, phreatophytes provide the most practical and scalable means for large-scale GDV mapping due to their detectability in vegetation inventories and remote sensing (Rohde et al., 2024a; Eamus et al., 2016). In arid and semi-arid environments, phreatophytes adopt different strategies depending on their groundwater access: facultative phreatophytes extract water from both the phreatic (saturated groundwater zone) and vadose (unsaturated soil above the water table) zones, whereas obligate phreatophytes must remain in constant contact with groundwater. This ability enables phreatophytes to decouple ecosystem net primary productivity from precipitation, resulting in increased growth and biodiversity compared to the surrounding vegetation, forming azonal vegetation (Glanville et al., 2023; Hultine et al., 2020).

Since GDV is crucial for biodiversity and ecosystem functioning, especially in drought-stressed environments, monitoring this vegetation is essential. However, effective monitoring, management, and protection of GDV is constrained by the lack of high-resolution maps. This issue is particularly pronounced in the Mediterranean biome, but it extends to GDV worldwide. Local-scale GDV mapping methods, such as species inventories or isotope analysis, provide on-ground assessments but are time- and labour-intensive (Jones et al., 2019; Killroy et al., 2008). To overcome these limitations and scale-up mapping efforts, the integration of spatial modelling with remote sensing and geospatial data has become increasingly important (Pérez Hoyos et al., 2016). These approaches integrate environmental variables, such as topography, soil properties, and climate, which control GDV occurrence (Bai et al., 2024; Martínez-Santos et al., 2021), with remote sensing products, either optical or radar. Nevertheless, while optical data has already been used to detect GDV as 'green, wet, and cool islands' during dry periods (Rohde et al., 2024b; Barron and Van Niel, 2009; Akasheh et al., 2008), the application of radar data, which enhances time series density and provides structural and moisture information on ecosystems, is lagging behind (Castellazzi et al., 2023).

Despite advances in remote sensing and geospatial data integration, large-scale GDV mapping faces several challenges. First, there is a lack of species-based community-level validation. The occurrence of phreatophyte species indicates GDV and is often implemented as training and validation data in remote sensing mapping approaches (e.g., Rohde et al., 2024b; Box et al., 2022; Duran-Llacer et al., 2022). However, phreatophyte species are usually identified through regional expert-based inventories (Lewis, 2011; Gomes Marques et al., 2019), which limits their applicability in large-scale mapping. Producing a consolidated list of phreatophyte species is therefore crucial for improving the identification of GDV from community data. Second, while current mapping efforts reveal broad spatial patterns of potential GDV, they dismiss small-scale GDV due to coarse spatial resolutions. For example, Doody et al. (2017) identified moderate groundwater-dependent ecosystems potential in 34% of Australia using an expert-based index integrating remote sensing and geospatial data. Link et al. (2023) produced the first global groundwater-dependent ecosystems potential map at 0.5° resolution, integrating groundwater discharge, xylem groundwater fraction, and potential inflow dependence, with 32% of arid regions potentially containing groundwater-dependent ecosystems. More recently, Rohde et al. (2024b) developed a mapping approach which identified global dryland groundwater-dependent ecosystems at 30 m resolution. However, many areas in the Mediterranean biome have remained unmapped, possibly due to the lack of accurate ground-truth vegetation data, as most previous mapping exercises have relied exclusively on species-level phreatophyte presence information, neglecting critical aspects related to species co-occurrences that can only be derived from vegetation plot data (Rohde et al., 2024b; Bruelheide et al., 2019). Furthermore, the potential of using citizen science data to increase the coverage of phreatophyte occurrence records, coupled with the use of radar imagery, has remained unexplored on large scales (Castellazzi et al., 2024).

In this study, we use a novel hierarchical approach to create the first high-resolution GDV distribution map across the Mediterranean biome. Our approach integrates species-level data, i.e., records of individual

phreatophyte species occurrence from GBIF, with community-level data, i.e., species co-occurrence and relative-occupancy data collected through vegetation plots from the sPlot database (Bruelheide et al., 2019), and biome-level data, i.e., remote sensing and other geospatial data. Our study is driven by three main objectives: I) Compiling a list of phreatophyte indicator species from literature to support global GDV identification; II) Applying this list to classify GDV within plant community databases, analyse spatial determinants of GDV distribution, and create a ground-truthed dataset for spatial modelling; and III) Training and validating a spatial model using optical and radar remote sensing imagery combined with geospatial data to map GDV distribution across the five regions of the Mediterranean biome.

2. Material and methods

2.1. Study extent

The Mediterranean biome is characterised by high seasonality with hot, dry summers and mild, wet winters, and contains 20% of global plant diversity (Pereira et al., 2004; Médail and Quézel, 1999). The Mediterranean climate is characterised by mean annual temperatures between 14 °C and 18 °C and mean annual precipitation between 400 mm and 1200 mm, with 65% of rain falling in winter (Pignatti, 2003). The biome is formed by five regions comprising six floristic realms (Liu et al., 2023): 1) the Mediterranean Basin (Holarctic, Saharo-Arabian), 2) the Chilean Matorral (Chile-Patagonia), 3) the California Chaparral (Holarctic, Neotropical), 4) the Western Cape of South Africa (African), and 5) Southwest Australia (Australian), and 40 ecoregions (Dinerstein et al., 2017). The zonal vegetation consists of evergreen and deciduous forests and shrublands dominated by sclerophyllous woody plants (Pignatti, 2003).

We excluded climates without an annual dry season (Beck et al., 2018) from the study, as these do not align with the characteristic Mediterranean climate. We mapped only (semi-) natural vegetation formed by vascular plants to ensure the focus remained on ecosystems rather than human-modified landscapes. Thus, pixels representing cropland, built-up areas, snow and ice, permanent water bodies, or mosses and lichens were excluded prior to modelling using the ESA WorldCover v200 map (Zanaga et al., 2022) and the Global Human Settlement map (Pesaresi and Politis, 2023). This resulted in a reduction of the total area by 45%, with the final study extent covering 1.7 million km^2. The GDV mapping workflow (Fig. 1) is described below.

2.2. Species-community dataset

First, to build a global species-community dataset for the Mediterranean biome, we conducted a systematic literature review to identify plant species that use groundwater in different environments. We used two search terms: 1) [phreatophyte] AND [species] (all fields) and 2) [“groundwater indicator”] OR [“groundwater dependent vegetation”] OR [“groundwater dependent ecosystem”] AND [plant] AND [species] (title, abstract, keywords) in the Web of Science database, which returned 839 publications (Web of Science, 2024, December). A Google query using the first term supplemented our results with additional grey literature. After screening titles and abstracts for relevance and species level data, we reviewed 91 full texts and compiled a list of global phreatophytes (Supplementary 2). Phreatophyte species were classified in two categories: either as obligate or facultative. When not specified,

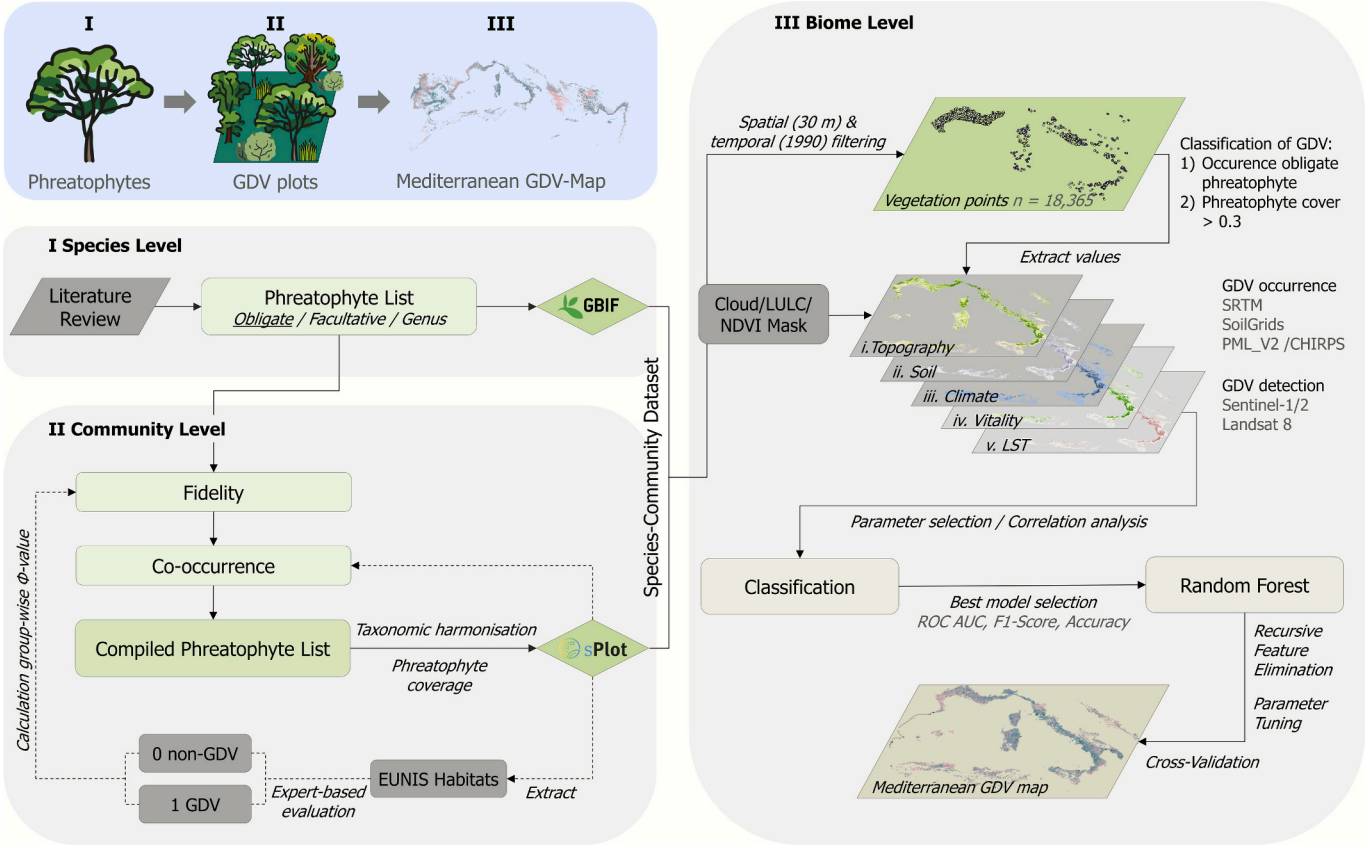


Fig. 1. Schematic overview of the mapping workflow. A simplified workflow is shown in the blue box. The detailed workflow includes all processing and analysing steps at: I) Species level, II) Community level, and III) Biome level. The study extent includes the five regions of the Mediterranean biome. All datasets used in the study are reported in Supplementary 1. Abbreviations: GDV - Groundwater-Dependent Vegetation, LULC - Land Use/Land Cover, NDVI - Normalized Difference Vegetation Index, LST - Land Surface Temperature. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

facultative category was assumed (592 species) (Hultine et al., 2020). We then standardised taxonomic names using the 'TNRS' package in R (Maitner et al., 2024), removed duplicates and hybrids, and assigned growth forms and IUCN Red List categories using the 'rredlist' package (Gearty and Chamberlain, 2024).

Second, we integrated 187,261 plant community records within the Mediterranean biome from sPlot (Bruelheide et al., 2019), expanding previous groundwater-dependent ecosystem mapping exercises only relying on the sPlotOpen dataset (Rohde et al., 2024b), which is an environmentally-balanced subset of sPlot (Sabatini et al., 2021). To extend our literature-based phreatophyte list, we calculated the phi coefficient (Φ - a measure of species fidelity) (De Cáceres et al., 2008; Bruelheide, 2000) to derive the level of association of other species with GDV. We based this calculation on the co-occurrence analysis for the 3.7 million species occurrences extracted from sPlot. To do so, we first reclassified 106 EUNIS habitats associated to the 132,359 vegetation plots in Mediterranean Europe into GDV and non-GDV groups based on habitat descriptions and species occurrences (Chytrý et al., 2020; Davies et al., 2004). We excluded habitats classified at the first EUNIS hierarchical level (e.g., grasslands, woodlands, and forests) and anthropogenic vegetation (e.g., plantations and croplands) from the Φ calculation. We calculated Φ from a contingency table that summarises the occurrences of a species in a target site group (GDV) versus its occurrences in another site group (non-GDV) (Chytrý et al., 2002) using the 'indicspecies' package in R (De Cáceres and Legendre, 2009) and considered only species with Φ above the median as groundwater-associated species. We note that, although the assignment of vegetation plots to EUNIS habitats relied on the overall species composition of the plot, this approach aimed to identify species which are indicators of groundwater-associated habitats, to expand our preliminary list. This was therefore an iterative process, where habitats were classified at the coarser EUNIS level based on a qualitative ecological interpretation, and phi calculations and thresholds were then applied at the plot level as a quantitative step. We then evaluated the extracted species in terms of distribution and habitat and excluded two generalist taxa (*Rubus vagabundus* and *Taraxacum sect. vulgaria)* and *Cydonia oblonga* as they do not indicate groundwater usage. We further expanded our checklist by identifying species that co-occur with phreatophytes in at least 60% of their occurrences. Frequent co-occurrence with known phreatophytes is interpreted as evidence of ecological association with GDV, rather than obligate physiological dependence. Hence, all newly identified species were ultimately treated as facultative phreatophytes. Based on the compiled list, we then calculated plot-wise phreatophyte coverage from relative species abundance for all sPlot vegetation plots and filtered records for location uncertainty ($\leq$ 30 m) and date (from 1990).

Third, to address underrepresented regions in the sPlot data, specifically in California, Western Australia, and Central Spain, and to increase the number of sites representing GDV, we extracted occurrence records for obligate phreatophytes identified in our list (shrubs and trees) from GBIF using the 'rgbif' package in R (Chamberlain et al., 2024; Chamberlain and Boettiger, 2017). As obligate phreatophytes are in constant contact with groundwater, their presence is a strong indicator of GDV. Combining GBIF occurrence point records with standardised sPlot vegetation plots allowed us to expand both the spatial coverage and the number of GDV sites. We downloaded occurrence records from 1990 onward with a spatial accuracy of at least 30 m, retaining only records without geospatial issues (GBIF.org, 2024). In effect, both datasets were treated as presence-only, confirming the occurrence of relevant species at each location. This made the integration of occurrence records (GBIF) and community survey plots (sPlot) a robust approach to expand coverage despite differing survey designs. Initial data cleaning removed invalid, rounded and zero coordinates and records within a 2000 m buffer of country and province centroids, capitals, and biodiversity institutions using the 'CoordinateCleaner' package in R (Zizka et al., 2019). To account for spatial uncertainties in the GBIF records we conducted manual data cleaning. In California, we first selected records within known groundwater-dependent ecosystems (California Department of Water Resources, 2022). Based on visual interpretation of high-resolution true-colour satellite imagery (Earthstar Geographics, 2024; Google, 2024), we then assessed and adjusted the spatial reliability of the records. Taking into account landscape and land cover, all records were flagged according to the following categories: i) RES: located in a vegetation unit that is smaller than the target resolution of 30 m, ii) AL: located in anthropogenic landscapes, such as urban or agricultural areas, iii) VC: located at forest edges or in bare ground river banks where vegetation is close, iv) W: located in water bodies, v) GOOD: located inside natural, large vegetated areas. Unsuitable records were removed from the dataset. About 50% of the records where the original location did not accurately reflect vegetation were shifted to a more suitable location within the closest vicinity (median 17 m; max 500 m).

Finally, we merged 15,155 sPlot vegetation plots with 3201 GBIF occurrence records to build our training and validation dataset (species-community dataset). We applied two rules to classify GDV points: i) occurrence of at least one obligate phreatophyte, or ii) total facultative phreatophyte cover exceeding predefined thresholds at 0.25, 0.3, 0.4, 0.5, and 0.75. The quartile-based thresholds (0.25, 0.5 and 0.75) being conventional benchmarks for classification, while additional intermediate values (0.3, 0.4) were tested after preliminary modelling indicated reduced accuracy at 0.5 (see Section 2.4 and Supplementary 3 for details). Regional threshold testing was not feasible due to uneven data distribution and gaps in the sPlot data. All other points were classified as non-GDV, resulting in a binary classification for modelling.

2.3. Remote sensing and geospatial data

2.3.1. Variables controlling GDV occurrence

We included three variables to account for topographically driven water accumulation that determines GDV occurrence (Rampheri et al., 2023; Münch and Conrad, 2007): elevation, slope (30 m resolution), and the multiscale topographic position index (270 m resolution) (Theobald et al., 2015). These were all derived from the SRTM digital elevation model (Farr et al., 2007). Then, we calculated mean sand content (250 m resolution) from three topsoil layers (0–30 cm) as a proxy for water holding capacity (Martínez-Santos and Renard, 2020) using the SoilGrids250 2.0 dataset (Poggio et al., 2021). To incorporate climate, we developed a novel approach (Eq. 1) to calculate potential inflow dependence (pIDE) (Doody et al., 2017) from daily CHIRPS precipitation data (Funk et al., 2015):

$$pIDE = transpiration/(precipitation - interception) \quad (1)$$

Interception and transpiration rates were derived from the PML_V2 product (Zhang et al., 2019) at 500 m resolution. We then calculated the mean pIDE from annual sums for the period 2018–2023. Pixels where transpiration rates exceed the actual precipitation reaching the ground are more likely to have an external water resource and contain GDV (Doody et al., 2017).

2.3.2. Variables controlling GDV detection

We followed the 'green island' concept to delineate GDV from surrounding non-GDV (Akasheh et al., 2008). To identify these groundwater-dependent 'green islands' we considered GDV to show: 1) high vitality during the annual dry season, 2) low interannual, and 3) low seasonal changes in vitality (Gou et al., 2015), and 4) higher vitality compared to surrounding land cover. Two vegetation indices targeting plant greenness and density were calculated to account for vegetation vitality: i) the Normalized Difference Vegetation Index (NVDI) from optical imagery and ii) the Radar Vegetation Index (RVI) from radar imagery. The calculation of the indices includes Sentinel-1&2 data at 30 m resolution, and combines both spectral and structural characteristics of vegetation, crucial for GDV detection (Castellazzi et al., 2024; Schulte Bühne and Pettorelli, 2018).

We calculated mean NDVI (Rouse et al., 1974; Eq. 2) during defined climatic dry seasons (El-Hokayem et al., 2023c) for the period 2018–2023 in Google Earth Engine (Gorelick et al., 2017) using Sentinel-2 scenes with cloud cover less than 20%. Cloudy pixels were masked using the QA60 band.

$$NDVI = (NIR - Red)/(NIR + Red) \quad (2)$$

The RVI measures randomness of radar scattering, and thus works as a proxy for vegetation cover and water content (Castellazzi et al., 2024; Castellazzi et al., 2023). Higher RVI values indicate tall and dense vegetation, as higher structural complexity causes radar signals to scatter in many directions (Szigarski et al., 2018; Kim et al., 2012). Sentinel-1 radar images were selected for the respective dry seasons from 20,182,023. Pre-processing of the images includes terrain flattening and mono-temporal speckle filtering as described by Mullissa et al. (2021). RVI was calculated as the ratio of co-polarisation (VV) and cross-polarisation (VH) bands (Charbonneau et al., 2005), described in Eq. 3.

$$RVI = (4*VH)/(VV + VH) \quad (3)$$

To account for low interannual changes in plant vitality, we calculated the mean standard deviation for NDVI and RVI during the annual dry season between 2018 and 2023. Seasonal change in vitality was calculated as the mean annual standard deviation of NDVI for the same period (Liu et al., 2021). Border noise removal could not be applied to Sentinel-1 images, resulting in stripe artefacts (Ali et al., 2018) for seasonal RVI changes, which were therefore excluded from the final model.

'Cool islands' (Barron and Van Niel, 2009) were detected by lower Land Surface Temperature during the dry season compared to the surrounding land cover (Rohde et al., 2024b). We calculated spatial Land Surface Temperature anomalies during the dry season from Landsat-8 as described by Rohde et al. (2024b). Finally, Land Surface Temperature anomalies were generated by averaging the differences in Land Surface Temperature between a focal pixel and the surrounding areas (270 m^2, 2700 m^2, and 5400 m^2) and then calculating the mean temperature anomaly between 2018 and 2023. The same workflow was applied to calculate spatial anomalies for NDVI and RVI during the dry season.

As the remote sensing metrics can be influenced by shading or inference of other mesic vegetation, we included environmental variables controlling GDV occurrence in the first place. A justification for the choice of predictor variables is provided in Supplementary 1. Prior to model selection and implementation, all predictor variables were harmonised in terms of extent, spatial resolution (30 m) and projection (WGS 1984, EPSG:4326) and merged into a single multiband image. All pixels with mean NDVI values <0.1 during the dry season were excluded from the analysis as they do not represent vegetation.

2.4. Modelling GDV distribution

We evaluated 14 different classification models to predict GDV and non-GDV classes using the 'caret' package in R (Kuhn, 2008). First, we extracted the 13 predictor variable values for all points in our species-community dataset, selected one unique point per 30 m pixel and excluded points with no data, resulting in a final training and validation dataset of 18,365 points. As the Pearson correlation was below 0.6 for all predictors at 500,000 random points across the study extent, all variables were included in the models. Second, we trained the models, splitting the data in 80% training and 20% testing data and using 10-fold cross-validation to minimise overfitting and ensure model stability. Third, we evaluated model performance based on area under the receiver operating characteristic curves (ROC-AUC), F1-score, and balanced accuracy to find the best model (detailed model performance for all models can be found in Supplementary 3). All models were trained five times with different thresholds for phreatophyte coverage, resulting in 70 runs in total. A Random Forest classifier (Breiman, 2001) where phreatophyte cover >0.3 represented GDV outperformed all other classifiers with a ROC-AUC of 0.84, a F1-score of 0.74, user accuracy for GDV of 0.79 and a balanced accuracy of 0.76 and was therefore selected for spatial modelling.

Prior to model implementation, we performed Recursive Feature Elimination using the 'caret' package in R (Kuhn, 2008), and thus excluded interannual and spatial RVI anomalies from the predictors. A Wilcoxon test was used to assess significant differences between non-GDV and GDV for each predictor. We then trained and optimised several Random Forest classifiers in Google Earth Engine using the 18,365 ground-truth points and eleven final predictor variables as input data to predict GDV. First, the species-community dataset was divided into training (80%) and testing (20%). Second, hyperparameter tuning was implemented to find the best fitting model based on accuracy. The final Random Forest model was built using 130 trees, 3 variables per split, one minimum leaf population, 0.9 bag fraction and the maximum number of nodes available. The out of bag error for the final model was 0.27 and the output included a binary classification of GDV and non-GDV, and the probability of each pixel (0–100%) being GDV.

2.5. Cross-validation

We used three cross-validation (CV) strategies to validate the final model. Standard 10-fold CV randomly splits the data into ten equal folds, using one for testing and the rest for training (Meyer et al., 2018). However, this approach may overfit ecological data due to spatial and temporal autocorrelation (Mushagalusa et al., 2024; Roberts et al., 2017). To address this, we applied block CV, creating folds based on environmental (region-wise) and spatial (grid-wise) patterns (Roberts et al., 2017; Valavi et al., 2019), always excluding one block for testing from model training. To prepare the species-community dataset for CV, we extracted ecoregions (Dinerstein et al., 2017) and assigned groundwater landscapes to each point of the dataset. These groundwater landscapes relate topography and the location of water bodies to describe typical groundwater interaction with vegetation. Five groundwater landscapes were defined hierarchically based on global datasets: 1) wetlands (Tootchi et al., 2019); 2) riverine landscapes (max 200 m distance from rivers) (Lin et al., 2021) with slope < 10; 3) coastal areas (max 1000 m distance from coastline) (Wessel and Smith, 1996) and elevation <10 m a.s.l.); 4) terrestrial lowlands (elevation ≤ 200 m a.s.l.), and 5) terrestrial uplands (elevation >200 m a.s.l.). CV was performed using 35 ecoregions, five groundwater landscapes, and a 250 km grid. The Random Forest model was trained 153 times, and the final performance was assessed using overall testing accuracy and user accuracy for GDV (true positives/total number of points classified as GDV).

3. Results

3.1. Global phreatophyte checklist

The literature review identified 1021 phreatophyte species, of which 265 were classified as obligate phreatophytes and 761 as facultative phreatophytes and six genera (*Phragmites*, *Populus*, *Prosopis*, *Salicornia*, *Salix* and *Tamarix*). An additional 38 species were selected based on phi calculation, and 12 species that frequently co-occur with phreatophytes within the dataset. In total, 1071 phreatophyte indicator and groundwater-associated species were compiled in a species checklist, of which 574 occur within the Mediterranean biome subset of the sPlot dataset. The five publications from which most species were extracted were Londo (1975), Evaristo and McDonnell (2017), Thomas (2014), Fan et al. (2017) and Chikishev (1965). Detailed results of the literature review and the phreatophyte list can be found in Supplementary 2.

3.2. Environmental characteristics of GDV occurrences

The phreatophyte list was used to classify the species-community

dataset resulting in 8950 non-GDV and 9415 GDV ground-truth points. Most points represented terrestrial vegetation (63%), followed by wetlands (27%), and riverine landscapes (10%). GBIF-derived phreatophyte occurrence points dominated in California, in the savannas and woodlands of Southwest Australia, and in the *Acacia-Argania* forests of Morocco. In Chile, South Africa, Southeastern Australia, North Africa, and the Middle East, the majority of ground-truth points were non-GDV, derived from sPlot vegetation data. However, points classified as GDV were more common in Iberian, Southern France, and Italian forests (Fig. 2).

Except seasonal changes in NDVI, all predictor variables extracted at the 18,365 ground-truth points showed highly significant differences. The effect size (r) was small for all variables, except for dry season spatial NDVI anomalies, where it was medium (Fig. 3). GDV sites are more often located in lowland valleys with flatter slopes, with a median elevation of 245 m a.s.l., which is 137 m lower than the median for non-GDV sites. Non-GDV sites are more often associated with very sandy soils (sand content >65%), although the median sand content for both classes is similar at 32–33%. Nevertheless, GDV sites are more often found on soils with a sand content between 40% and 60%. The pIDE values are only slightly higher for GDV sites, although there is an increase in the presence of GDV above the threshold of 1, where transpiration exceeds precipitation. Moreover, the median NDVI for GDV sites is 0.67 compared to 0.56 for non-GDV sites, whereas GDV sites are 0.7 °C colder at the median, underpinning the concept of GDV forming 'green and cool islands'. Seasonal changes in vitality show less pronounced differences between the two classes.

3.3. Global occurrence of Mediterranean GDV areas

We mapped 482,000 km^2 of likely GDV (28% of the analysed pixels) across the five regions of the Mediterranean biome, with 306,000 km^2 located within the Mediterranean Basin. The highest proportion of GDV was found in California (42%) and Chile (40%). As much as 90% of all GDV represented terrestrial vegetation, while wetlands and riverine landscapes accounted for 5% and 4%, respectively (Supplementary 3). The map shows large areas of likely GDV in the Western Iberian Peninsula, Southern France and Corsica, California, Southern Mediterranean Chile and the West Coast of Australia. Large non-GDV areas were found in North Africa, the Middle East region and inland Australia (Fig. 4). Currently, only about 25% of Mediterranean GDV lies within protected areas (UNEP-WCMC and ICUN, 2025), underscoring the need to integrate GDV mapping into conservation planning.

The detailed maps highlight the diversity in modelled and observed GDV in different Mediterranean landscapes. In the "Laguna de Ruidera" in Central Spain (Fig. 4a), GDV is embedded in agricultural landscapes and typical xerophytic macchia vegetation, forming rather isolated ecosystems. In the sparsely vegetated dry woodlands and steppes ecoregion in Morocco, GDV is associated with oases, where surface water is also available (Fig. 4b), or with phreatophytic steppe vegetation, where small patches of vegetation or even single woody individuals are dependent on groundwater (Fig. 4c). Larger areas of likely GDV habitats dominate in the "Cilento and Vallo di Diano National Park" in Southern Italy (Fig. 4d), where oak forests grow on shallow groundwater tables. In valleys where rivers are dry during summer droughts, riverine GDV formed by *Fraxinus angustifolia*, *Populus*, and *Salix* species are dominant, while non-GDVs were found at higher elevations on calcareous rocks with deep groundwater circulation.

The final Random Forest model had a testing accuracy of 0.75. The user accuracy for GDV (true positives) was 0.79 and for non-GDV (true negatives) 0.72. The three most important predictor variables were sand content, dry season spatial NDVI anomalies, and elevation. Standard K-fold CV showed stable model performance, with accuracies ranging between 0.73 and 0.77. The accuracy varied across Mediterranean ecoregions (Fig. S3.5), with high accuracy (0.73–0.97) in the Iberian, Aegean and Western Turkish forests, Chile, and California. Moderate accuracy (0.61–0.7) was achieved in North Africa, Southeastern Australia, Southwestern Iberian and Tyrrhenian-Adriatic forests. Lower accuracy (0.17–0.48) was found in the Middle East, Italian forests, Pindus Mountain forests and Southwestern Australian forests. CV using

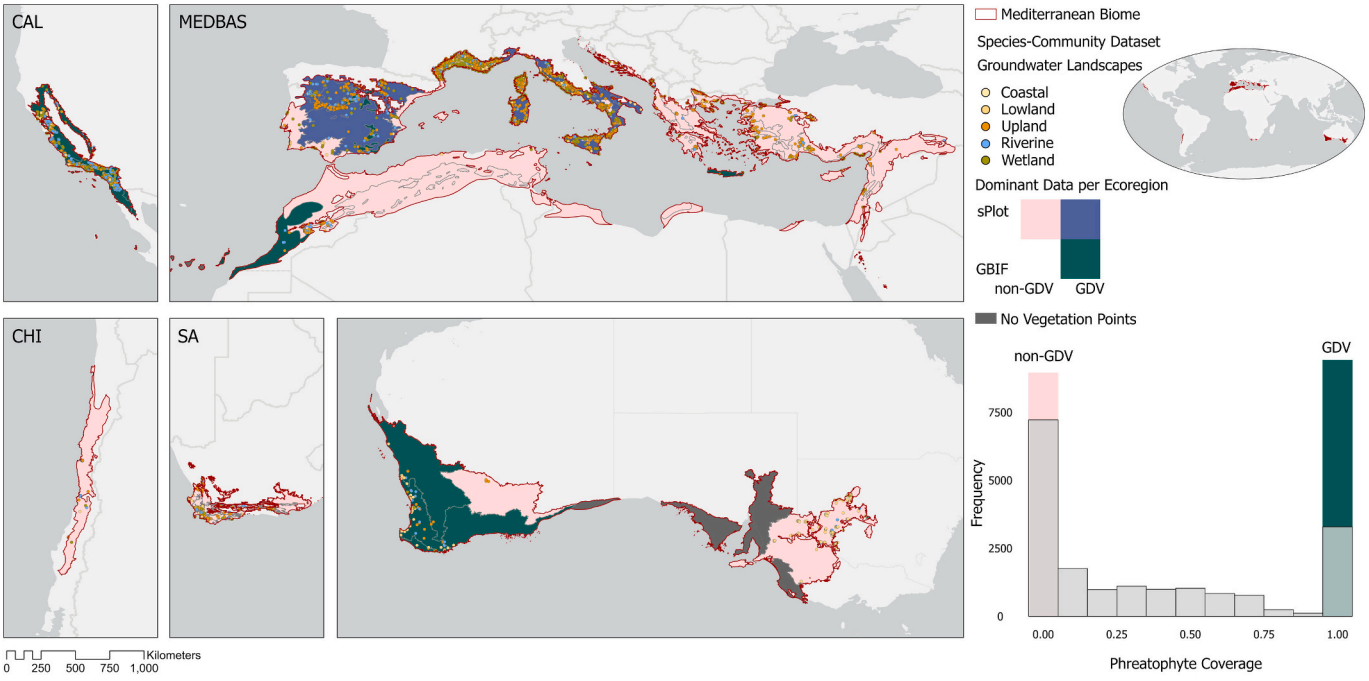


Fig. 2. Distribution of species-occurrence records (GBIF) and vegetation plots (sPlot) used as training and validation data to map GDV across 40 ecoregions in the Mediterranean biome. Ecoregions are categorised based on dominant (>50% of points) data source and non-GDV/GDV classes. Ecoregions where non-GDV points from sPlot dominate are shown in pink, those where GDV points from GBIF dominate in teal, and those where GDV points from sPlot dominate in purple. From GBIF no non-GDV plots were extracted. Abbreviations: CAL - California, MEDBAS - Mediterranean Basin, CHI - Chile, SA - South Africa, AUS - Australia, GDV - groundwater-dependent vegetation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

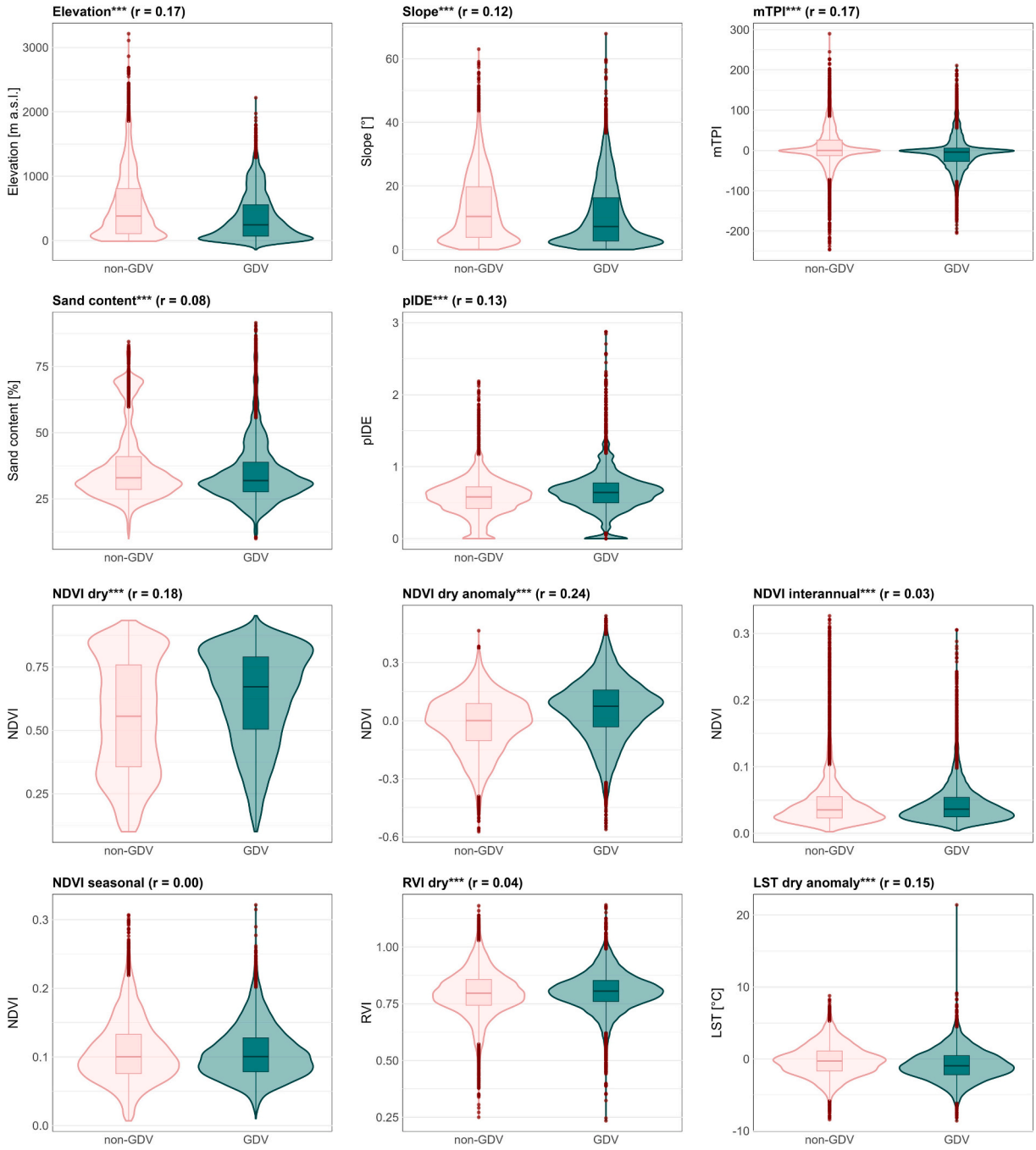


Fig. 3. Value distribution for eleven predictor variables extracted at the 18,365 GDV and non-GDV sites of the species-community dataset. A Wilcoxon test was performed to test for statistically significant differences (asterisks) and effect size (r) between the two classes. Red dots represent outliers in the data and the violin plots show the data distribution per class and variable. Abbreviations: GDV - Groundwater-Dependent Vegetation, NDVI - Normalized Vegetation Index, RVI - Radar Vegetation Index, LST - Land Surface Temperature, mTPI - Multiscale Topographic Position Index, pIDE - potential Inflow Dependence. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

groundwater landscapes performed well for coastal, lowland, riverine, and wetland sites. However, accuracy was only moderate when upland sites were excluded (Table S3.9). Detailed results for the Random Forest model and CV are presented in Supplementary 3. To address regional variability (over/underestimation) and better capture transitional ecosystems, we provide a GDV probability map (0–100%) that enables users to test different classification thresholds. Additionally, we provide a GDV map generated at a less conservative 0.4 phreatophyte coverage threshold (overall GDV coverage 21%; user accuracy GDV 0.72). We also provide data and code to test ecoregion-specific calibration.

4. Discussion

We present the first high-resolution GDV map for the Mediterranean biome. This map shows that GDV covers 482,000 km^2 across the entire Mediterranean biome, with highest proportions in California (42%) and Chile (40%). Our map is based on the first global list of phreatophyte indicator species (1071 species) and six genera, derived from literature and vegetation analysis. From this checklist, our ground-truthed dataset of 18,365 points across the Mediterranean biome, indicates that GDV is generally densely vegetated, more drought-resilient, and located in lowland areas where water accumulates. This new understanding of the

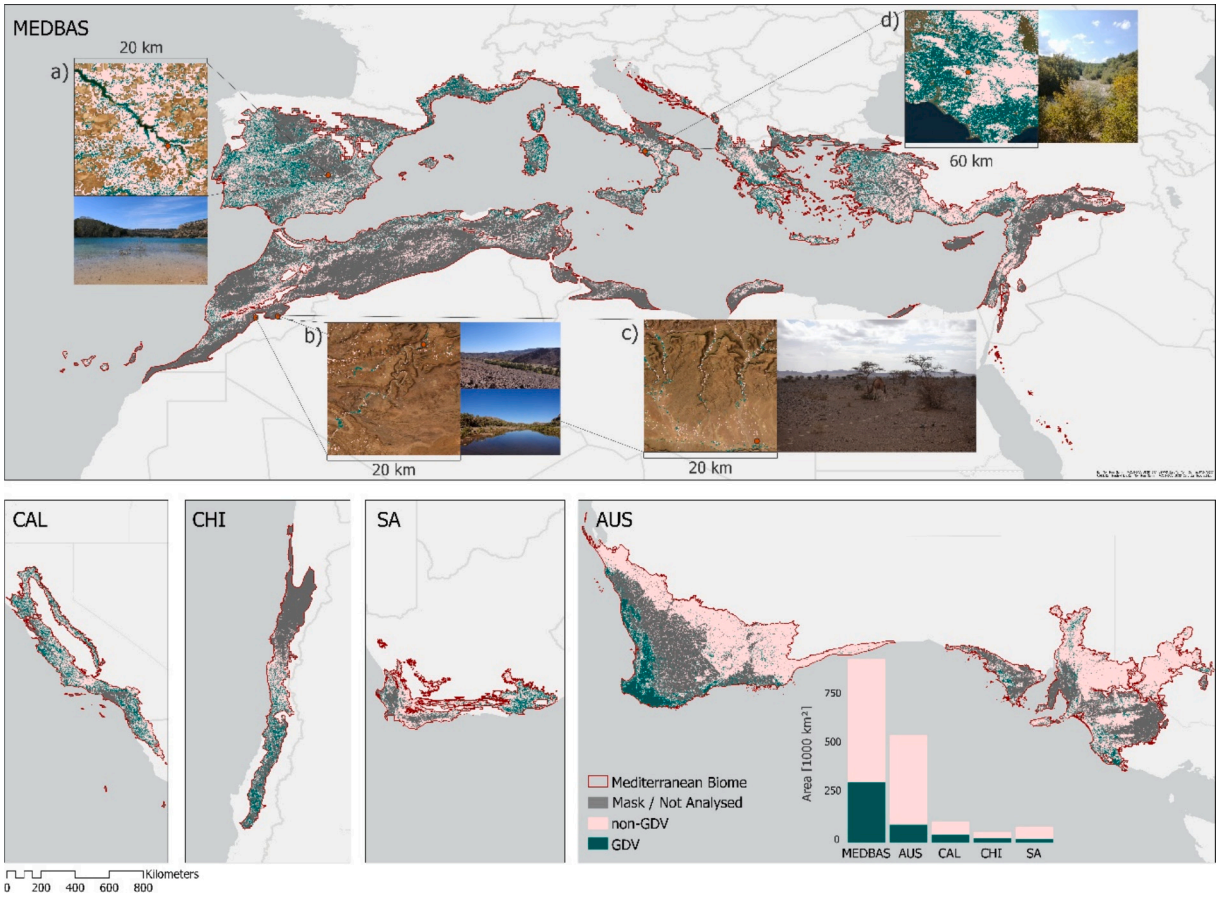


Fig. 4. Distribution of likely GDV and non-GDV in the Mediterranean biome, along with four local examples from the Mediterranean Basin. a) "Laguna de Ruidera" (Ramsar site) with typical phreatophytic vegetation including *Juncus maritimus*, *Carex* spp., *Molinia caerulea*, *Populus* spp. and *Phragmites australis*. The lagoon provides habitat for migratory birds and is a biodiversity hotspot; b) "Oasis de Fint" with *Phoenix dactylifera* used for the oasis economy by local communities; c) "Ougnat Massif" with sparse phreatophytic steppe vegetation such as *Tamarix* spp., *Acacia* spp. or *Retama raetam,* providing food for animals such as dromedaries; and d) "Cilento and Vallo di Diano National Park" with typical riverine GDV including *Populus nigra, Salix alba*, and *Fraxinus angustifolia*, as well as large undisturbed *Quercus* forests on shallow groundwater tables. Full-scale interactive maps are accessible at: https://geoecohalle.users.earthengine.app/view/mediterranean-gdv including all predictor variables, GDV classes and probability maps for different phreatophyte coverages and assigned groundwater landscapes. Abbreviations: CAL - California, MEDBAS - Mediterranean Basin, CHI - Chile, SA - South Africa, AUS - Australia, GDV - Groundwater-Dependent Vegetation.

global extent and distribution of GDV in the drought-stressed Mediterranean biome provides a unique resource for monitoring ecosystem response to global change, prioritising areas for accurate local-scale assessments, guiding conservation efforts, and informing sustainable groundwater management.

Our GDV hotspots coincide with many biodiversity and geodiversity hotspots such as the California Floristic Province, Central Chile, the Mediterranean Basin, Southwest Australia, and the Eastern Cape in South Africa (Polman et al., 2024; Myers et al., 2000). Here, GDV plays a critical ecological role by sustaining high biodiversity, including rare and endemic species, especially in Southwest Australia and South Africa, and by acting as thermal and hydrologic refugia during dry seasons and climatic extremes (Glanville et al., 2023; Dinerstein et al., 2017; Kløve et al., 2014). In the Mediterranean, where climate change impacts are amplified in lowlands (Saatkamp et al., 2023), GDV often occupies narrow, fragmented habitats such as riparian corridors and springs. These habitats act as stepping-stones that maintain ecological connectivity, as seen for example in the Ebro Delta (Spain) and Camargue wetlands (France), which provide critical refugia for waterbirds (Ramsar, 2025).

We produced a global checklist of phreatophyte species and genera, which represents a key asset in future GDV mapping and analyses. Yet, compiling this list did not come without difficulties, as different criteria for identifying phreatophytes might yield different results. For example, Rohde et al. (2024b) identified *Abies alba* and *Achillea millefolium* as common phreatophytes in global drylands based on expert assessment, but our literature review and vegetation analysis did not support this. Similarly, we derived 340 phreatophytes from a list developed for the Netherlands (Londo, 1975), where the water table is shallow (< 2 m for most of the country) (de Vries, 2007). However, forest species such as *Aconitum vulparia*, *Circaea alpina,* and *Lathraea squamaria*, which can access groundwater in the Netherlands, may not be dependent on it in regions with deeper water tables. Although phreatophytes form azonal vegetation driven by hydrogeology, their groundwater use varies with location (Thomas, 2014; Ahring and Steward, 2012). A species may be facultative in one area but obligate in another, or switch to soil moisture depending on local conditions (Hultine et al., 2020; Orellana et al., 2012). These variations complicate classification and highlight the need for site-specific assessments. Integrating plant traits, such as rooting depth or indicator values, into expert systems (Chytrý et al., 2020) could strengthen large-scale GDV classifications (El-Hokayem et al., 2023a).

Groundwater access likely decouples GDV productivity from precipitation-driven dynamics. As a result, GDV sites appear more vital and colder than surrounding areas, forming 'green and cool islands' in the landscape (Barron and Van Niel, 2009; Akasheh et al., 2008). Our data show that GDV sites in the Mediterranean biome are common on sandy soils at lower elevations, where water accumulates and sustains vegetation during the dry seasons, validating common assumptions

applied for GDV detection (Eamus et al., 2016; Gou et al., 2015). However, soil-groundwater-vegetation interactions can be complex. For example, sandy soils allow for faster groundwater infiltration and easier root penetration, while soils with high clay and silt content promote capillary rise (Schulte Bühne and Pettorelli, 2018). Nevertheless, sandy soils have lower water holding capacity, so that vital vegetation is more likely to depend on groundwater rather than soil moisture during dry seasons (El-Hokayem et al., 2023a). That said, sandy soils are not necessarily required for GDV occurrence. Instead, the observed patterns may also reflect broader landscape settings, such as coastal plains and river valleys, rather than strict edaphic controls (Gomes Marques et al., 2019). GDV is predominantly found at lower elevations (Saccò et al., 2023; Doody et al., 2017; Gou et al., 2015), a pattern reflected by the high variable importance of elevation in model predictions. However, this finding may partly reflect sampling bias as most of the GDV ground-truth points are concentrated at lower elevations. This may lead to an underrepresentation of GDV at higher elevations, where snowmelt and fractured or weathered hard-rock aquifers can sustain riparian vegetation or meadows (Cabello et al., 2026; Lowry et al., 2011).

Our study found an overall likely GDV coverage of 28%, similar to the 31% reported by El-Hokayem et al. (2023b), but lower than the 50% in a recent meta-analysis (Evaristo and McDonnell, 2017). While the meta-analysis estimates plant groundwater use, our approach focuses on identifying persistent GDV areas, which may explain the differences. Observed GDV occurrence patterns align with previous studies (e.g., Rohde et al., 2024b; Doody et al., 2017) and support the idea that GDV forms distinct ecological units shaped by groundwater availability and access (Barron et al., 2014). Comparison with Rohde et al. (2024b) is challenging since their GDEs include waterbodies. Still, broad-scale patterns align: e.g., groundwater-dependent *Eucalyptus* forests in Western Australia, extensive non-GDV areas east of Adelaide, GDV patches north of the Atlas Mountains, the Camargue wetlands in Southern France, coastal Croatia and a north-south gradient in Portugal. Differences appear in Southeastern Anatolia, where Rohde et al. (2024b) map extensive GDEs including dry grasslands on rocky terrain, while our map classifies GDV mainly along waterbodies, such as the Euphrates and Tigris rivers. On Mount Etna (Italy), they identify GDV, but the vegetation there likely depends on seasonal soil moisture in volcanic substrates, which can mimic GDV signals (El-Hokayem et al., 2023a). Agreement with the Mediterranean GDV potential map (El-Hokayem et al., 2023b), suggests that spatial patterns may arise from variables controlling the occurrence of GDV, such as sand content and elevation, rather than detection. These predictors highlight the importance of geodiversity, where variations in soil, topography, and geology influence groundwater availability for GDV (Brilha, 2014). Certain environmental conditions support GDV occurrence even though GDV is not present. However, regional mismatches, such as in Western Australia and South Africa, show the importance of high-resolution detection to capture local variation.

Cross-validation results reveal regional differences in model uncertainty and reliability, influenced by training data availability, Random Forest overfitting, and GDV characteristics (Barreñada et al., 2024). Accuracy is highest in regions where large, uniform GDV were mapped, such as Chile, parts of California, and Southwest Australia. However, in California, the model tends to overpredict GDV compared to regional datasets (California Department of Water Resources, 2022), likely due to the inclusion of GDV-only GBIF records and a conservative phreatophyte coverage threshold (> 0.3). Reliability decreases in transitional zones, which are areas that represent both ecological and climatic gradients. Examples include the transition from semi-arid to arid regions, or from and semi-arid to temperate regions in the Eastern Mediterranean Basin, or France, and oasis-desert-river transitional zones in North Africa. These zones often exhibit intermediate GDV probabilities, reflecting the intersection of biome transitions and GDV/non-GDV boundaries, thereby amplifying classification uncertainty. Accuracy is also lower in fragmented landscapes like Italy and the Middle East, where GDV can be small, scattered or share spectral and environmental characteristics with irrigated agriculture or mesic vegetation on high moisture soils (El-Hokayem et al., 2023a). Identification is further complicated in arid regions, where GDV often consists of sparse, ephemeral vegetation dependent on fluctuating groundwater (Abdullah et al., 2024; Wang et al., 2011). Under-sampled regions, such as North Africa, the Eastern Mediterranean, Chile, and coastal or mountainous areas reduce model reliability because the full range of GDV types could not be captured. These imbalances introduce both geographic and ecological biases, potentially limiting the ability to generalise across the full range of GDV types in the Mediterranean biome. Although the Random Forest model can capture regional patterns, it cannot fully compensate for data gaps or ecological variation. The provided GDV probability map can be used to further explore transitional zones and areas of higher uncertainty.

The integration of NDVI anomalies and radar data allowed our model to provide a more accurate assessment of vegetation vitality during the dry season, which represents a substantial improvement in GDV mapping. Considering spatial NDVI anomalies, in particular, allows a more precise detection of 'green islands' (Akasheh et al., 2008) during dry seasons by decoupling absolute NDVI values from regional and species-specific variability. For example, groundwater-dependent *Phoenix dactylifera* habitats in Morocco show lower NDVI values (0.6) compared to groundwater-dependent *Quercus* forests in Southern Italy (0.8) (Fig. 4). However, as semi-arid GDV often occurs in sparsely vegetated environments or is associated with surface water (e.g., river banks, wetlands, oases), high NDVI is not always a reliable predictor due to soil and water disturbance. We have addressed limitations of NDVI such as weather dependence, saturation in densely vegetated areas, or sensitivity to soil and water by integrating radar data (Joshi et al., 2016; Bergen et al., 2009). Our mapping approach successfully showcases how Sentinel-1 radar data can be integrated with optical remote sensing to increase the time-series density available for classification on large scales. This approach may hold potential for extending GDV mapping to other biomes, where, due to more regular cloud cover, optical remote sensing is substantially more limited (Mahmood et al., 2025).

Despite these methodological advances, certain limitations remain. First, our phreatophyte checklist is not complete and biased toward Europe due to reliance on EUNIS habitats, leaving groundwater-associated species in other Mediterranean regions underrepresented (Davies et al., 2004). Community-based GDV classification is further complicated by the variable groundwater dependence of facultative phreatophytes (Hultine et al., 2020), though using cover rather than abundance mitigates some uncertainty. Second, the regional bias in community data was partly addressed by combining sPlot and GBIF. However, GBIF introduces its own spatial clustering bias toward accessible areas (Thompson et al., 2023; Bird et al., 2013). Using sPlot and GBIF data from 1990 expands spatial coverage but risks overestimating GDV in degraded areas due to historical classification. Yet, restricting data to 2018–2023 would reduce the available ground-truth data by 90%. Third, low-resolution, modelled datasets, such as mTPI, precipitation, transpiration, and sand content, introduce gaps or inaccuracies, particularly in coastal and riverine landscapes. Also, uncertainties (e.g., sand content in Southern Spain, Italy or North Africa) are likely transferred to our model (Poggio et al., 2021; Zhang et al., 2019; Funk et al., 2015). Conceptually, pIDE reflects broader water balance deficits rather than direct groundwater use, potentially overestimating GDV where runoff accumulates or in areas with high soil moisture storage or lateral inflows (Castellazzi et al., 2024; Doody et al., 2017). Excluding these predictors reduced model training accuracy to 0.65, yet their global availability supports similar applications. Additionally, masking errors may arise from inaccuracies in land cover data, such as including irrigated crops or excluding GDV. We excluded groundwater depth, a key factor influencing GDV, due to the resolution limitations of available global datasets (Reinecke et al., 2024; Huang et al., 2019; Zhu et al., 2012). Adding lithology or aquifer data could improve ecological robustness, particularly for regional differences in groundwater

availability (WorldBank, 2025; Hartmann and Moosdorf, 2012). While we deemed the trade-offs between resolution and data completeness acceptable, especially given the extensive spatial coverage of our model, we note that small and isolated GDV (<30 m) such as spring vegetation or single trees may have gone undetected.

5. Conclusion

This study provides the first map of likely groundwater-dependent vegetation (GDV) in the Mediterranean biome at 30 m resolution. By integrating ground-truth species and community data, remote sensing, and machine learning, we addressed significant gaps in large-scale GDV mapping. The multi-level approach included compiling 1071 phreatophyte indicator species and six genera from literature and vegetation analysis into a novel species checklist. The checklist was used to create a species-community dataset from sPlot and GBIF databases for training and validating a Random Forest model predicting GDV distribution across the biome. We found that 482,000 km^2 of (semi-) natural vegetation in the Mediterranean biome is likely to be GDV, covering 28% of the study extent, with hotspots identified in Chile, California, the Iberian Peninsula, and Southwest Australia. Only about a quarter of this GDV currently lies within protected areas. Key predictors for GDV distribution include sand content, dry season vegetation vitality and elevation. Our results highlight the interconnected roles of groundwater availability, geodiversity, and vegetation characteristics in shaping GDV distribution, and emphasise their importance in conservation planning for biodiversity and geodiversity hotspots. Future research should integrate temporal trends in vegetation vitality, groundwater fluctuations, and land use change with the GDV distribution map to pinpoint high-risk areas and guide targeted conservation efforts throughout the Mediterranean biome.

CRediT authorship contribution statement

Léonard El-Hokayem: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Gabriella Damasceno:** Writing – review & editing, Formal analysis, Data curation, Conceptualization. **Francesco M. Sabatini:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Helge Bruelheide:** Writing – review & editing, Supervision, Resources, Methodology, Conceptualization. **Gianmaria Bonari:** Writing – review & editing, Validation, Data curation. **Ciara Dwyer:** Writing – review & editing, Data curation. **Fernando Gonçalves:** Writing – review & editing, Data curation. **Borja Jiménez-Alfaro:** Writing – review & editing, Data curation. **Jonathan Millett:** Writing – review & editing, Data curation. **Josep Peñuelas:** Writing – review & editing, Data curation. **Jens-Christian Svenning:** Writing – review & editing, Data curation. **Jan Altman:** Writing – review & editing, Data curation. **Han Y.H. Chen:** Writing – review & editing, Data curation. **Tetiana Dziuba:** Writing – review & editing, Data curation. **Mohamed A. El-Sheikh:** Writing – review & editing, Data curation. **Behlül Güler:** Writing – review & editing, Data curation. **Daniel Hending:** Writing – review & editing, Data curation. **Bruno Hérault:** Writing – review & editing, Data curation. **Mohamed Z. Hatim:** Writing – review & editing, Data curation. **Florian Jansen:** Writing – review & editing, Data curation. **Carsten Meyer:** Writing – review & editing, Supervision, Funding acquisition. **Sharif A. Mukul:** Writing – review & editing, Data curation. **Remigiusz Pielech:** Writing – review & editing, Data curation. **Flávio Rodrigues:** Writing – review & editing, Data curation. **Hua-Feng Wang:** Writing – review & editing, Data curation. **Christopher Conrad:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research was funded by the Federal State of Saxony-Anhalt via the MLU|BioDivFund. sPlot is a strategic project of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig funded by the German Research Foundation (FZT 118, 202548816). The authors thank the iDiv Data & Code Unit for assistance with curation and archiving of the dataset. We thank Melisa Giorgis for her helpful comments on the first draft, Michele Di Musciano for providing the code to calculate plot-wise co-occurrence from sPlot data, Gerd Schmidt for pitching the idea to use groundwater landscapes for cross-validation, and Johannes Löw for support with Sentinel-1 data. GB was funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4 - Call for tender No. 3138 of 16 December 2021, rectified by Decree n. 3175 of 18 December 2021 of Italian Ministry of University and Research funded by the European Union – NextGenerationEU; Award Number: Project code CN_00000033, Concession Decree No. 1034 of 17 June 2022 adopted by the Italian Ministry of University and Research, CUP B63C22000650007, Project title "National Biodiversity Future Center - NBFC". FG was supported by the Swiss National Science Foundation Postdoctoral Fellowships (TMPFP2_217531). BJ-A was funded by grant IDE/2024/000720 from Principality of Asturias-Sekuens-EU-FEDER. JP was supported by the Catalan Government grant AGAUR2023 CLIMA 00118. JCS was supported by the Center for Ecological Dynamics in a Novel Biosphere (ECONOVO), funded by the Danish National Research Foundation (grant DNRF173). JA was funded by research grant 23-05272S of the Czech Science Foundation and the long-term research development project No. RVO 67985939 of the Czech Academy of Sciences. CM acknowledges funding by the Volkswagen Foundation through a Freigeist Fellowship (A118199).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2026.114669.

Data availability

Data and codes are openly available in the iDiv Data Repository at: https://doi.org/10.25829/idiv.3602-q75799. The vegetation plot raw data contained in the sPlot database are available upon request.

References

Abdullah, M.M., Al-Ali, Z.M., Blanton, A., Charabi, Y., Abulideh, A., Al-Awadhi, T., Srinivasan, S., Fadda, E., Mohan, M., 2024. UAVs for improving seasonal vegetation assessment in arid environments. Front. Environ. Sci. 12. https://doi.org/10.3389/fenvs.2024.136671.

Ahring, T.S., Steward, D.R., 2012. Groundwater surface water interactions and the role of phreatophytes in identifying recharge zones. Hydrol. Earth Syst. Sci. 16 (11), 4133–4142. https://doi.org/10.5194/hess-16-4133-2012.

Akasheh, O.Z., Neale, C.M.U., Jayanthi, H., 2008. Detailed mapping of riparian vegetation in the middle Rio Grande River using high-resolution multi-spectral airborne remote sensing. J. Arid Environ. 72, 1734–1744. https://doi.org/10.1016/j.jaridenv.2008.03.014.

Ali, I., Cao, S., Naeimi, V., Paulik, C., Wagner, W., 2018. Methods to remove the border noise from Sentinel-1 synthetic aperture radar data: implications and importance for time-series analysis. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11 (3), 777–786. https://doi.org/10.1109/JSTARS.2017.2787650.

Bai, T., Wang, X.-S., Han, P., 2024. Controls of groundwater-dependent vegetation coverage in the yellow river basin, China: insights from interpretable machine learning. J. Hydrol. 631, 130747. https://doi.org/10.1016/j.jhydrol.2024.130747.

Barreñada, L., Dhiman, P., Timmerman, D., Boulesteix, A.-L., van Calster, B., 2024. Understanding overfitting in random forest for probability estimation: a visualization and simulation study. Diagnostic and Prognostic Research 8 (1), 14. https://doi.org/10.1186/s41512-024-00177-1.

Barron, O., Van Niel, T.G., 2009. Application of thermal remote sensing to delineate groundwater discharge zones. Water 5 (2), 109–124. https://doi.org/10.1504/IJW.2009.028721.

Barron, O.V., Emelyanova, I., Van Niel, T.G., Pollock, D., Hodgson, G., 2014. Mapping groundwater-dependent ecosystems using remote sensing measures of vegetation and moisture dynamics. Hydrol. Process. 28, 372–385. https://doi.org/10.1002/hyp.9609.

Beck, H., Zimmermann, N., McVicar, T., et al., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci. Data 5, 180214. https://doi.org/10.1038/sdata.2018.214.

Bergen, K.M., Goetz, S.J., Dubayah, R.O., Henebry, G.M., Hunsaker, C.T., Imhoff, M.L., Nelson, R.F., Parker, G.G., Radeloff, V.C., 2009. Remote sensing of vegetation 3-D structure for biodiversity and habitat: review and implications for lidar and radar spaceborne missions. J. Geophys. Res. 114, G00E06. https://doi.org/10.1029/2008JG000883.

Bird, T.J., Bates, A.E., Lefcheck, J.S., Hill, N.A., Thomson, R.J., Edgar, G.J., Stuart-Smith, R.D., Wotherspoon, S., Krkosek, M., Stuart-Smith, J.F., Pecl, G.T., Barrett, N., Frusher, S., 2013. Statistical solutions for error and bias in global citizen science datasets. https://doi.org/10.1016/j.biocon.2013.07.037.

Box, J.B., Leiper, I., Nano, C., Stokeld, D., Jobson, P., Tomlinson, A., Cobban, D., Bond, T., Randall, D., Box, P., 2022. Mapping terrestrial groundwater-dependent ecosystems in arid Australia using Landsat-8 time-series data and singular value decomposition. Remote Sensing in Ecology and Conservation 8 (4), 464–476. https://doi.org/10.1002/rse2.254.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Brilha, J., 2014. Inventory and quantitative assessment of Geosites and geodiversity sites: a review. Geoheritage 8, 119–134. https://doi.org/10.1007/s12371-014-0139-3.

Bruelheide, H., 2000. A new measure of fidelity and its application to defining species groups. J. Veg. Sci. 11 (2), 167–178. https://doi.org/10.2307/3236796.

Bruelheide, H., Dengler, J., Jiménez-Alfaro, B., 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.

Cabello, J., Escudero-Clares, M., Martos-Rosillo, S., Casas, J.J., Cintas, J., Zakaluk, T., Salinas-Bonillo, M.J., 2026. Groundwater-dependent vegetation in semi-arid Mediterranean mountains: the hidden role of weathered hard-rock aquifers. J. Hydrol. 664, 134407. https://doi.org/10.1016/j.jhydrol.2025.134407.

California Department of Water Resources, 2022. i02 NaturalCommunitiesCommonlyAssociatedWithGroundwater, California Department of Water Resources, Sacramento, USA. https://gis.data.cnra.ca.gov/maps/e0cd9c77db7648b49bf8dafa8d9f104c (Dataset).

Canadell, J., Jackson, R.B., Ehleringer, J.R., Mooney, H.A., Sala, O.E., Schulze, E.D., 1996. Maximum rooting depth of vegetation types at the global scale. Oecologia 108 (4), 583–595. https://doi.org/10.1007/BF00329030.

Castellazzi, P., Gao, S., Pritchard, J., Ponce-Reyes, R., Stratford, D., Crosbie, R., 2024. Detecting springs and groundwater-dependent vegetation in data-scarce regions of Australia combining citizen science, GRACE, and optical/radar imagery. Remote Sens. Environ. 313, 114345. https://doi.org/10.1016/j.rse.2024.114345.

Castellazzi, P., Prichard, J., Stratford, D., 2023. Mapping drought-resilient and groundwater-Dependant vegetation with Sentinel-1 Sar data. In: IGARSS 2023–2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, USA, pp. 2886–2889. https://doi.org/10.1109/IGARSS52108.2023.10282802.

Chamberlain, S., Barve, V., Mcglinn, D., Oldoni, D., Desmet, P., Geffert, L., Ram, K., 2024. Rgbif: Interface to the Global Biodiversity Information Facility API. https://doi.org/10.32614/CRAN.package.rgbif (R package version 3.8.1).

Chamberlain, S., Boettiger, C., 2017. R Python, and ruby clients for GBIF species occurrence data. PeerJ PrePrints. https://doi.org/10.7287/peerj.preprints.3304v1.

Charbonneau, F., Trudel, M., Fernandes, R., 2005. Use of dual polarization and multi-incidence SAR for soil permeability mapping. In: proceedings of the 2005 advanced synthetic aperture radar (ASAR) workshop, St-Hubert, QC, Canada, 15-17.

Chikishev, A.G. (Ed.), 1965. Plant Indicators of Soils, Rocks, and Subsurface Waters. Consultants Bureau, New York.

Chytrý, M., Tichý, L., Hennekens, S., et al., 2020. EUNIS habitat classification: expert system, characteristic species combinations and distribution maps of European habitats. Appl. Veg. Sci. 23, 648–675. https://doi.org/10.1111/avsc.12519.

Chytrý, M., Tichý, L., Holt, J., Botta-Dukát, Z., 2002. Determination of diagnostic species with statistical fidelity measures. J. Veg. Sci. 13, 79–90. https://doi.org/10.1111/j.1654-1103.2002.tb02025.x.

Davies, C.E., Moss, D., Hill, M.O., 2004. EUNIS Habitat Classification. European Environment Agency, Copenhagen.

De Cáceres, M., Font, X., Oliva, F., 2008. Assessing species diagnostic value in large data sets: A comparison between phi-coefficient and Ochiai index. J. Veg. Sci. 19, 779–788. https://doi.org/10.3170/2008-8-18446.

De Cáceres, M., Legendre, P., 2009. Associations between species and groups of sites: indices and statistical inference. Ecology 90, 3566–3574. https://doi.org/10.1890/08-1823.1.

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

Doody, T.M., Barron, O.V., Dowsley, K., Emelyanova, I., Fawcett, J., Overton, I.C., Pritchard, J.L., Van Dijk, A., Warren, G., 2017. Continental mapping of groundwater dependent ecosystems: A methodological framework to integrate diverse data and expert opinion. Journal of Hydrology: Regional Studies 10, 61–81. https://doi.org/10.1016/j.ejrh.2017.01.003.

Duran-Llacer, I., Arumí, J.L., Arriagada, L., Aguayo, M., Rojas, O., González-Rodríguez, L., Rodríguez-López, L., Martínez-Retureta, R., Oyarzún, R., Singh, S.K., 2022. A new method to map groundwater-dependent ecosystem zones in semi-arid environments: A case study in Chile. Sci. Total Environ. 816, 151528. https://doi.org/10.1016/j.scitotenv.2021.151528.

Eamus, D., Fu, B., Springer, A.E., Stevens, L.E., 2016. Groundwater dependent ecosystems: Classification, identification techniques and threats. In: Jakeman, A.J., Barreteau, O., Hunt, R.J., Rinaudo, J.D., Ross, A. (Eds.), Integrated Groundwater Management. Springer, Cham. https://doi.org/10.1007/978-3-319-23576-9_13.

El-Hokayem, L., De Vita, P., Conrad, C., 2023a. Local identification of groundwater dependent vegetation using high-resolution Sentinel-2 data – a Mediterranean case study. Ecol. Indic. 146, 109784. https://doi.org/10.1016/j.ecolind.2022.109784.

El-Hokayem, L., De Vita, P., Usman, M., Link, A., Conrad, C., 2023b. Mapping potentially groundwater-dependent vegetation in the Mediterranean biome using global Geodata targeting site conditions and vegetation characteristics. Sci. Total Environ. 898, 166397. https://doi.org/10.1016/j.scitotenv.2023.166397.

El-Hokayem, L., De Vita, P., Usman, M., Link, A., Conrad, C., 2023c. Potential groundwater dependent vegetation in the Mediterranean. PANGAEA. https://doi.org/10.1594/PANGAEA.961765 [Dataset].

Evaristo, J., McDonnell, J., 2017. Prevalence and magnitude of groundwater use by vegetation: a global stable isotope meta-analysis. Sci. Rep. 7, 44110. https://doi.org/10.1038/srep44110.

Fan, Y., Miguez-Macho, G., Jobbágy, E.G., Jackson, R.B., Otero-Casal, C., 2017. Hydraulic regulation of plant rooting depth. PNAS 114, 40. https://doi.org/10.1073/pnas.1712381114.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D.E., 2007. The shuttle radar topography mission. Rev. Geophys. 45 (2), RG2004. https://doi.org/10.1029/2005RG000183.

Funk, C., Pete, P., Martin, L., Pedreros, D., Verdin, J., Shraddhanand, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes. Sci. Data 2, 150066. https://doi.org/10.1038/sdata.2015.66 [Dataset].

GBIF.org, 25 September 2024. GBIF Occurrence Download. Dataset. https://doi.org/10.15468/dl.x3nmbp.

Gearty, W., Chamberlain, S., 2024. rredlist: 'IUCN' Red List Client. doi:10.32614/CRAN.package.rredlist, R package version 0.7.1.9000, https://docs.ropensci.org/rredlist/.

Geographics, Earthstar, 2024. Imagery Basemap ArcGis pro. Dataset.

Glanville, K., Sheldon, F., Butler, D., Capon, S., 2023. Effects and significance of groundwater for vegetation: A systematic review. Sci. Total Environ. 875, 162577. https://doi.org/10.1016/j.scitotenv.2023.162577.

Gomes Marques, I., Nascimento, J., Cardoso, R.M., Miguéns, F., Condesso De Melo, M.T., Soares, P.M.M., Gouveia, C.M., Kurz Besson, C., 2019. Mapping the suitability of groundwater-dependent vegetation in a semi-arid Mediterranean area. Hydrol. Earth Syst. Sci. 23, 3525–3552. https://doi.org/10.5194/hess-23-3525-2019.

Google, 2024. Google Satellite Imagery. [Dataset].

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Gou, S., Gonzales, S., Miller, G.R., 2015. Mapping potential groundwater-dependent ecosystems for sustainable management. Groundwater 53 (1), 99–110. https://doi.org/10.1111/gwat.12169.

Hartmann, J., Moosdorf, N., 2012. Global Lithological Map. Database v1.0. https://doi.org/10.1594/PANGAEA.788537 (gridded to 0.5° spatial resolution) [dataset]. PANGAEA.

Howard, J.K., Dooley, K., Brauman, K.A., Klausmeyer, K.R., Rohde, M.M., 2023. Ecosystem services produced by groundwater dependent ecosystems: a framework and case study in California. Frontiers in Water 5, 1115416. https://doi.org/10.3389/frwa.2023.1115416.

Ramsar, 2025. Ramsar Sites and the List of Wetlands of International Importance. Gland, Switzerland, The Ramsar Convention Secretariat. https://rsis.ramsar.org/.

Huang, F., Zhang, Y., Zhang, D., Chen, X., 2019. Environmental groundwater depth for groundwater-dependent terrestrial ecosystems in arid/semiarid regions: A review. Int. J. Environ. Res. Public Health 16, 763. https://doi.org/10.3390/ijerph16050763.

Hultine, K.R., Friend, R., Blasini, D., Bush, S.E., Karlinski, M., Koepke, D.F., 2020. Hydraulic traits that buffer deep-rooted plants from changes in hydrology and climate. Hydrol. Process. 34 (2), 209–222. https://doi.org/10.1002/hyp.13587.

Jasechko, S., Seybold, H., Perrone, D., et al., 2024. Rapid groundwater decline and some cases of recovery in aquifers globally. Nature 625, 715–721. https://doi.org/10.1038/s41586-023-06879-8.

Jones, C., Stanton, D., Hamer, N., Denner, S., Singh, K., Flook, S., Dyring, M., 2019. Field investigation of potential terrestrial groundwater-dependent ecosystems within Australia's great Artesian Basin. Hydrgeol. J. 28, 1–25. https://doi.org/10.1007/s10040-019-02081-1.

Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M.R., Kuemmerle, T., Meyfroidt, P., Mitchard, E.T.A., Reiche, J., Ryan, C.M., Waske, B., 2016. A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. Remote Sens. 8 (1), 70. https://doi.org/10.3390/rs8010070.

Killroy, G., Dunne, F., Ryan, J., O'Connor, Á., Daly, D., Craig, M., Coxon, C., Johnston, P., Moe, H., 2008. A Framework for the Assessment of Groundwater-Dependent Terrestrial Ecosystems under the Water Framework Directive. Environmental

Protection Agency, Wexford. https://www.epa.ie/publications/research/water/erc-report-12.php.
Kim, Y., Jackson, T., Bindlish, R., Lee, H., Hong, S., 2012. Radar vegetation index for estimating the vegetation water content of Rice and soybean. IEEE Geosci. Remote Sens. Lett. 9 (4), 564–568. https://doi.org/10.1109/LGRS.2011.2174772.
Kløve, B., Ala-Aho, P., Bertrand, G., Gurdak, J.J., Kupfersberger, H., Kvaerner, J., Muotka, T., Mykrä, H., Preda, E., Rossi, P., Bertrachi, C., Velasco, E., Pulido-Velazquez, M., 2014. Climate change impacts on groundwater and dependent ecosystems. J. Hydrol. 518, 250–266. https://doi.org/10.1016/j.jhydrol.2013.06.037.
Kløve, B., Ala-aho, P., Bertrand, G., et al., 2011. Groundwater dependent ecosystems. Part I: Hydroecological status and trends. Environmental Science & Policy 14 (7), 770–781. https://doi.org/10.1016/j.envsci.2011.04.002.
Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28 (5), 1–26. https://doi.org/10.18637/jss.v028.i05.
Lewis, J., 2011. The application of Ecohydrological groundwater indicators to hydrogeological conceptual models. Groundwater 50 (5), 679–689. https://doi.org/10.1111/j.1745-6584.2011.00899.x.
Lin, P., Pan, M., Wood, E.F., Yamazaki, D., Allen, G.H., 2021. A new vector-based global river network dataset accounting for variable drainage density. https://doi.org/10.6084/m9.figshare.c.5052635.v1 [Dataset].
Link, A., El-Hokayem, L., Usman, M., Conrad, C., Reinecke, R., Berger, M., Wada, Y., Coroama, V., Finkbeiner, M., 2023. Groundwater-dependent ecosystems at risk –global hotspot analysis and implications. Environ. Res. Lett. 18, 094026. https://doi.org/10.1088/1748-9326/acea97.
Liu, C., Liu, H., Yu, Y., Zhao, W., Zhang, Z., Guo, L., Yetemen, O., 2021. Mapping groundwater-dependent ecosystems in arid Central Asia: implications for controlling regional land degradation. Sci. Total Environ. 797. https://doi.org/10.1016/j.scitotenv.2021.149027.
Liu, Y., Xu, X., Dimitrov, D., Pellissier, L., Borregaard, M.K., Shrestha, N., Su, X., Luo, A., Zimmermann, N.E., Rahbek, C., Wang, Z., 2023. An updated floristic map of the world. Nat. Commun. 14, 2990. https://doi.org/10.1038/s41467-023-38375-y.
Londo, G., 1975. Nederlandse lijst van hydro-, freato- en afreatofyten. Rijksinstituut voor Natuurbeheer, Leersum. https://edepot.wur.nl/367558.
Lowry, C.S., Loheide II, S.P., Moore, C.E., Lundquist, J.D., 2011. Groundwater controls on vegetation composition and patterning in mountain meadows. Water Resour. Res. 47, W00J11. https://doi.org/10.1029/2010WR010086.
Mahmood, T., Usman, M., Conrad, C., 2025. Selecting relevant features for random Forest-based crop type classifications by spatial assessments of backward feature reduction. PFG – journal of photogrammetry, remote sensing and Geoinformation. Science. https://doi.org/10.1007/s41064-024-00329-4.
Maitner, B., Boyle, B., Efren, P., 2024. Taxonomic Name Resolution Service. https://doi.org/10.32614/CRAN.package.TNRS, R-Package Version 0.3.6.
Martínez-Santos, P., Díaz-Alcaide, S., De la Hera-Portillo, A., Gómez-Escalonilla, V., 2021. Mapping groundwater-dependent ecosystems by means of multi-layer supervised classification. J. Hydrol. 603, 126873. https://doi.org/10.1016/j.jhydrol.2021.126873.
Martínez-Santos, P., Renard, P., 2020. Mapping groundwater potential through an ensemble of big data methods. Groundwater 58, 4. https://doi.org/10.1111/gwat.12939.
Médail, F., Quézel, P., 1999. Biodiversity hotspots in the Mediterranean Basin: setting global conservation priorities. Conserv. Biol. 13, 1510–1513. https://doi.org/10.1046/j.1523-1739.1999.98467.x.
Meinzer, O.E., 1927. Plants as indicators of groundwater. USGS Water-Suppl. Pap. 577. https://doi.org/10.3133/wsp577.
Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., Nauss, T., 2018. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. Environ. Model. Software 101, 1–9. https://doi.org/10.1016/j.envsoft.2017.12.001.
Milano, M., Ruelland, D., Fernandez, S., et al., 2013. Current state of Mediterranean water resources and future trends under climatic and anthropogenic changes. Hydrol. Sci. J. 58 (3), 498–518. https://doi.org/10.1080/02626667.2013.774458.
Mullissa, A., Vollrath, A., Odongo-Braun, C., Slagter, B., Balling, J., You, Y., Gorelick, N., Reiche, J., 2021. Sentinel-1 SAR backscatter analysis ready data preparation in Google earth engine. Remote Sens. 13 (10), 1954. https://doi.org/10.3390/rs13101954.
Münch, Z., Conrad, J., 2007. Remote sensing and GIS based determination of groundwater dependent ecosystems in the Western cape, South Africa. Hydrgeol. J. 15, 19–28. https://doi.org/10.1007/s10040-006-0125-1.
Mushagalusa, C.A., Fandohan, A.B., Glèlè Kakaï, R., 2024. Random forest and spatial cross-validation performance in predicting species abundance distributions. Environ. Syst. Res. 13, 23. https://doi.org/10.1186/s40068-024-00352-9.
Myers, N., Mittermeier, R., Mittermeier, C., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858. https://doi.org/10.1038/35002501.
Orellana, F., Verma, P., Loheide, S.P., Daly, E., 2012. Monitoring and modeling water-vegetation interactions in groundwater-dependent ecosystems. Rev. Geophys. 50, 3003. https://doi.org/10.1029/2011RG000383.
Pereira, J.S., David, J.S., David, T.S., Caldeira, M.C., Chaves, M.M., 2004. Carbon and water fluxes in Mediterranean-type ecosystems — constraints and adaptations. In: Esser, K., Lüttge, U., Beyschlag, W., Murata, J. (eds) Progress in botany. Progress in botany, 65, Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-18819-0_19.
Pérez Hoyos, I.C., Krakauer, N.Y., Khanbilvardi, R., Armstrong, R.A., 2016. A review of advances in the identification and characterization of groundwater dependent ecosystems using geospatial technologies. Geosciences 6, 17. https://doi.org/10.3390/geosciences6020017.
Pesaresi, M., Politis, P., 2023. GHS-BUILT-C R2023A - GHS settlement characteristics, derived from Sentinel2 composite (2018) and other GHS R2023A data. European Commission. https://doi.org/10.2905/3c60ddf6-0586-4190-854b-f6aa0edc2a30. Joint Research Centre (JRC).
Pignatti, S., 2003. The Mediterranean ecosystem. Bocconea 16 (1), 29–40.
Poggio, L., de Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., Rossiter, D., 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. SOIL 7, 217–240. https://doi.org/10.5194/soil-7-217-2021.
Polman, E.M.N., Seijmonsbergen, A.C., Versteegh, H., Kissling, W.D., 2024. Global geodiversity components are not equally represented in UNESCO global Geoparks. Phil. Trans. R. Soc. A 382, 20230054. https://doi.org/10.1098/rsta.2023.0054.
Rampheri, M.B., Dube, T., Dondofema, F., Dalu, T., 2023. Identification and delineation of groundwater dependent ecosystems (GDEs) in the Khakea–Bray transboundary aquifer region using geospatial techniques. Geocarto Int. 38 (1), 2172217. https://doi.org/10.1080/10106049.2023.2172217.
Reinecke, R., Gnann, S., Stein, L., Bierkens, M., de Graaf, I., Gleeson, T., Oude Essink, G., Sutanudjaja, E.H., Ruz Varags, C., Verkaik, J., Wagener, T., 2024. Uncertainty in model estimates of global groundwater depth. ERL 19, 11. https://doi.org/10.1088/1748-9326/ad8587.
Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B.A., Hartig, F., Dormann, C.F., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40, 913–929. https://doi.org/10.1111/ecog.02881.
Rohde, M.M., Albano, C.M., Huggins, X., et al., 2024b. Groundwater-dependent ecosystem map exposes global dryland protection needs. Nature 632, 101–107. https://doi.org/10.1038/s41586-024-07702-8.
Rohde, M.M., Stella, J.C., Singer, M.B., Roberts, D.A., Caylor, K.K., Albano, C.M., 2024a. Establishing ecological thresholds and targets for groundwater management. Nat Water 2, 312–323. https://doi.org/10.1038/s44221-024-00221-w.
Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., Harlan, J.C., 1974. Monitoring the Vernal Advancement of Retrogradation of Natural Vegetation. NASA/GSFC, Type III, Final Report, p. 371.
Saatkamp, A., Argagnon, O., Noble, V., Finocchiaro, M., Meineri, E., 2023. Climate change impacts on Mediterranean vegetation are amplified at low altitudes. Glob. Ecol. Biogeogr. 32, 1113–1126. https://doi.org/10.1111/geb.13682.
Sabatini, F.M., Lenoir, J., Hattab, T., et al., 2021. sPlotOpen – an environmentally balanced, open-access, global dataset of vegetation plots. Glob. Ecol. Biogeogr. 30 (9), 1740–1764. https://doi.org/10.1111/geb.13346.
Saccò, M., Mammola, S., Altermatt, F., et al., 2023. Groundwater is a hidden global keystone ecosystem. Glob. Chang. Biol. 30 (1), e17066. https://doi.org/10.1111/gcb.17066.
Schulte Bühne, H., Pettorelli, N., 2018. Better together: integrating and fusing multispectral and radar satellite imagery to inform biodiversity monitoring, ecological research and conservation science. Methods Ecol. Evol. 9, 849–865. https://doi.org/10.1111/2041-210X.12942.
Soares, P.M.M., Lima, D.C.A., 2022. Water scarcity down to earth surface in a Mediterranean climate: the extreme future of soil moisture in Portugal. J. Hydrol. 615, 128731. https://doi.org/10.1016/J.JHYDROL.2022.128731.
Szigarski, C., Jagdhuber, T., Baur, M., Thiel, C., Parrens, M., Wigneron, J.-P., Piles, M., Entekhabi, D., 2018. Analysis of the radar vegetation index and potential improvements. Remote Sens. 10, 1776. https://doi.org/10.3390/rs10111776.
Theobald, D.M., Harrison-Atlas, D., Monahan, W.B., Albano, C.M., 2015. Ecologically-relevant maps of landforms and physiographic diversity for climate adaptation planning. PloS One 10 (12), e0143619.
Thomas, F.M., 2014. Ecology of Phreatophytes. In: Lüttge, U., et al. (Eds.), Progress in Botany 75. Springer, Berlin/Heidelberg, pp. 335–375.
Thompson, M.M., Moon, K., Woods, A., Rowley, J.J.L., Poore, G.B., Kingsford, R.T., Callaghan, C.T., 2023. Citizen science participant motivations and behaviour: Implications for biodiversity data coverage. https://doi.org/10.1016/j.biocon.2023.110079.
Tootchi, A., Jost, A., Ducharne, A., 2019. Multi-source global wetland maps combining surface water imagery and groundwater constraints. Earth System Science Data 11 (1), 189–220. https://doi.org/10.5194/essd-11-189-2019.
UNEP-WCMC & IUCN, 2025. Protected Planet: The World Database on Protected Areas (WDPA) and World Database on Other Effective Area-Based Conservation Measures (WD-OECM) [Dataset]. UNEP-WCMC & IUCN. www.protectedplanet.net, Cambridge, UK.
Valavi, R., Elith, J., Lahoz-Monfort, J.J., Guillera-Arroita, G., 2019. blockCV: an r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. Methods Ecol. Evol. 10 (2), 225–232. https://doi.org/10.1111/2041-210X.13107.
Wang, P., Zhang, Y., Yu, J., Fu, G., Ao, F., 2011. Vegetation dynamics induced by groundwater fluctuations in the lower Heihe River basin, northwestern China. J. Plant Ecol. 4 (1–2), 77–90. https://doi.org/10.1093/jpe/rtr002.
Web of Science, 2024. 839 results from Web of Science Core Collection for developed search query. https://www.webofscience.com/wos/woscc/summary/f4785da8-1b68-4d0f-bb6a-73b39097f793-0131badf4f/relevance/1. [Dataset].
Wessel, P., Smith, W.H.F., 1996. A global self-consistent, hierarchical, high-resolution shoreline database. J. Geophys. Res. 101, 8741–8743. https://doi.org/10.1029/96JB00104.
World Bank, 2025. A Global Dataset of Aquifer Typologies and Groundwater Resources : The Hidden Wealth of Nations - the Economics of Groundwater in Times of Climate

Change : Background Paper. World Bank Group, Washington, D.C.. https://doi.org/10.1596/43749

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsendbazar, N.E., Xu, P., Ramoino, F., Arino, O., 2022. ESA WorldCover 10 m 2021 v200. [Dataset]. https://doi.org/10.5281/zenodo.7254221.

Zhang, Y., Kong, D., Gan, R., Chiew, F.H.S., McVicar, T.R., Zhang, Q., Yang, Y., 2019. Coupled estimation of 500m and 8-day resolution global evapotranspiration and gross primary production in 2002-2017. Remote Sens. Environ. 222, 165–182. https://doi.org/10.1016/j.rse.2018.12.031.

Zhu, J., Yu, J., Wang, P., Zhang, Y., Yu, Q., 2012. Interpreting the groundwater attributes influencing the distribution patterns of groundwater-dependent vegetation in northwestern China. Ecohydrology 5 (5), 628–636. https://doi.org/10.1002/eco.249.

Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D., Farooq, H., Herdean, A., Ariza, M., Scharn, R., Svanteson, S., Wengstrom, N., Zizka, V., Antonelli, A., 2019. CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. Methods in Ecology and Evolution 7. https://doi.org/10.1111/2041-210X.13152, R package version 3.0.1.