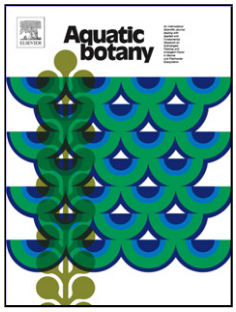
World distribution, diversity and endemism of aquatic macrophytes

Murphy, K., Efremov, A., Davidson, T. A., Molina-Navarro, E., Fidanza, K., Crivelari Betiol, T. C., Chambers, P., Tapia Grimaldo, J., Varandas Martins, S., Springuel, I., Kennedy, M., Mormul, R. P., Dibble, E., Hofstra, D., Lukács, B. A., Gebler, D., Baastrup-Spohr, L. & Urrutia-Estrada, J.

Original citation & hyperlink:
Murphy, K, Efremov, A, Davidson, TA, Molina-Navarro, E, Fidanza, K, Crivelari Betiol, TC, Chambers, P, Tapia Grimaldo, J, Varandas Martins, S, Springuel, I, Kennedy, M, Mormul, RP, Dibble, E, Hofstra, D, Lukács, BA, Gebler, D, Baastrup-Spohr, L & Urrutia-Estrada, J 2019, 'World distribution, diversity and endemism of aquatic macrophytes' Aquatic Botany, vol. 158, 103127.
https://dx.doi.org/10.1016/j.aquabot.2019.06.006

DOI 10.1016/j.aquabot.2019.06.006
ISSN 0304-3770

Publisher: Elsevier



Title: World distribution, diversity and endemism of aquatic macrophytes

Authors: Kevin Murphy, Andrey Efremov, Thomas A. Davidson, Eugenio Molina-Navarro, Karina Fidanza, Tânia Camila Crivelari Betiol, Patricia Chambers, Julissa Tapia Grimaldo, Sara Varandas Martins, Irina Springuel, Michael Kennedy, Roger Paolo Mormul, Eric Dibble, Deborah Hofstra, Balázs András Lukács, Daniel Gebler, Lars Baastrup-Spohr, Jonathan Urrutia-Estrada

PII: S0304-3770(19)30030-0
DOI: https://doi.org/10.1016/j.aquabot.2019.06.006
Reference: AQBOT 3127

To appear in: *Aquatic Botany*

Received date: 26 January 2019
Revised date: 20 May 2019
Accepted date: 15 June 2019

Please cite this article as: Murphy K, Efremov A, Davidson TA, Molina-Navarro E, Fidanza K, Crivelari Betiol TC, Chambers P, Tapia Grimaldo J, Varandas Martins S, Springuel I, Kennedy M, Mormul RP, Dibble E, Hofstra D, Lukács BA, Gebler D, Baastrup-Spohr L, Urrutia-Estrada J, World distribution, diversity and endemism of aquatic macrophytes, *Aquatic Botany* (2019), https://doi.org/10.1016/j.aquabot.2019.06.006

World distribution, diversity and endemism of aquatic macrophytes

Kevin Murphy[a,*], Andrey Efremov[b], Thomas A. Davidson[c], Eugenio Molina-Navarro[c,1], Karina Fidanza[d], Tânia Camila Crivelari Betiol[d], Patricia Chambers[e], Julissa Tapia Grimaldo[a], Sara Varandas Martins[a], Irina Springuel[f], Michael Kennedy[g], Roger Paolo Mormul[d], Eric Dibble[h], Deborah Hofstra[i], Balázs András Lukács[j], Daniel Gebler[k], Lars Baastrup-Spohr[l], Jonathan Urrutia-Estrada[m,n,o]

[a] *University of Glasgow, Glasgow G12 8QQ, Scotland*
[b] *Omsk State Pedagogical University, 14, Tukhachevskogo nab., 644009 Omsk, Russia*
[c] *Lake Group, Dept of Bioscience, Silkeborg, Aarhus University, Denmark*
[d] *NUPELIA, Universidade Estadual de Maringá, Maringá, PR, Brazil*
[e] *Environment and Climate Change Canada, Burlington, Ontario, Canada*
[f] *Department of Botany & Environmental Science, Aswan University, 81528 Sahari, Egypt*
[g] *School of Energy, Construction and Environment, University of Coventry, Priory Street, Coventry, CV1 5FB, UK*
[h] *Department of Wildlife, Fisheries and Aquaculture, Mississippi State University, Starkville, MS 39762, USA*
[i] *National Institute of Water and Atmospheric Research (NIWA), Hamilton, New Zealand*
[j] *Department of Tisza River Research, MTA Centre for Ecological Research, DRI, 4026 Debrecen Bem tér 18/C, Hungary*
[k] *Poznan University of Life Sciences, Wojska Polskiego 28, 60637 Poznan, Poland*
[l] *Institute of Biology, Freshwater Biology Section, University of Copenhagen, Denmark*
[m] *Departamento de Ciencias Biológicas y Químicas, Facultad de Recursos Naturales, Universidad Católica de Temuco, Casilla 15-D, Temuco, Chile*
[n] *Núcleo de Estudios Ambientales, Facultad de Recursos Naturales, Universidad Católica de Temuco, Casilla 15-D, Temuco, Chile*
[o] *Laboratorio de Invasiones Biológicas, Facultad de Ciencias Forestales, Universidad de Concepción, Casilla 160-C, Concepción, Chile.*

[1] Current address: *Hydrological Resources and Aquatic Ecosystems, Department of Geology, Geography and Environment, University of Alcalá, 28802 Alcalá de Henares (Madrid), Spain*

E-mail addresses: mearnskevin@googlemail.com (K. Murphy), stratiotes@yandex.ru (A. Efremov), ab9280@coventry.ac.uk (M. Kennedy), eugenio.molinanavarro@gmail.com (E. Molina-Navarro), thd@bios.au.dk (T. Davidson), julstg@gmail.com (J. Tapia Grimaldo), martinssaraster@gmail.com (S. Varandas Martins), irina44@yahoo.com (I. Springuel), taniacrivelari@hotmail.com (T.C. Crivelari Betiol), roger.mormul@gmail.com (R.P. Mormul), karina.fidanza@gmail.com (K. Fidanza), patricia.chambers@canada.ca (P. Chambers), deborah.hofstra@niwa.co.nz (D. Hofstra), lukacs.balazs@okologia.mta.hu (B.A. Lukács), jurrutiaestrada@gmail.com (J. Urrutia Estrada), e.dibble@msstate.edu (E. Dibble), daniel.gebler@up.poznan.pl (D. Gebler), lbaastrupspohr@bio.ku.dk (L. Baastrup-Spohr).

* Corresponding author. Tel: +44 1479 873541

Highlights:

- Aquatic macrophytes generally show narrow distributional ranges at world scale
- Only 42 species (1.21%) of the total 3457 macrophyte species have broad world range
- A non-linear latitude-diversity gradient for macrophytes peaks in the sub-tropics
- Macrophytes exhibit pronounced endemism at both world ecozone and smaller scales
- The richest world hotspots for macrophyte diversity all lie in the Neotropics

Abstract

To test the hitherto generally-accepted hypothesis that most aquatic macrophytes have broad world distributions, we investigated the global distribution, diversity and endemism patterns of 3457 macrophyte species that occur in permanent, temporary or ephemeral inland freshwater and brackish waterbodies worldwide. At a resolution of 10x10° latitude x longitude, most macrophyte species were found to have narrow global distributions: 78% have ranges (measured using an approach broadly following the IUCN-defined concept "extent of occurrence") that individually occupy <10% of the world area present within the six global ecozones which primarily provide habitat for macrophytes. We found evidence of non-linear relationships between latitude and macrophyte α- and γ-diversity, with diversity highest in sub-tropical to low tropical latitudes, declining slightly towards the Equator, and also declining strongly towards higher latitudes. Landscape aridity and, to a lesser extent, altitude and land area present per gridcell also influence macrophyte diversity and species assemblage worldwide. The Neotropics and Orient have the richest ecozone species-pools for macrophytes, depending on γ-diversity metric used. The region around Brasilia/Goiás (Brazil: gridcell 10-20°S; 40-50°W) is the richest global hotspot for macrophyte α-diversity (total species α-diversity, S_T: 625 species/gridcell, 350 of them Neotropical endemics). In contrast, the Sahara/Arabian Deserts, and some Arctic areas, have the lowest macrophyte α-diversity (S_T <20 species/gridcell). At ecozone scale, macrophyte species endemism is pronounced, though with a >5-fold difference between the most species-rich (Neotropics) and species-poor (Palaearctic) ecozones. Our findings strongly support the assertion that small-ranged species constitute most of Earth's species diversity.

Key words: aquatic plants, biodiversity hotspots, latitudinal diversity gradient, world ecozones, macroecology

1. Introduction

Aquatic macrophytes are "aquatic photosynthetic organisms, large enough to see with the naked eye, that actively grow permanently or periodically submerged below, floating on, or up through the water surface" of inland freshwater or brackish waterbodies (Chambers et al., 2008). This definition includes aquatic plants that live in permanent, temporary, and ephemeral inland waterbodies and watercourses. Permanent inland waterbodies (including lakes, rivers, canals, reservoirs and other inland waters that rarely, if ever, dry up) self-evidently offer potential macrophyte habitat, but temporary and ephemeral waterbodies are more open to discussion regarding their status as macrophyte habitats (e.g., Lukács et al., 2013; see also Appendix A1). Temporary waterbodies are those that dry up and re-fill following a reasonably predictable (often annual) temporal precipitation cycle, e.g., many English chalk streams, supporting species such as *Ranunculus penicillatus* (Dumort.) Bab. (Ranunculaceae): Haslam (1978). Ephemeral waterbodies are less predictable in their occurrence and may go several years between filling, often being dependent upon quantity of rainfall in a given year. The macrophytes of temporary and ephemeral systems tolerate periodic drought, but usually have a physiological requirement for partial or total inundation during the course of their life-cycle (for example, to initiate propagule production or germination). An example is the rare Neotropical endemic *Isoetes jamaicensis* Hickey (Isoetaceae) which is found in ephemeral pools in open xerophytic scrublands of Jamaica. The dormant

corms of this species only sprout in years when enough seasonal rainfall has occurred to fill the pools (Hickey, 1981). Our definition excludes terrestrial plant species that occur in wetland systems, such as Brazilian várzea floodplain grasslands (e.g., Lewis et al., 2008; Carvalho et al., 2013). These species are adapted to tolerate periodic flooding, but usually do not have a physiological requirement for inundation as an essential part of their population survival strategy (e.g., Pezeshki et al., 2001).

Aquatic macrophytes are generally assumed to show mostly broad world distributional ranges (Cook, 1985; Santamaría, 2002; Les et al., 2003). However, no study has hitherto critically tested this hypothesis. Santamaría's study, for example, was based on a sample of <1% of the total number of macrophyte species identified by Chambers et al. (2008). Cook (1985) provided useful information on the native versus introduced status of 172 macrophyte species across 16 phytogeographic regions of the world, but this still accounts for only a small percentage of the known number of macrophyte species.

Over the past century numerous studies have documented the factors, acting at scales from local to global, which drive macrophyte diversity and assemblage patterns (including the classic studies by Butcher, 1933; Gessner, 1955; and Haslam, 1978). Latitude and altitude (elevation above sea level: a.s.l.) are good examples of spatio-environmental drivers (Crow, 1993; Jones et al., 2003; Tapia Grimaldo et al., 2016). Other large-scale drivers are hydrological regime, including annual evapotranspiration and flood patterns, especially flood pulse duration, size, and frequency (van Geest et al., 2005; Varandas Martins et al., 2013); alkalinity/acidity (Vestergaard and Sand-Jensen, 2000); and land use influences (particularly nutrient inputs from agriculture and other sources influencing trophic conditions in inland waters: e.g., Akasaka et al., 2010). Other studies have looked in more general terms at drivers of macrophyte diversity and distribution in different types of waterbodies: lakes and reservoirs (e.g., Rørslett, 1991; Pulido et al., 2015; Alahuhta et al., 2017), rivers and canals (e.g., Murphy and Eaton, 1983; Kennedy et al., 2015; Tapia Grimaldo et al., 2016, 2017); and wetland waterbodies (e.g., Santos and Thomaz, 2007; Zhang et al., 2018). Finally, several recent studies have examined how a range of different potential driving factors may influence macrophyte diversity and/or assemblage at large scales, for example across Europe (e.g., Chappuis et al., 2012), and regions of Africa and the Neotropics (e.g., Kennedy et al., 2015; Morandeira and Kandus, 2015; Tapia Grimaldo et al., 2017).

No previous study has assessed either the global distribution of the total world pool of aquatic macrophyte species, or the global-scale drivers of the patterns of macrophyte diversity, assemblage and endemism seen at world scale. Macroecological studies aimed at understanding the factors controlling large-scale patterns of species distribution and diversity have had only a limited previous application to macrophytes (Jacobs and Wilson, 1996; Carvalho et al., 2009). Several factors are thought to be important in influencing the macroecology of aquatic plants in inland waterbodies. For example, both endo- and ectozoochorous vectors of macrophyte propagule movement, acting primarily via migratory birds, are increasingly recognised as factors that play a role in explaining macrophyte global distributions (e.g., Coughlan et al., 2017). Historic (late Quaternary) climate-change impact, and species range contractions and expansions following such major climatic changes, have also been shown to predict the current distribution of (especially) high-latitude narrow-range endemic species. This is relevant both for macrophytes (Stuckey, 1993; Nies and Reusch, 2005) and also some less-mobile animal species (Sandel et al., 2011; Morueta-Holme et al., 2013). Older planetary-scale phytogeographical vicariance factors also undoubtedly play a role in determining macrophyte global distributions, at least for some species (e.g., Les et al., 2003; Volkova et al., 2018). A good example is *Pistia stratiotes* L. (Araceae), an ancient species first known from fossil seeds dated to the Late Oligocene/Early Miocene, some 20-18 million years ago (Renner and Zhang, 2004). This plant is thought to be native to both the Neotropics and Afrotropics, with a lineage that may stretch back to Gondwana, according to IUCN (http://www.iucnredlist.org/details/168937/0). However, studies which have used molecular estimates of divergence time to compare closely-related aquatic macrophyte taxa that show discontinuous intercontinental distributions do not generally support ancient divergence as an explanation for macrophyte distribution. These

studies have generally found divergence times that are "far too recent to implicate continental drift as a major determinant of discontinuous distributions in aquatic plants" (Les et al., 2003).

Given the current multiplicity of threats, such as global climate change, eutrophication, pollution, and habitat loss (Zhang et al., 2017), to the continued survival of plants that live in inland freshwater and brackish habitats, there is a clear need to examine the existing world status of macrophyte species. This includes identifying the locations of macrophyte diversity hotspots across the planet; assessing the global distributions of these plants and the species assemblages that they form at a world scale; and determining the potential large-scale drivers (both current and historic) of their diversity, endemicity and species distributions. Such information can usefully inform predictions of how world, and more local-scale, macrophyte diversity and distribution patterns may respond to current and future global change issues.

In this study we aimed to determine the worldwide distribution of all known vascular macrophyte species recorded in freshwater and brackish inland waterbodies and watercourses between 80°N and 60°S, which are the generally-accepted latitudinal limits of macrophyte colonization (Chambers et al., 2008). Some macroalgae (notably charophytes) and aquatic bryophytes (e.g., Lang and Murphy, 2011) also meet the definition of aquatic macrophytes; however, because of limited geospatial records (often with poor taxonomic resolution), these plants were not included in this study.

Our first objective was critically to address the assumption (Cook, 1985; Santamaría, 2002; Les et al., 2003) that macrophytes generally have broad global distributions. Our second objective was then to consider the following specific questions:

i. what proportions of macrophyte species, and which ones, have broad, intermediate and narrow global range sizes?

ii. can we identify sets (assemblages) of macrophyte species which tend to occur together geographically, at world scale? If so, what is the species composition of these world-scale assemblages; where do they occur; and are their geographical locations associated with potential large-scale spatial and environmental drivers of macrophyte community assemblage?

iii. how important is large-scale (global ecozone), and smaller-scale endemism in macrophytes?

iv. where are the global hotspots and coldspots of macrophyte α-diversity located?; and

v. can macrophyte α- and γ-diversity be related to geoclimatic and other world-scale gradients associated with latitude, and potentially also other spatial and environmental variables; and if so, how?

2. Material and methods

2.1. Sources of species records, inclusion and exclusion criteria

Post-1950 world geospatial records (Appendix A2) were collated for all known vascular aquatic macrophyte species. Native, introduced, naturalised and invasive records were all included, but not records known to be for plants growing in cultivated locales such as ornamental parks or gardens. Records were primarily extracted from the Global Biodiversity Information Facility (GBIF: www.gbif.org) world species distribution database. The world coverage of this resource for macrophyte species is generally good except in parts of Africa, and also across much of the area occupied by the former USSR: e.g., see distribution maps for *Stuckenia pectinata* (L.) Börner (Potamogetonaceae) provided in Appendix A3. The data were supplemented by information from primary datasets held by the authors, and records from other online geospatial plant species distribution information resources covering large areas of the planet (notably Flora of China, Flora of North America, e-

Monocot, CJB African Plant Database, IUCN Red List, www.plantarium.ru, plant.depo.msu.ru, Flora do Brazil, Flora Zambesiaca, www.binran.ru/resursy/informatsionnyye-resursy/tekuschie-proekty/caucasian-flora, and Cátalogo de angiospermas acuáticas de México), as well as numerous sources providing more local distributional coverage (see Appendix A4).

Records were included from inland freshwater and brackish waterbodies, but not coastal or marine low-salinity aquatic habitats such as coastal salt-marsh channels, and the Upper Baltic Sea; nor highly-saline or highly-alkaline ("soda lake") inland waterbodies such as the Great Salt Lake (Utah, USA), and Lake Bogoria (Kenya). Records were collated from sites located in the six major global ecozones which primarily provide macrophyte habitat. These are the Palaearctic: Pa; Nearctic: Nea; Neotropics: Neo; Orient (Indomalaya): Or; Australasia: Aus; and Afrotropics: Afr (Olson et al., 2001). Records were not included from the two world ecozones which only provide very limited areas of freshwater habitat: Antarctica, and the Pacific islands of Oceania.

For a few species, for which only pre-1950 records are held by GBIF, information from the IUCN Red List, or other reliable sources, was used where appropriate to provide evidence for the continued current existence of populations of these species in the field. An example is *Utricularia vitellina* Ridl. (Lentibulariaceae), for which GBIF holds only two records, both collected from Malaya (as the country was then known) in 1911 (www.gbif.org/occurrence/search?taxon_key=4087456). Although it is difficult to read the collector's handwriting on the label of the 1911 voucher specimen held by the Royal Botanic Gardens, Kew, England, the collection locality can be read as "Gunung Tahan" which is Malay for Mount Tahan, a mountain in the State of Pahang in peninsular Malaysia (GBIF transcribes the locality as "Guunia Taliak", which is meaningless in Malay and highly unlikely to be correct). The 2018 IUCN Red List information for this species states that the plant is "... presently known from the summit areas of two mountains (Mount Tahan and Mount Korbu) in Peninsular Malaysia", growing "...at the margins of streambanks". Mount Korbu lies approximately 100 km from Mount Tahan, in the adjacent Malaysian State of Perak. This case emphasises the importance of the detective work sometimes needed in the data collation exercise to extract relevant information from the records.

A number of species were excluded from the database for a range of reasons, which are outlined in detail in Appendix A5. In brief these grounds for exclusion included the following:

i. Obviously-inaccurate or impossible records were ignored if other evidence, such as original site location information on voucher specimen sheets, could not be found to correct the error. These spatially-erroneous records are usually produced as a result of incorrect assignment by the database system(s) of latitude and longitude coordinates;
ii. Also not used in the study were fossil records, and records derived from cultivated locales such as Botanic Gardens and garden centres selling aquatic plants;
iii. Some species were omitted from the study because they lack post-1950 geospatial records (and any corroborative information on more-recent occurrences, from the IUCN Red Data List or other reliable sources), or because available information about their location(s) was too vague to allow them to be correctly spatially positioned in one or more world gridcells (see below);
iv. A few macrophyte species were identified which have ranges entirely restricted to areas outwith the target ecozones. These are species endemic to parts of Oceania (e.g., the Hawaiian Islands), or the sub-Antarctic islands (Appendix A6). These species were not used in the study. Similarly, the records from Oceania or sub-Antarctic islands for a few other species which occurred in one or other of these areas, as well as having records within the target ecozones, were also excluded from the analysis.

2.2. Species nomenclature

Species names were verified against The Plant List (www.theplantlist.org), except in a few cases, mainly where newly-described species had not yet been included in The Plant List, or where newly-published taxonomic

studies (up to 2018) had resulted in very recent revision of species nomenclature. In these source publications were followed for nomenclature, e.g., Cheek and Lebbie (2018) for a newly-described macrophyte species (*Lebbiea grandiflora* Cheek (Podostemaceae); and Cheek et al. (2017b) for their recent taxonomic revision of the genus *Inversodicraea* Engl. (Podostemaceae). Hybrids, e.g., *Potamogeton* × *cooperi* (Fryer) Fryer (Potamogetonaceae) were not included in the study. Infraspecies, e.g., *Ranunculus penicillatus* subsp. *pseudofluitans* (Syme) S.D. Webster (Ranunculaceae), were included only as records within their parent species distribution. Plant species with broad ecological niches, and often broad ranges (Slatyer et al., 2013; Kennedy et al., 2017), and which often occupy a wide range of habitat types (i.e., aquatic to terrestrial), were included in this study only when it was clear that they occurred in an aquatic habitat. Such species were marked (see Appendix A2) to show that a part of their distributional range comprises non-aquatic records. Examples include *Ranunculus repens* L. (Ranunculaceae), *Corrigiola litoralis* L. (Molluginaceae) and *Echinochloa crus-galli* (L.) P.Beauv. (Poaceae). Species with unresolved names (according to The Plant List) were included only if at least one other major data source (e.g., IUCN Red List: www.iucnredlist.org; GBIF: www.gbif.org) also lists the species. We discuss nomenclatural issues potentially affecting the dataset further in Appendix A5.

2.3. World ecozones, gridcells and species range calculation

Records were compiled for 238 10x10° latitude x longitude gridcells (see Appendix A7) spanning the six ecozones of the world where macrophytes primarily occur. Ten of these gridcells straddled ecozone boundaries, and were hence each subdivided into two smaller gridcells for record collation purposes (see below); so the total number of sampling units was actually 248. We follow the generally-accepted demarcation of major terrestrial ecozones (Olson et al., 2001), but use a simplified version of the recent division of the Nearctic/ Neotropics (Escalante et al., 2010), along the 30°N latitude line; and also mainly use this line of latitude to divide the Palaearctic from the Orient, which again is a simplification of the actual dividing line. Records were included between 60°S and 80°N (the effective latitudinal limits of macrophyte colonisation: Chambers et al., 2008), totalling 182.47 x 10^6 km^2 total world gridcell area (i.e., area of gridcells containing land within the six ecozones included in the study). This value is larger than the normally-accepted value for total land world area (approximately 135 x 10^6 km^2, excluding Antarctica and the Pacific islands of Oceania) because moderate to substantial areas of sea occupy part of the gridcell in 126 of our total 248 world gridcells. Gridcells subdivided for record collation purposes included eight gridcells split between the Palaearctic and Afrotropics, with the dividing line primarily following the Tropic of Cancer through the Sahara and Arabian Deserts. Two more gridcells were subdivided (approximately along the Wallace Line) between the Orient and Australasia. The two individual parts of each of these 10 gridcells were treated as separate units for most purposes in the study, but were combined for mapping purposes. Gridcells were assigned codenames combining their parent ecozone and an alphabetic code for the 10x10° unit (listed in Appendix A7). For example, gridcell Or F, within the Orient ecozone (20-30°N; 120-130°E), contains the island of Taiwan, adjoining Ryukyu Islands, and a coastal strip of the mainland China province of Zhejiang.

A single species record, in a single year during the study period, within a gridcell unit was considered the minimum needed to give a positive result ("gridcell hit") for a given species in that unit of planet surface. To correct for variation in gridcell area (calculated for each 10° latitude band using spherical trigonometry, taking account of the fact that the Earth is a non-perfect sphere) with increasing latitude (approximately 1.222 x 10^6 km^2 at the Equator (0-10°N or S), dropping to 0.317 x 10^6 km^2 at 70-80°N: see Appendices A2 and A7), actual extent of occurrence (gridcell EOO: broadly following the concept used in the IUCN Red List;) was calculated as the sum total area of gridcells in which a given species is present (see also: Gaston, 1994; Carvalho et al., 2009; Morueta-Holme et al., 2013). Although the concept is basically similar, our approach differs in detail from that of IUCN, which states that "Extent of occurrence (EOO) is defined as the area contained within the shortest continuous imaginary boundary that can be drawn to encompass all the known, inferred or projected sites of present occurrence of a taxon". We only included "known" post-1950 record sites, not "inferred or projected"

ones, and we did not attempt to draw "the shortest continuous imaginary boundary" for individual species distributions, but simply summed the gridcell areas of those 10x10° gridcells in which each species occurred. Within the EOO of a species IUCN go on to state that "Area of occupancy (AOO) is defined as the area within its extent of occurrence which is occupied by a taxon, excluding cases of vagrancy". The number of species present in each of seven bands of gridcell EOO (from >50% down to <5% of total world (six ecozones) gridcell area, within the geographical criteria laid out above) was calculated. The world total macrophyte species pool (γ-diversity for all gridcells combined) and ecozone species pool values (γ-diversity for all gridcells per ecozone) were also calculated, for all species and for ecozone-endemic species (see below).

2.4. Endemic and threatened species

Numbers of ecozone-endemic species (defined as species with world distribution limited to only one of the six global ecozones included in the study) were determined. Total number of species per gridcell (S_T), and of ecozone-endemic species per gridcell (S_E), were assessed as measures of species-richness (S: α-diversity measured at gridcell level), allowing us to identify regional hotspots for macrophyte α-diversity. Species which were present in only one of the six ecozones included in the study, but which also had records from one or both of the two ecozones not considered (islands of Oceania and Antarctica) were not included in the endemics list: e.g., *Ranunculus pseudotrullifolius* Skottsb. (Ranunculaceae). An additional index of macrophyte species diversity (IH: Carrara et al., 2017), sometimes called the "endemic diversity index", was also calculated, following the method of Kier and Barthlott (2001). This is calculated by first counting the number of gridcells in which each species occurs (N), then calculating the inverse of this number (1/N), then finally summing the inverse values calculated for each species present per gridcell, for each of the 248 gridcell units included in the study. The IH index is useful in separating gridcells which have the same value of S_T, but contain species of differing geographic range sizes. Those with a high proportion of cells containing species with small range size (more likely to be endemic) will score higher than those which contain species with broader range size (Kerr, 1997).

Information from the IUCN Red List was used to identify macrophyte species most threatened: i.e. listed as Near Threatened, Vulnerable, Endangered or Critically Endangered (NT/ VU/ EN/ CR) by IUCN. For most of these species an estimated AOO (km^2) is given by the IUCN Red List. These data were extracted and used to compile a list of macrophyte species and families which may be considered at greatest risk of loss (Appendix A8).

2.5. Spatio-environmental variables and mapping

Data were collated for four global-scale spatio-environmental variables measured per gridcell:

- **i.** mid-gridcell latitude (LAT: absolute °N or S of the Equator) of each of the 228 10x10° latitude x longitude gridcells and the 20 subdivided gridcells, obtained from world atlas maps.
- **ii.** total land cover (LC) expressed as the % area of gridcell (or subdivided gridcell) not occupied by the sea. This serves as an indicator of the coastal/island vs. continental status of each gridcell.
- **iii.** area of waterbodies (WB) present in a gridcell (or subdivided gridcell) expressed as the % of land area covered by inland waterbodies such as lakes, reservoirs, rivers and canals. This serves as a surrogate measure of landscape aridity.
- **iv.** area of high-altitude land (ALT) per gridcell (or subdivided gridcell) expressed as the % of gridcell land area >1000 m above sea level.

For the LC and WB metrics, land use maps (400 m resolution at the equator) were obtained for each continent from the WaterBase United Nations University project (www.waterbase.org). The layers were merged into one

and re-projected to the Behrmann Equal Area Cylindrical Projection using ArcMap™ v. 9.3.1. Then, using a 10° gridcell shapefile and the Zonal Histogram tool in the ArcMap™ Spatial Analyst toolbox, the area of each land use (and subsequently total land cover) per gridcell was obtained. The same procedure was followed for the ALT metric, deriving the area of high-altitude land per gridcell, using the USGS Global Multi-Resolution Terrain Elevation Data 2010 (Danielson and Gesch, 2011), downloaded at 30 arc-sec resolution (https://topotools.cr.usgs.gov/GMTED_viewer/) and reclassified to values 0 and 1 (below and above 1000 m a.s.l., respectively). Data for ALT, WB, and the world distribution of the three diversity indices (S_T, S_E, IH), were subsequently mapped at 10x10° resolution using ESRI® ArcMap™ v. 9.3.1 Continents, with the 10° grid shapefile obtained from the ArcGIS website (www.arcgis.com).

2.6. Statistical analyses

Classification of the gridcells x species matrix, using a subset of the broadest-range species within the full dataset (species with gridcell EOO >25% of world area), was undertaken using the hierarchical divisive procedure Two-Way Indicator Species Analysis (TWINSPAN: Hill, 1979), to determine end-sets of gridcell sample-groups (produced with a minimum division eigenvalue of 0.1; and with minimum sample group size for division set to 40 to prevent the production of large numbers of small end-groups); and species assemblages occurring in gridcell-groups (see Appendix A9). Two Saharan gridcells (AfrN B: western Algeria/ northern Mauritania; and AfrN E: south-east Libya/ south-west Egypt: both only including land south of the Tropic of Cancer) were excluded from the analysis because they had no records of any of the common species present. AfrN B is the only world gridcell with no macrophytes present at all, in our dataset. AfrN E has only a single macrophyte species recorded, the limited-range (EOO <25% of world area) *Marsilea aegyptiaca* Willd. (Marsiliaceae).

For inferential statistical comparisons among TWINSPAN gridcell-groups, Ryan-Joiner testing of the data for inland waterbody coverage (WB), land cover (LC), altitude (ALT), and mid-gridcell latitude (LAT) showed these to be non-normal variables, which could not be normalised by transformation. These were hence assessed using the non-parametric Kruskal-Wallis procedure, with post-hoc separation, for variables showing significant outcomes ($p < 0.05$), utilising Dunn's test. Three other variables (gridcell total species α-diversity: S_T; gridcell endemic species α-diversity: S_E; and gridcell "endemic diversity index": IH) were also compared among TWINSPAN gridcell-groups and ecozones. Ryan-Joiner testing, followed by assessment of homogeneity of variance using Levene's Medians test, showed that these variables were suitable for comparison by one-factor analysis of variance (ANOVA), with post-hoc mean separation using Tukey's least significant difference test. Inferential analyses were performed using Excel with the Real Statistics Data Analysis Tools add-in. Classification, data summaries and inferential statistical analyses are based on the dataset as compiled in October 2018, but updates continued to be made to the master datafile thereafter (see Appendix A2), up to the date of final ms submission (June 2019).

Polynomial curve-fitting procedures (using Excel functions) were utilised to provide an initial indication of relationships between macrophyte α-diversity and latitude. We then developed this further by using boosted regression trees (BRTs), which are not affected by the distribution of the data (De'ath, 2007), to determine the environmental variables that best predict variation in species-richness and endemism across the full data set of world gridcells. The approach of Elith et al. (2008) was employed to find the optimal number of trees. Tree complexity was set at three with a learning rate of 0.001, and with the bag fraction set at 0.75, meaning each individual tree was constructed using 75% of the data, with its predictive ability tested on the remaining 25% (Elith et al., 2008). A poisson distribution was used for the count-type data of species and endemic richness (S_T, S_E) and a gaussian distribution for the continuous endemism index (IH). Whilst BRTs are excellent at finding patterns in large complex data sets, using thousands of small trees to find variables that best predict, in this case, world macrophyte species-richness and endemism metrics, they do not provide a good means to visualise

the data. Thus, we used a single univariate regression tree (De'ath, 2002), pruned using a cost-complexity measure, to show how the different explanatory variables relate to patterns in S_T, S_E and IH.

3. Results

3.1. World macrophyte species pool: family composition and species distributions

In total, geospatial records were collated for 3457 species within 456 genera, and 93 families of vascular macrophytes (Table 1, Appendix A2). The taxonomic breakdown is clubmosses and horsetails: 6 species (0.17% of total); ferns and fern-allies: 209 species (6.05%); dicotyledons: 1520 species (43.97%); and monocotyledons: 1722 species (49.81%).

There is a clear trend in the data (Table 1) in terms of proportions of the total world macrophyte species pool with broad, intermediate and narrow global ranges. Only 42 species (1.21% of total) have broad ranges (EOO >50% of world total gridcell area for the six ecozones included in the study); 718 species (20.77%) have intermediate ranges (EOO 10-50% of world area); and 2697 species (78.02 %) each have narrow range-size, with EOO <10% of world area. Of the total, at least 667 species (mostly with broad or intermediate ranges) commonly occur across a range of non-aquatic as well as aquatic habitats (Appendix A2), e.g., *Cyperus rotundus* L. (Cyperaceae), *Echinochloa colona* (L.) Link (Poaceae), and *Juncus effusus* L. (Juncaceae). This inflates the range-size of these species over and above their true aquatic range area. A similar effect is likely for macrophyte species which are also grown as food crops, or otherwise cultivated, e.g., rice: *Oryza sativa* L. (Poaceae). Unfortunately it was not feasible to check every record given by GBIF for information on habitat or cultivation status (even if provided, which often was not the case, from the sub-samples of records examined) of these very common species. There are simply too many records to make this possible: in July 2018, for example, GBIF held 365,622 occurrence records for *O. sativa* alone. This point should be borne in mind when considering the range data calculated here for such macrophyte species: their true aquatic (and non-cultivated, where relevant) ranges are certainly overestimated.

At the other extreme, 658 species individually have ranges limited to a single gridcell. These extreme narrow-range endemic macrophyte species are dominated by five families: Podostemaceae (168 species: 25.5% of single-gridcell species); Araceae, especially *Cryptocoryne* spp. (65 species: 9.8%); Isoetaceae (55 species: 8.4%); Cyperaceae (46 species: 7.0%); and Eriocaulaceae (35 species: 5.3%). The peak latitudinal occurrence of single-gridcell endemic species occurs in the tropical latitude bands of 0-10 and 10-20°absolute (with 222 and 220 species, respectively), followed by 20-30 and 30-40°absolute (93 species each), 40-50°absolute (28 species), and just two species in latitude band 50-60°absolute. No single-gridcell endemic macrophyte species occur at latitudes >60°absolute.

Of the total set of macrophyte species, a small proportion (41 species: 1.2% of world total: Appendix A2) are considered to be major invasive exotic species in at least one part of the world (several more are considered to be locally-invasive in one or more regions). Almost all of these are broad or intermediate-range species, including many well-known examples, such as *Eichhornia crassipes* (Mart.) Solms (Pontederiaceae), which has an EOO of 57.28% of world area, and is native to the Neotropics.

The world macrophyte species pool is dominated by 12 high-diversity families, each having >100 macrophyte species. With the exception of Isoetaceae, these families all include >1 genera containing aquatic macrophytes (Table 2). Cyperaceae is the family with the greatest number of macrophyte species worldwide, though many of these are multiple-habitat species, which also occur in non-aquatic conditions (see Appendix A2). Comparing among ecozones the Neotropics has the greatest number of species per family for eight families (Cyperaceae, Poaceae, Podostemaceae, Plantaginaceae, Eriocaulaceae, Lentibulariaceae, Isoetaceae and Alismataceae); the Afrotropics has two families with greater macrophyte species diversity than in any other ecozone

(Hydrocharitaceae and Lythraceae); and the Orient (Araceae), and Palaearctic (Potamogetonaceae) one apiece. With one exception the high-diversity families all have moderate to high numbers of species occurring in every ecozone. However, Podostemaceae has large numbers of species present only in sub-tropical to tropical areas, especially in the Neotropics, and to a lesser extent in the Afrotropics and Orient. At least one species of Podostemaceae occurs in all three of the other ecozones. In the Nearctic this family is represented solely by *Podostemum ceratophyllum* Michx., which has a restricted distribution in the eastern and southern USA. Podostemaceae species are slightly better-represented in the Palaearctic and Australasia.

At the other extreme, 14 families are represented in the dataset by single macrophyte species. These are Lycopodiaceae (sole macrophyte species: *Lycopodiella inundata* (L.) Holub); Polypodiaceae (*Microsorum pteropus* (Blume) Copel.); Philydraceae (*Philydrum lanuginosum* Banks and Sol. ex Gaertn.); Rapateaceae (*Spathanthus bicolor* Ducke); Restionaceae (*Platycaulos mahonii* (N.E.Br.) H.P.Linder and C.R.Hardy); Cannaceae (*Canna glauca* L.); Zingiberaceae (*Hedychium coronarium* J.Koenig); Droseraceae (*Aldrovanda vesiculosa* L.); Portulacaceae (*Neopaxia australasica* (Hook. f.) Ö. Nilsson); Balsaminaceae (*Hydrocera triflora* (L.) Wight and Arn.); Apocyanaceae (*Oxystelma esculentum* (L. f.) Sm.); Pedaliaceae (*Trapella sinensis* Oliv.); Rosaceae (*Comarum palustre* L.); and Molluginaceae (*Corrigiola litoralis* L.).

The world macrophytes dataset includes 1292 species (37.5% of the total) whose conservation status had been assessed by IUCN, up to October 2018 (Appendix A2). Of these, 93 were assessed as Data Deficient (DD), and 954 were allocated Least Concern (LC) status. The remaining 244 species (7.2% of the total set of world macrophyte species; or 18.9% of the macrophyte species assessed up to 2018 by IUCN) are considered to be at risk (assessed by IUCN as Near Threatened NT; Vulnerable VU; Endangered EN; or Critically Endangered CR). All of these at-risk species have narrow ranges, in some cases with an extremely small area of occupancy (AOO). A notable example is *Ledermaniella yiben* Cheek (Podostemaceae), known from only a single gridcell (AfrN Q: Sierra Leone). This has an estimated AOO of just 1 km^2 and the species' survival is seriously threatened by reservoir construction (Cheek et al., 2017a). Table 3 and Appendix A8 provide details of the 29 families, and their constituent species, which are most at risk, based on possessing both limited range size (AOO <20,000 km^2) and listing by IUCN as NT/VU/EN/CR. Of these, 28 species had a total AOO on the planet's surface 1000-20,000 km^2, 38 species 100-1000 km^2, 43 species 10-100 km^2 , 57 species 1-10 km^2, and 41 species <1 km^2. The macrophyte species with the smallest known AOO is *Polypleurum longicaule* M.Kato (Podostemaceae): just 100 m^2, with its only known population being located in a waterfall near Udon Thani, northeastern Thailand (in gridcell Or C: https://www.iucnredlist.org/species/194789/8901730). Podostemaceae formed the biggest single group of limited-range threatened species. The extreme within-ecozone endemism of many Podostemaceae species is particularly noticeable in the dataset, where species distributions which are limited to a single river system in a single gridcell are far from unusual.

3.2. Macrophyte distributions and diversity

Most macrophyte species are limited to a single world ecozone, with only a few occurring in all six ecozones examined in this study, but this is not so at genus level (Fig. 1). Most macrophyte genera show limited occurrence, in one to four ecozones, but there are a substantial number which have broad ecozone distribution (with 100 genera occurring in all six ecozones), a fact which has long been known (e.g., Gessner, 1955; Chambers et al, 2008). Many of these broad-range genera contain numerous species that occupy non-aquatic as well as aquatic habitats (at least 40% of the 174 genera found in 4-6 ecozones fall into this category: see Appendix A2), such as *Cyperus* L. (Cyperaceae) and *Brachiaria* (Trin.) Griseb. (Poaceae). However it remains the case that more macrophyte genera (62%) have "narrow" distributions (occurring in no more than half the world's six ecozones that primarily support macrophyte populations) than those with broader distributions.

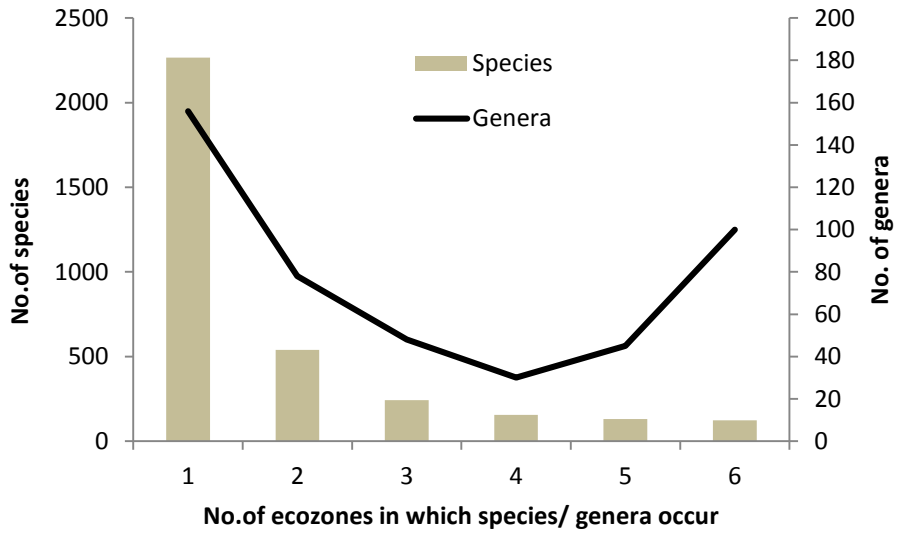


Fig. 1. Number of macrophyte species and genera occurring in one to six world ecozones

Good examples of the extremes of macrophyte species world ranges are provided by *Isoetes eludens* J.P.Roux, Hopper and Rhian J.Sm. (Isoetaceae: EOO: 0.550% of global gridcell area, within the six world ecozones examined here) and *S. pectinata* (EOO: 67.530%). GBIF world species distribution maps for these and selected other species are provided in Appendix A3. Most macrophyte species have EOO values intermediate between these extremes, with the majority (>75%) individually occupying <10% of global land gridcell area (Table 1; Appendix A2). A typical example is *Hydrostachys polymorpha* Klotzsch (Hydrostachyaceae), an Afrotropical endemic with an EOO of 3.892% (see Appendix A3).

Many macrophyte species (excluding those few which exhibit broad-range worldwide distributions, and are usually present in all ecozones: for example, *S. pectinata*) show global distribution patterns restricted to a single area of the planet. These range from the tight distributions seen for narrow-range species such as *Marsilea capensis* A. Br. (Marsiliaceae), with an EOO of 4.221%, all in southern Africa; to intermediate-range distributions spanning more than one ecozone, but still clearly occupying a single area of the planet's surface, e.g., *Bolbitis heteroclita* (Presl.) Ching (Blechnaceae). This species has an EOO of 10.795%, with records scattered across Australasia and the Orient (as well as some records from islands of Oceania located close to the Pacific margin of these two ecozones, which were not included in its calculated EOO).

Some species show more disjunct distributions. Perhaps the clearest examples are plants which occur solely in the Neotropics and Afrotropics, with a substantial trans-oceanic distance separating them across the Atlantic, e.g., *Torenia thouarsii* (Cham. & Schltdl.) Kuntze (Linderniaceae), with an EOO of 20.239%, recorded from 16 Afrotropic and 15 Neotropic gridcells.

A third principal global distribution pattern is seen for Northern Hemisphere high-latitude species which have circumboreal distributions around the planet through the Palaearctic and Nearctic. A range of near-continuous through to more disjunct patterns is evident. A good example is *Carex chordorrhiza* L.f. (Cyperaceae), an intermediate-range species (EOO: 26.495%), which shows substantial variation in commonness across its range, from very rare in the UK, occurring only in northern Scotland, to common in Fennoscandia, Iceland, Canada, and Alaska, though less so in Siberia. For world distribution maps of this and other species mentioned as examples in this section, see Appendix A3.

The most diverse gridcells in the world (macrophyte diversity hotspots) are located in tropical to subtropical areas of the Neotropics (Fig. 2). Comparisons of the significance of differences in mean values for the three diversity indices (see Appendix 10 for detailed results), across ecozones, show a similar picture for S_T and S_E. The

Neotropics have significantly higher gridcell diversity, using these metrics, than all other ecozones. However, there is greater variation in IH among ecozones. The Neotropics still have the highest diversity using this index but the value is not significantly different from that of the Orient. In addition, the Palaearctic and Nearctic have significantly lower IH values than all other ecozones except Australasia.

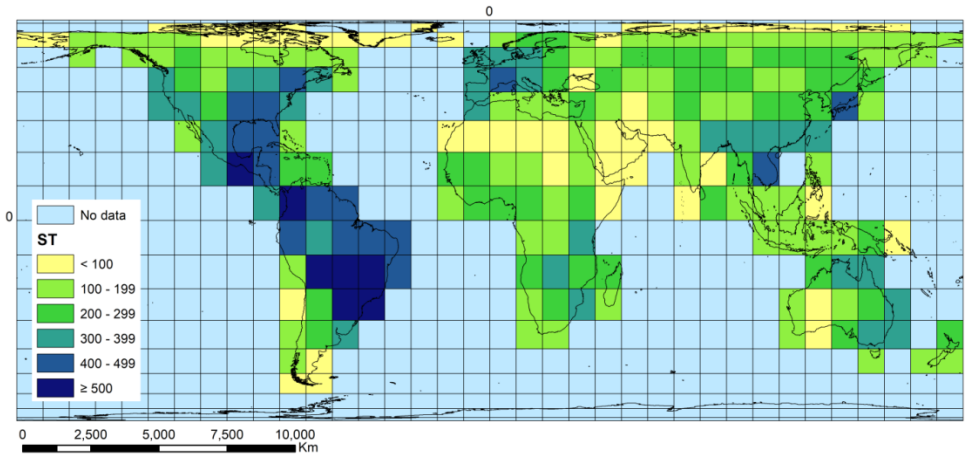
An example of the value of the IH index in helping to separate the species-richness of gridcells showing similar values of S_T is seen in the adjacent Nearctic gridcells NeaS A (Washington State/Vancouver Island) and NeaS I (California). These have identical values for S_T at 386 species per gridcell (Appendix A2). However, the IH index shows that NeaS I is more diverse, in terms of supporting macrophyte species with more restricted range, having a value of 34.07 compared to 25.42 for NeaS A. California has a number of restricted-range endemic macrophyte species, such as *Navarretia bakeri* H. Baker (Polemoniaceae), which do not occur elsewhere on the planet, and these plants contribute to the higher value of IH for this gridcell.

The richest gridcell on Earth for macrophyte diversity is NeoN N: Brasilia/Goiás (Brazil: 10-20°S; 40-50°W), with 625 species, 350 of them endemic to the Neotropics. This also has the highest calculated value for IH of any world gridcell. Amongst the numerous Neotropical endemic macrophytes which occur in NeoN N some have ranges which are solely restricted to this gridcell. Examples include *Eleocharis angustispicula* R.Trevis., *Rhynchospora testacea* Boeckeler (both Cyperaceae), and *Apinagia dissecta* (Montagn.) Engl. (Podostemaceae).

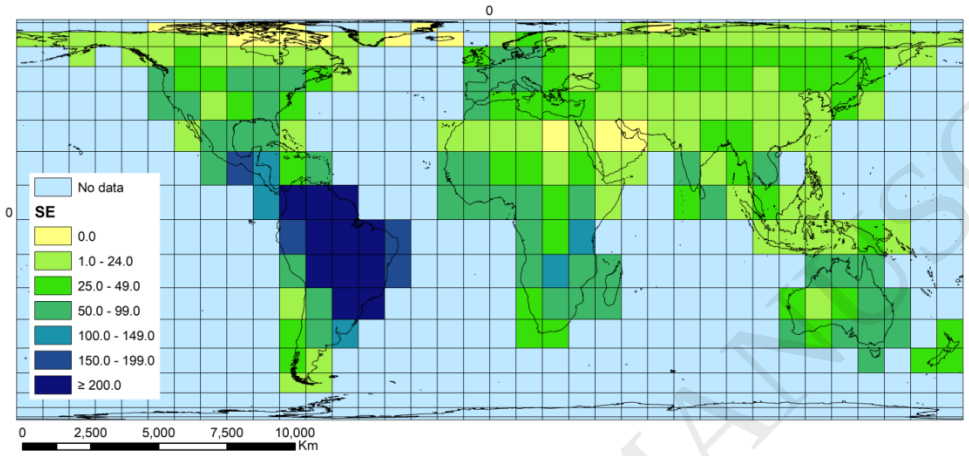
The Neotropics also has, by a large margin, the richest ecozone total species γ-diversity ("regional species pool") of the six world ecozones examined: 1566 species, with 837 ecozone-endemic species present (Appendix 10). Next are the Palaearctic (1009 species/ 262 endemics) and Orient (934 species/ 293 endemics). However when we standardise γ-diversity by total gridcell area per ecozone (to give values of S_T as number of species present per 10^6 km^2) the picture changes somewhat (Appendix A10). Now the Orient comes out top, Neotropics second and Australasia third, while the Palaearctic drops to last place on this metric.

The macrophyte hotspots with lowest values of S_T, compared across ecozones (though all still with >300 species per gridcell) were seen in the Afrotropics, whilst hotspot IH values were also relatively low in this ecozone (Fig. 2). However, this was not the case for endemic α-diversity. Afrotropical hotspot S_E values were second only to the Neotropics, as was the Afrotropical value for total ecozone-endemic species pool (Fig. 2; Appendix A10).

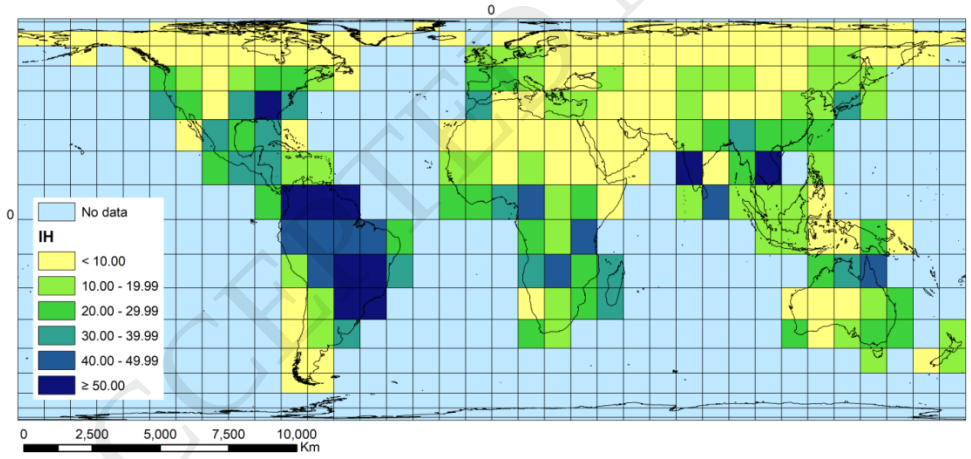
Comparing the Neotropics and Afrotropics in more detail (see Appendix A2), it is noticeable that a trend of rich to very rich macrophyte diversity occurs in Neotropical gridcells across almost the full set of gridcells present in the ecozone. Even the least species-rich Neotropical gridcell, (NeoS K: Tierra del Fuego and the Falkland Islands/Islas Malvinas) still supports 55 macrophyte species. In contrast, the Afrotropics has a tail of eight gridcells of low to very low S_T (diversity coldspots), all with ≤15 species (AfrN A-H), which are located in the Sahara or Arabian Deserts, immediately south of the Tropic of Cancer. The other two macrophyte diversity coldspots with very low S_T (each with 9 species per gridcell) are both Arctic: PaC T (the Yamal Peninsula in Siberia, Russia), and NeaN CC (the Boothia Peninsula of northern Nunavut, Canada).



(A)



(B)



(C)

Fig 2. World maps for distribution of three indices of macrophyte α-diversity in 10x10° latitude x longitude gridcells: A. All species diversity: S_T; B. Ecozone-endemic species diversity: S_E; C. "Endemic species diversity" index: IH. Values are number of species recorded per gridcell for S_T and S_E, and calculated index value for IH. Detailed data for the three indices are given in Appendix A2. Position of Greenwich Meridian (0° longitude) and Equator (0° latitude) marked as "0". Scale bars are indicative only, referring to average longitudinal distance on the map.

3.3. TWINSPAN classification of world gridcells using common species

TWINSPAN classification of the gridcell samples x species dataset, utilising the 258 broadest-range species (those with a world range >25% gridcell EOO), produced a set of nine gridcell end-groups (labelled A-I: with division eigenvalues for final group sub-division of 0.10 or more), and supporting eight species assemblages (labelled I-VIII: Table 4). The generally low eigenvalues imply a substantial overlap of species occurrence among gridcell-groups, and similarly for sample-occurrence overlap between species assemblages: Appendix A9. From Table 4 it is, however, clear that the species comprising each assemblage show marked differences in their preferential occurrence among the nine gridcell-groups. For example, species common in tropical and subtropical areas (Assemblages I-III) occur more frequently in Groups B, C and D. A typical example is *Fuirena umbellata* Rottb. (Cyperaceae), an Assemblage I plant common in the gridcells of Groups B-D. It occurs in only one Group E gridcell and two Group F gridcells, and is wholly absent from the gridcells comprising Groups A, G, H and I.

Species which are more common in higher latitudes (Assemblages V-VIII) characteristically dominate gridcell Groups A, G, H and I. An example is *Potamogeton gramineus* L. (Potamogetonaceae) an Assemblage VII species, common in most of the gridcells comprising Groups G-I, and also present in 22.5% of the gridcells making up Group A. However this species is absent from Group B, rare in the gridcells comprising Groups C, D and F, and is present in only 27.2% of the gridcells of Group E. Assemblage VII species are notably predominant in the gridcells making up Group A.

Assemblage IV is not predominant in any gridcell-group, but contains numerous species with wide range distributions covering both tropical and temperate gridcell-groups. Examples include the invasive species *Egeria densa* Planch. and *Hydrilla verticillata* (L.f.) Royle (both Hydrocharitaceae); as well as cosmopolitan species native to both warm and cooler regions of the planet: e.g., *Lemna gibba* L. (Araceae) and *Najas marina* L. (Hydrocharitaceae): see Appendix A3 for species world distribution maps.

Substantial differences exist between the sets of gridcells making up each gridcell-group determined by TWINSPAN analysis (Table 6) in terms of the three macrophyte α-diversity metrics used here (S_T, S_E, IH), the ecozone occurrence of their component gridcells, and also the four environmental variables detailed in Appendix A2.

Taking ecozone geographical composition first, only one gridcell-group comprised units entirely located in a single ecozone: Group D was made up solely of Neotropical units (Fig. 3). Two others (Groups C: Australia, Orient; H: Nearctic, Palaearctic) contained gridcells drawn from two ecozones, though these were completely different in each case. The rest had a more varied ecozone composition. Group E was the most variable in terms of ecozones represented, with only Oriental gridcells being absent from this group.

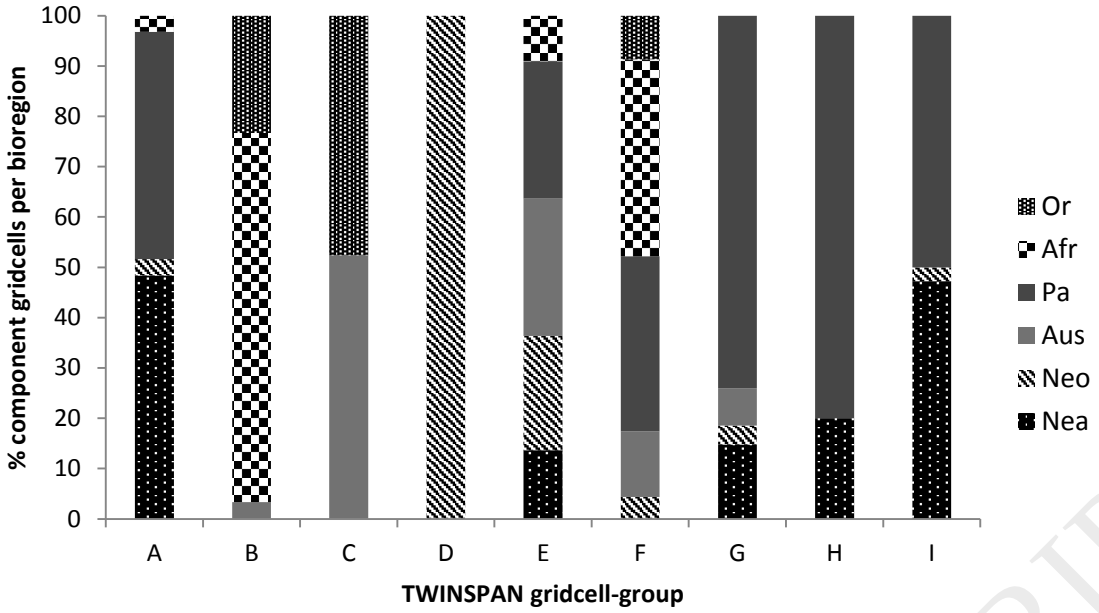


Fig. 3. Percentage composition of TWINSPAN gridcell-groups in terms of component gridcell ecozone location: Or: Orient; Afr: Afrotropics; Pa: Palaearctic; Aus: Australasia; Neo: Neotropics; Nea: Nearctic.

Considering the results of the TWINSPAN sample classification (see Table 5 for supporting inferential statistics for the results presented below) primarily in relation to gridcell species α-diversity, there is strong evidence that Group D has the highest diversity, by a substantial margin Group D entirely comprises Neotropical gridcells, with a mean latitude of 15°absolute (mid-cell latitude range for this mainly Southern Hemisphere group is 5-35°absolute: i.e., tropical to sub-tropical). The S_T value for Group D is about 10 times higher than the mean S_T value of Group A. Group D has an intermediate WB score suggesting a moderate occurrence of available aquatic habitat for macrophyte growth, and also a low-moderate value for ALT indicative of mainly low-lying relief. Group D comprises a fairly high proportion of coastal gridcells and those containing oceanic islands, with relatively low LC values.

The set of gridcells comprising Group E has the second highest mean S_T. This group occurs in latitude range 15-45°absolute, comprising gridcells from both the Northern and Southern Hemispheres. Many of the component gridcells of the group are located in Mediterranean biome regions, both in the Mediterranean area itself and also in Australia and the Cape region of South Africa. Of the six target ecozones, only the Orient is unrepresented in this group (Fig. 3). Group E shows a full range of terrain from low-lying to quite high elevation. In terms of WB score the group is low-intermediate, suggesting generally semi-arid to humid conditions, and with reasonably abundant aquatic habitat available for macrophyte colonization.

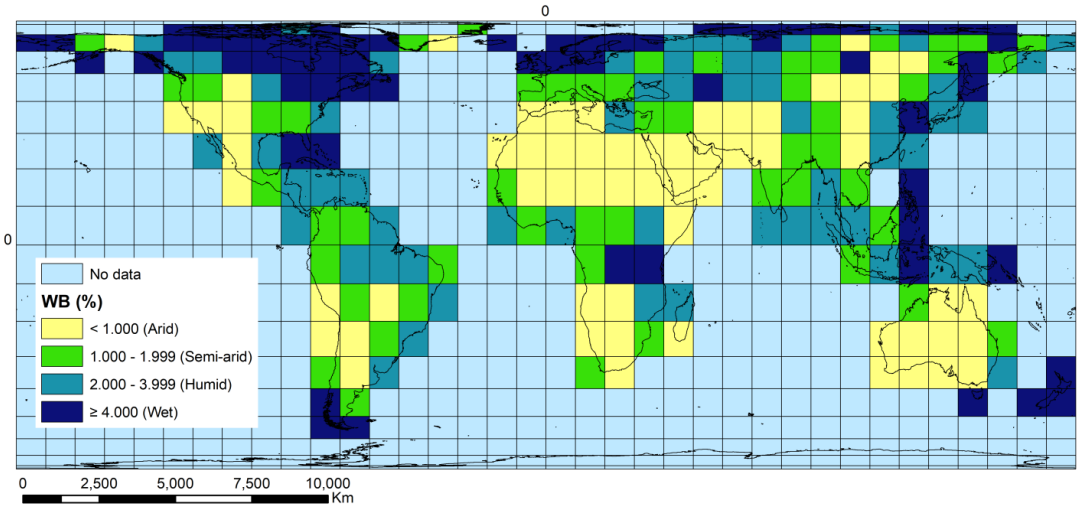
Gridcell-groups B, C, G and H are generally not significantly different from each other in terms of S_T and IH. These four groups also have similar, intermediate mean S_E. However, the four groups are located at widely differing latitudes. Two (Groups B and C) are tropical, with mean mid-cell latitudes around 12°absolute, while the other two are located in temperate latitudes, with their gridcells centred in the range 40-50°absolute. There are no significant differences in WB values (Table 5) between the four groups, and they generally comprise a mix of semi-arid to humid gridcells, likely to have a moderate to abundant presence of aquatic systems. Similarly, there is no significant difference in mean ALT across the four groups. Geographically, the four gridcell-groups are more separated from each other, though each contains a mix of units from at least two ecozones.

Finally, Groups A, F and I form a set of gridcells which can be considered collectively in the light of their low macrophyte diversity. The mean values for S_T of all three groups are much lower than for the rest, lying in the range 34–137species, compared with a range of 188–433 species per gridcell for the higher diversity gridcell-groups. Group A also has extremely low S_E, with no gridcell in the group having more than 11 endemic species present. Groups F and I have slightly higher, but still low, mean values for S_E. Clearly, one or more environmental factors adversely impacting macrophyte colonisation and growth are likely to be affecting the areas containing the gridcells making up these three groups.

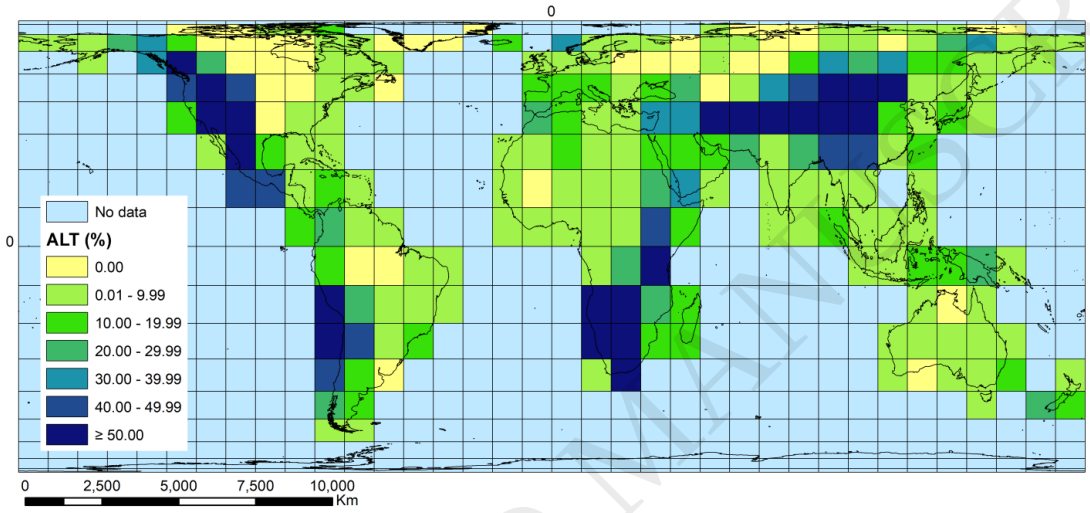
In the case of Group A, a strong candidate for such an adverse factor is LAT. The harsh climate (highly-stressful to macrophyte survival and growth) of the sub-Arctic to Arctic gridcells which largely comprise this group is highly likely to have a severe impact on macrophyte colonisation success and survival, and hence diversity. Group A gridcells are located at higher latitude (mean: 67.9°N) than the other groups,. Group I is also a high-latitude group, with all its gridcells located in the northern boreal region, within both the Palaearctic and Nearctic ecozones. In contrast, Group F is not a high-latitude group (mean mid-cell latitude: 24.5°absolute) so is unlikely to be severely influenced by extreme cold stress conditions.

Landscape aridity, as measured by the WB index (see Fig. 4; Table 5), is highly likely to be a factor influencing macrophyte diversity for one of the three low-diversity gridcell-groups. Whilst Groups A and I are the two most humid gridcell-groups on this measure, this is not so for Group F. This group has a WB index much lower than all other gridcell-groups. Nearly all the gridcells of Group F are located in arid to semi-arid areas, located in five of the six world ecozones examined (Fig. 4: no Nearctic gridcells are present in Group F). A good example of a Group F gridcell which has very few inland aquatic systems, and hence extremely low values of WB, is AfrN D (a Sahara Desert gridcell located in south-west Libya: south of the Tropic of Cancer). While inland waterbodies are not completely absent from this gridcell (a few oasis pools and ephemeral streams are present, and a small lake (Mare de Zoui) is located in the Tibesti Mountains, within the gridcell: https://en.wikipedia.org/wiki/Tibesti_Mountains#Flora) they are sparsely distributed, and offer only very limited habitat for aquatic plant growth.

Finally, rather unlikely as a factor influencing the low diversity of these three groups is ALT (see Fig. 4; Table 5). Although Group A is predominantly composed of low-lying landscapes, Groups F and I contain a mix of landscapes from low to high altitude, similar to the rest of the gridcell-groups, with little evidence that environmental factors associated with altitude might be influencing their plant communities.



(A)



(B)

Fig. 4. World maps for distribution of two spatio-environmental variables in 10x10° latitude x longitude gridcells: A. Inland waterbodies presence (WB: % of gridcell land area occupied by waterbodies such as lakes, reservoirs, rivers and canals); B. High-altitude land presence (ALT: % of gridcell land area >1000 m above sea level, a.s.l.). Detailed data for the two variables are given in Appendix A2. Position of Greenwich Meridian (0° longitude) and Equator (0° latitude) marked as "0". Scale bars are indicative only, referring to average longitudinal distance on the map.

3.4. Global macrophyte diversity distribution

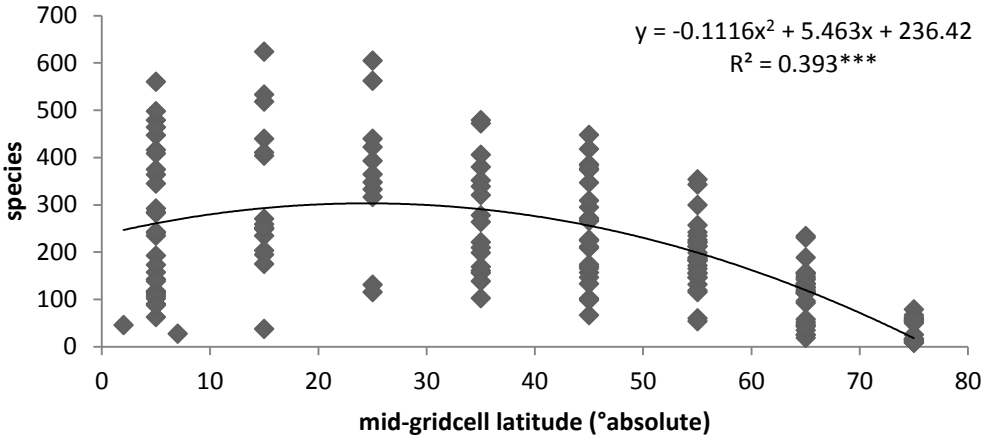
3.4.1. Macrophyte gridcell α-diversity per 10° latitude band

Macrophyte total species α-diversity per gridcell (S_T), excluding arid gridcells (with WB ≤ 1% of gridcell land area), shows a significant polynomial relationship with latitude. Maximum values of S_T occur in latitude band 10-30°absolute: i.e., tropical to subtropical conditions, both North and South of the Equator, dropping off slightly in the Tropics closest to the Equator, and also steadily declining through the temperate zone (in both hemispheres), sub-Arctic, and into the Arctic in the Northern Hemisphere (Fig. 5). For ecozone endemics only (S_E), again excluding arid gridcells, the relationship is noisier but still displays a significant polynomial curve ($R = 0.525$, $n=177$, $p<0.001$), with maximum S_E values again seen in gridcells located in the latitude band 10-30°absolute. A significant polynomial relationship is also seen for IH ($R = 0.637$, $n=177$, $p<0.001$), in non-arid

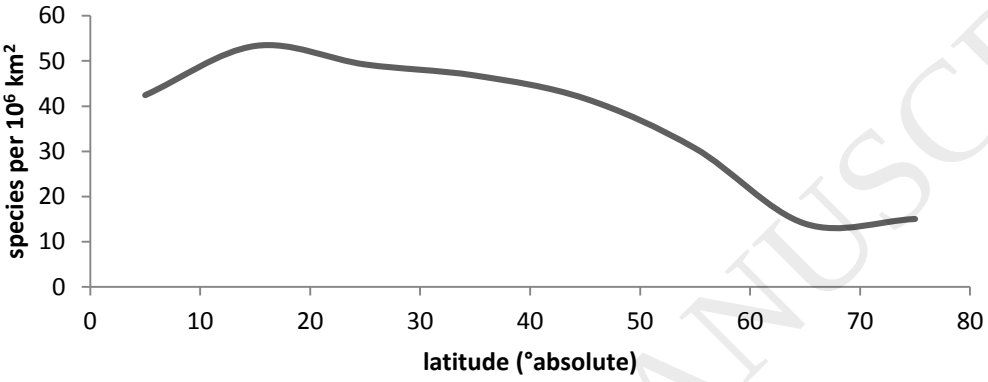
gridcells, though in this case the fitted curve shows evidence for a somewhat flatter peak of IH, throughout the tropical to subtropical latitude band (0-30°absolute).

3.4.2. Total macrophyte γ-diversity per 10° latitude band

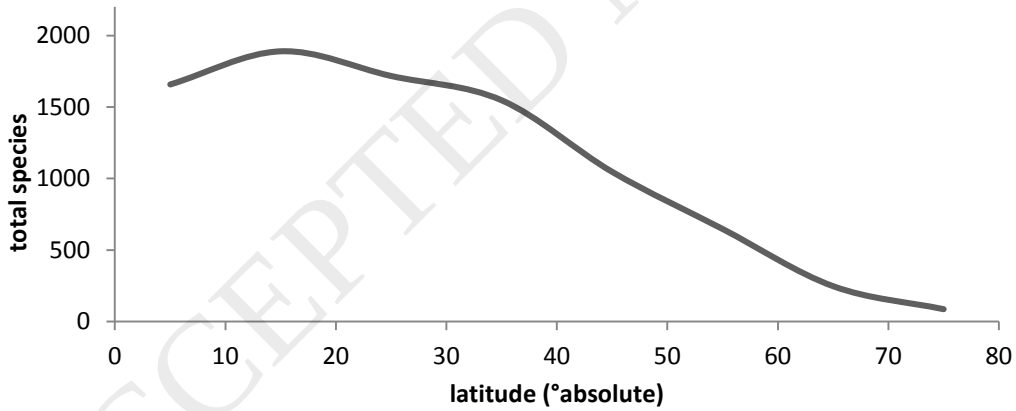
The γ-diversity versus latitude curves !Fig. 5) show similar patterns for both total number of species recorded, for all gridcells located in each 10° latitude band ("species pool per latitude band") from the Equator to 80°absolute, adjusted for total area of gridcells present (within the six target ecozones) in each band (to give values as species per 10^6 km^2); and also for unadjusted data (total number of species recorded per 10° latitude band). In both cases the maximum occurs in latitude band 10-30°absolute.



(A)



(B)



(C)

Fig. 5. A. Gridcell all-species macrophyte α-diversity (S_T) vs. latitude: excluding arid gridcells (WB ≤ 1% of gridcell area): R = 0.627, n=177, p<0.001: B. Total number of species recorded for gridcells located in each 10° latitude band from equator to 80°absolute, adjusted for total area of gridcells present in six ecozones located within each band (to give values as species per 10^6 km^2); C. As (B) but for unadjusted data (total number of species recorded per 10° latitude band).

3.4.3. Boosted Regression Trees analysis

The predictive power of the boosted regression trees (BRTs) is high, with the four explanatory variables (LAT, WB, ALT and LC) explaining >60% of the variance in each of the all-species richness (S_T), endemic richness (S_E) and endemic diversity index (IH) datasets. The pruned trees (see Appendix A10 for diagrams) contain 3, 2 and 2 variables for S_T, S_E and IH, respectively, and explain 64% (from Lat, WB and LC), 66% (from LC and Lat) and 63% (Lat and LC) of variability, respectively. Exclusion of altitude only slightly decreases the predictive ability of the trees. The single pruned classification trees per diversity measure (S_T, S_E, IH) explain less variance than their boosted equivalents (see Appendix 10) but provide a good visualisation of how the variables combined to shape diversity patterns. LAT, WB and LC are all important in shaping richness and endemism, but the values at which an effect was found differ between S_T and S_E. For total richness there is no great change across the lower latitudes with WB and LC more influential below 60° latitude. It should be noted that the pruned classification trees (Appendix A10) are partial dependency plots, which show the modelled effect of each variable while accounting for the effects of the other variables in the model. For richness patterns, WB is important at the low end (i.e., more arid conditions) of its gradient whereas LC is influential across the entire length of the gradient. The patterns for S_E and IH are similar, although the boosted regression tree suggests that LC is more important than LAT for S_E, whereas the reverse is true for IH. Both the simple regression and the boosted tree outcomes show that there are marked changes in endemism at lower latitudes, at 18.5 and 26 °absolute, respectively for S_E and IH, with LC and WB potentially important at these lower latitudes. The pattern of influence of LC is much clearer for IH, with greater endemism occurring the higher the LC value, along the length of the gradient, for values higher than around 30 % gridcell land cover, which is also reflected at low latitudes by the simple regression tree.

For S_E and IH there is some evidence to suggest a peak of diversity explained by latitude within the 10-20°absolute latitude band, with the response curves both tailing away slightly towards the Equator, and more strongly in the opposite direction, towards the Poles, though with indications (in the case of S_E) of a "shoulder" of relatively high diversity at around 30-40°absolute. This pattern also exists for S_T but is less pronounced, with maximum diversity predicted from the response curve across a broader range of latitude, 10-40°absolute, than is seen for the other two diversity measures.

4. Discussion

4.1. Extent of macrophyte global distributions

The evidence presented here suggests that small global ranges are the norm for most macrophyte species, as for other biota (e.g., Sandel et al. 2011). Reinforcing this finding is information on the limited-range of macrophyte species reported by IUCN as “at risk”, which we summarise in Section 3.1, above (Table 3; Appendix A8). All of these at-risk species have narrow ranges, in some cases extremely small. Moreover, range size based on actual area of occupancy (AOO) is substantially smaller than the gridcell EOO measures that were primarily used to determine world ranges of species in our study. The lowest EOO value found for aquatic macrophytes in this study was 0.703 x 10^6 km^2, which was recorded for two species. *Callitriche transvolgensis* Tzvelev (Plantaginaceae) is found in a single 50-60°N gridcell: PaW L (Nizhny Novgorod, Russia); and *Elatine ojibwayensis* Garneau (Elatinaceae) is also restricted to a single 50-60°N gridcell: NeaN S (western Québec, Canada). In comparison, we provide evidence that at least 207 macrophyte species have AOO values <20,000 km^2.

One previous study has offered evidence against the suggestion that macrophytes generally have narrow world distributions. In a re-examination of the data of Santamaría (2002) for 18 macrophyte species which were stated in that study to have “broad” world distributions (though the term “broad” was not quantitatively defined), we found: (i) two species that needed to be eliminated (*Zannichellia pedunculata* Rchb. is now a sub-

species of broad-range (EOO >91 x 10^6 km^2: range size >50% of total world gridcell area) *Zannichellia palustris* L. (Potamogetonaceae), and the wetland plant *Glyceria nubigena* W.A. Anderson (Poaceae) does not meet the definition of aquatic macrophyte); (ii) agreement on two species having narrow ranges, and six species having broad ranges; and (iii) disagreement on the range category of eight species, seven of which we determined to be in the intermediate-range (EOO 18.2-91 x 10^6 km^2: 10-50% of total world gridcell area) rather than broad-range category, while *Zannichellia major* Boenn. ex Reichenb. (Potamogetonaceae) is a narrow-range species (EOO <18.2 x 10^6 km^2: <10% of total world gridcell area). Hence, the majority of the species that Santamaría (2002) used as evidence of macrophytes showing broad world distributions do not, on the basis of our definitions of range-size categories, have broad-range EOO distributions.

Data additionally provided in Santamaría (2002) for a set of species from one genus across a limited area of the planet's surface (Europe) showed that 61% of aquatic *Ranunculus* species had a range of 20-40° latitude, within Europe, compared to 39% of aquatic *Ranunculus* species with ranges of ≤ 20° latitude. Our analysis of a similar constrained dataset (i.e., aquatic *Ranunculus* spp. occurring in 24 10x10° gridcells across the European part of the Palaearctic (West) ecozone (stretching from 20°W to 50°E, and from 30-70°N: see Appendix A7) showed reasonably comparable results: 48% of 25 European aquatic *Ranunculus* species had latitudinal ranges of up to 20° whereas the remaining 52% had ranges of 20-40° latitude across Europe. However, comparison of proportional gridcell occurrence for aquatic *Ranunculus* in Europe, using the data from Santamaría (2002) and our data, does not support this finding. There is a consistent trend in both datasets showing that only a few aquatic *Ranunculus* species have broad European distributions when range is measured on this basis (see Appendix 10 for data). This matches with our findings for the world set of aquatic *Ranunculus* species. On a global basis, only three of a total of 54 *Ranunculus* species in our database (Appendix A2) show broad world range.

There is evidence that broad-range species tend to show a rather high occurrence (compared with their proportional occurrence in the total world species pool) in the species-sets comprising the total species α-diversity of individual gridcells (S_T). This trend towards S_T dominance by species which do not exhibit narrow global range size occurs across the board, but is particularly pronounced in low α-diversity gridcells. For those with individual α-diversity in the range 1–11 species per gridcell, the average proportional occurrence of narrow-range species in the gridcell flora is 24.8%, compared with 75.2% for intermediate/ broad-range species. Explanation of this observation needs further work.

There is also evidence for the existence of a fairly substantial proportion (38%) of macrophyte genera having broad-range at ecozone level, with their constituent species occurring in 4 or more ecozones. This is substantially higher than the corresponding proportion of macrophyte species which occur in four or more ecozones (11.6% of world total species): Fig. 1. Part of the reason for this may be the prevalence of genera containing multi-habitat species (i.e., species that have populations occurring in both aquatic and non-aquatic conditions) within this broad-range ecozone category. At least 70 of these genera fall into that category, and it is possible that the generalist survival strategies that such broad-range species usually possess (Grime, 1979; Kennedy et al., 2017) may facilitate their ability to find suitable locales to colonise across multiple ecozones.

Although most macrophytes show narrow global distributions, the question of the extent to which introductions and invasions have influenced macrophyte world distributions certainly requires further investigation (Cook, 1985). Most of the major invasive macrophyte species have broad ranges (see Appendix 2) but an example of a species of more limited world distribution that causes invasive aquatic weed problems is *Lagarosiphon major* (Ridl.) Moss (Hydrocharitaceae), which has an intermediate EOO of 18.779 x 10^6 km^2 (10.29% of world gridcell area). Outwith its native range of southern Africa, where it never causes nuisance problems, this species is a major invasive weed in a few other parts of the world, notably New Zealand (http://www.nzpcn.org.nz/flora_details.aspx?ID=3356). The evidence suggests, however, that the vast majority

of aquatic macrophyte species are rarely, if ever, invasive: only a few are problem species in this context (e.g., Pieterse and Murphy, 1993; Brundu, 2015).

Whether anthropogenic movement of macrophytes which are not aggressively invasive is an important vector for macrophyte dispersal is not yet fully understood. Cook (1985) provides a rare example of a study of this issue. Dispersal of species used by hobby aquarists and in water gardens is certainly of importance in some cases, though such species often seem to find it difficult to establish in the wild in new locales, even if they successfully escape. Nevertheless there are plenty of examples of non-invasive macrophyte species that have successfully established in inland waterbodies, whether as a result of deliberate or accidental introduction, thereby extending their range into parts of the world where they are non-native. An example is water cress, *Nasturtium officinale*, which has expanded from its native European distribution to similar latitudes in the Southern Hemisphere, almost certainly by deliberate introduction, given its value as a crop plant: https://www.gbif.org/species/5708806).

4.2. World-scale macrophyte species assemblages

Although differences in the composition of macrophyte assemblages from tropical versus cooler areas of the planet have long been known, few previous studies have examined how latitude-associated factors influence macrophyte assemblages and their diversity. Our TWINSPAN classification of the 258 broadest-ranging macrophyte species (i.e., those with a world range >25% gridcell EOO) provides evidence for the existence of nine species assemblages at a world scale. Three of these assemblages comprise mainly tropical species, four encompass cool-temperate to Arctic flora, and two are more intermediate in terms of the latitudinal distributions of their component species. These world-scale assemblages tend to blur into each other in terms of the distributions of their component species, rather than being sharply defined in geographical location. Nevertheless, the species forming these assemblages are differentially distributed among eight gridcell sample-groups identified by the TWINSPAN world-scale classification, which also have significantly differing macrophyte α-diversity (though differing in detail amongst the three diversity metrics: S_T , S_E and IH).

Our finding of nine global-scale assemblages of macrophyte species extends the results of previous studies comparing macrophyte species diversity and assemblage composition between geographic regions. For example, Crow (1993) examined differences in distribution of selected macrophyte taxa between Costa Rica and the northern USA and found often similar high diversity for both regions. Jacobsen and Terneus (2001) reported on the macrophyte communities present in streams in the tropical Ecuadorian highlands compared to temperate lowland Denmark, and found differences in assemblage but, again, similar species diversity between the two areas. In contrast, Chappuis et al. (2012) examined variation in aquatic plant species distribution across Europe and northern Africa and observed high species richness in countries bordering the northern Mediterranean (Spain, France, Italy) compared to low richness along the southern and eastern Mediterranean, as well as along the eastern Baltic. Finally, Tapia Grimaldo et al. (2016) used an eigenfunction spatial analysis procedure to demonstrate that spatial and environmental drivers of macrophyte community composition differ substantially in their influence upon macrophyte communities of temperate and tropical calcareous rivers, respectively located in the British Isles and Zambia.

It would be interesting to extend our analysis to rarer macrophyte species, and also to examine in more detail intra-assemblage variation, and its causes, in the macrophyte species-sets present in different parts of the world. As well as the obvious geoclimatic differences across the Earth, factors such as habitat-type prevalence (e.g. proportional occurrence of lentic versus lotic waterbodies in a given unit of the planet's surface), connectivity, waterbody area present per unit, and relative age of waterbodies supporting macrophyte communities are all likely to be amongst the principal drivers of such variation within world-scale macrophyte assemblages (e.g., Rørslett, 1991; Santos and Thomas, 2007; Morandeira and Kandus, 2015; Pulido et al., 2015; Tapia Grimaldo et al., 2016, 2017; Alahuhta et al., 2017). This is discussed in more detail in Appendix A10.

4.3. Endemism in macrophytes

There are few previous studies of macrophyte endemism against which our results can be compared, and, with the exception of Chambers et al. (2008), they tend to be limited in geographical coverage (e.g., Cook, 1983; Moura Júnior et al., 2015; see also https://www.iucn.org/sites/dev/files/import/downloads/chapter_7_aquatic_plants.pdf). Hence our results provide substantial new information in this context. At global ecozone scale macrophyte endemism is pronounced, but with a >5-fold difference between the richest (Neotropics) and poorest (Palaearctic), expressed as ecozone-endemic γ-diversity adjusted for ecozone area (Appendix 10). Warm conditions (tropical to subtropical) appear strongly to favour macrophyte endemism.

The trend is the same for ecozone-endemic species α-diversity, measured at gridcell scale (S_E). The Neotropics ecozone has significantly higher endemic α-diversity than all other ecozones, and on this metric the Nearctic now joins the Palaearctic in showing low endemism. Whilst endemism at an ecozone scale is clearly an important feature of global macrophyte ecology, endemism is also prevalent at much smaller scales. This includes single-10x10° gridcell EEO occurrence for a substantial proportion (19%) of the world's macrophyte species, and even smaller endemic range (measured as AOO) for the IUCN "at risk" subset of macrophyte species, agreeing with the findings of previous studies on macrophytes with very restricted ranges (e.g., Camenish and Cook, 1996; Cheek et al., 2017a; Cheek and Lebbie, 2018). We discuss some possible causes of the observed world patterns of macrophyte endemism in more detail in Appendix A10.

4.4. Global hotspots and coldspots of macrophyte diversity

Our analysis and mapping of macrophyte α-diversity, measured as S_T, S_E and IH, together provide a clear picture of the location of macrophyte α-diversity hotspots on the planet's surface (Fig. 2; Appendix A10). The highest-diversity gridcells are all located in tropical to subtropical areas of the Neotropics, especially the cerrado area of Brazil, which is noted for its generally high biodiversity (Klink and Machado, 2005). This ecozone also supports the highest macrophyte γ-diversity ("regional species pool"), though only second-highest (after the Orient) when adjusted for ecozone land area. No ecozone diversity hotspots (assessed using any of the three α-diversity measures) are located at latitudes >50°absolute.

Although others have suggested that biodiversity hotspots result from co-existence of multiple species with small niche-breadth (e.g., Bastolla et al., 2005), the available evidence does not support this explanation of macrophyte diversity distribution. Our results suggest that most macrophyte species co-existing (though we use that term with caution since we did not separate out the individual habitats, within gridcells, in which species occurred) in individual gridcells do not generally have narrow global ranges (range being closely associated with niche-breadth: Slatyer et al., 2013). Similar results have been previously reported for macrophytes in African river systems, where species co-existence within individual habitats certainly was recorded (e.g., Kennedy et al., 2017). Alternative explanations of the factors driving macrophyte α-diversity in general, and hotspot formation in particular, across the planet need to be sought.

The question of whether sampling effort influences identification of global hotspots also needs to be addressed (see Appendix A5 for information on variability in macrophyte sampling effort across the world). Rarefaction analysis undertaken for individual geographic regions has shown that sampling effort influences macrophyte diversity estimates (e.g., for macrophytes of Afrotropical versus Neotropical calcareous rivers: Tapia Grimaldo et al., 2017). It is true that some (but by no means all) of the macrophyte diversity hotspots identified by our study are located in parts of the world well-surveyed by aquatic botanists. Examples include the eastern USA, France and southern Brazil (e.g., Pott and Pott, 2000). However, in the case of two regions (Western Europe and the USA) where macrophyte survey effort is both substantial and reasonably uniform, there is still considerable variation in hotspot presence within each area (Fig. 2). It is most unlikely, for example, that France and the

British Isles differ greatly in the amount and intensity of macrophyte field survey effort, yet France is much more of a macrophyte hotspot (for all three diversity metrics used here: Fig. 2). For areas of the world that are probably less intensively-sampled, such as south-east Asia, it is also probably unlikely that botanists put in a greater macrophyte sampling effort in Indochina compared with Malaysia (see Appendix A5). Yet Indochina is a richer macrophyte hotspot than mainland Malaysia (Fig. 2). It is probably safest to say that in all likelihood sampling effort does contribute to hotspot identification but only to a limited extent. Further work would be useful to investigate this issue in more detail.

4.5. Macrophyte diversity in relation to world-scale environmental and spatial gradients

There is considerable disagreement in the literature about what, if any, relationship exists between latitude and diversity of a variety of biota, including aquatic macrophytes (e.g., Stevens, 1989; Crow, 1993; Hillebrand, 2004; Cirtwill et al., 2015). Our results show a non-linear latitudinal gradient in aquatic macrophyte diversity, together with weaker relationships with three other spatio-environmental factors. Polynomial analyses of α-diversity (S_T, S_E and IH) versus latitude, for non-arid gridcells, indicate that peak diversity occurs in a latitude band somewhat higher than 20°absolute (Fig. 5). Gridcells located in arid regions were removed from analysis because arid conditions in themselves will reduce macrophyte diversity, regardless of latitude, as shown by the outcomes of our TWINSPAN and BRT analyses. A non-linear latitudinal gradient in macrophyte γ-diversity (as "species pool per latitude band") is also apparent (Fig. 5), again peaking in a latitude band around 20°absolute. Other evidence from our study supports this finding: namely a maximum for single-gridcell occurrence endemics around 10-20°absolute; and the outcomes of the BRT analyses, which show a strong influence of latitude on all three measures of macrophyte α-diversity. Our findings also agree with Tapia Grimaldo et al. (2017) who found that in African and American hardwater rivers, macrophyte α-diversity (for individual river sites, much smaller in area than our gridcells) was higher at sites in the 20-30° compared to the 10-20° latitude band. Chappuis et al. (2012), however, observed that along a 20° to 70°N gradient extending across northern Africa and western Europe, aquatic plant richness peaked somewhat further from the Equator, between 40° and 50°N (with low richness over the 20-40° N range here due to the presence of the Sahara Desert). Collectively, the evidence on a global basis suggests that macrophyte diversity is likely to be highest in tropical to sub-tropical latitudes, slightly lower closest to the Equator, and steadily decreases with higher latitude, moving away from the sub-tropics.

The importance of the three other spatio-environmental drivers of macrophyte diversity and distribution examined here was less than the latitude effect. Land cover (LC: the proportion of individual gridcells occupied by land, or large inland freshwater bodies such as the Great Lakes, Lake Victoria, and Lake Baikal, as opposed to sea) was only a weak driver of diversity and assemblage, though stronger for endemism (S_E), as already discussed above. Altitude (ALT) also showed relatively little effect on macrophyte diversity, though there was some evidence for an effect on species distribution, with one TWINSPAN gridcell sample-group in particular being associated with higher-altitude landscapes. However, landscape aridity (measured as the occurrence of inland waterbodies (WB) within gridcell land areas) had a noticeably strong relationship with both macrophyte diversity and assemblage. BRT outcomes for WB suggested that aridity was a factor affecting diversity primarily at fairly low latitudes. This is to be expected given that the main areas of the planet experiencing semi-arid to arid conditions lie at latitudes below 40°absolute (see Fig. 4: there are a few exceptions such as the Gobi Desert).

Our study utilised only a very limited set of the geospatial and other environmental factors (e.g., Mikulyuk et al., 2011) that might influence global diversity and distribution of aquatic macrophytes. Clearly, further research is required to determine the causal factors that influence macrophyte range size and distribution, and to permit meaningful prediction of how this may be affected by future climate, or other human-mediated, global changes.

4.6. Concluding statement

In conclusion, we here present substantial evidence that macrophytes generally have narrow, rather than broad world distributions. Our findings for aquatic macrophytes strongly support the statement by Sandel et al. (2011) that "small-ranged species constitute most of Earth's species diversity". We also found that endemism is an important feature of macrophyte ecology at world scale, and that latitude is a major driver of macrophyte α- and γ-diversity, appearing also to influence the global distribution of these plants and the world-scale assemblages that they form. However, further work is clearly needed to explain the reasons for the prevalence of the small ranges of most macrophyte species, and further to determine the spatio-environmental and biological factors (including plant dispersal and reproduction modes) which may influence global distribution, diversity and endemism in aquatic macrophytes. We predict that these drivers will turn out to include a combination of latitude-related (and possibly altitude-mediated) geoclimatic conditions, both current and historic; geographical differences in biotic pressures, such as herbivory, on macrophyte populations; variation in availability of suitable aquatic habitat for macrophyte colonisation across the world; biological characteristics of the plants themselves; and also perhaps the period of time for which human beings have, in different parts of the world, been influencing the global movement of macrophyte species and impacting inland aquatic habitats, along with the precise nature, location and intensity of such anthropogenic vectors and impacts.

Declarations of interest: none

Acknowledgements

Parts of the dataset used in this study were collected during fieldwork campaigns undertaken by some of the authors (KM, AE, TD, JTG, SVM, RPM, MK, IS, KF, TCCB) in Egypt, Russia, Khazakhstan, México, Trinidad, Florida (USA), Brasil, Argentina, South Africa, Zambia, Botswana, the UK, and Ireland, during 1987-2016. This field survey work was supported by primary funding from, amongst others: The British Council; CONACYT (México); CNPq (Process 478311/2013-3), NUPELIA, CAPES (Brazil); EC/ACP Science & Technology Programme (AFS/2009/219013); Marie Curie Intra European Fellowship No.255180 (PRECISE); UK DfID DelPHE Programme; UK DEFRA Darwin Programme; and The Carnegie Trust for the Universities of Scotland. We thank the agencies concerned, and colleagues who worked with us in the field. We acknowledge the efforts made by herbaria, worldwide, which have digitised, and made accessible online, images of plant voucher specimens that they hold, and we encourage others to do likewise: the readily-accessible availability of such images and their label information was extremely useful in our study. We also acknowledge those involved in the development and ongoing maintenance of the GBIF online database, without which this study would have been impossible. We thank Sidinei Thomaz (Universidade Estadual de Maringá, Brazil), Georg Janauer (University of Vienna, Austria), and Janne Alahuhta (University of Oulu, Finland) for critically reviewing our paper prior to submission; and Jane Drummond (University of Glasgow, Scotland), an expert in geomatics, for checking world gridcell areas. A preliminary version of this paper was presented as a poster and abstract at the 15th International Symposium on Aquatic Plants in Queenstown, New Zealand, 2018: we thank the symposium and organising committees for arranging an excellent meeting.

Appendices A1 – A10. Supplementary data

Supplementary data associated with this article can be found, as Appendices A1 – A10, attached to this online paper.

References

Akasaka, M., Takamura, N., Mitsuhashi H., Kadono, Y., 2010. Effects of land use on aquatic macrophyte diversity and water quality of ponds. Freshw. Biol. 55, 909-922.

Alahuhta, J., Kosten, S., Akasaka, M., Auderset, D., Azzella, M.M., Bolpagni, R., Bove, C.P, Chambers, P.A., Chappuis, E., Clayton, J., de Winton, M., Ecke, F., Gacia, E., Gecheva, G., Grillas, P., Hauxwell, J., Hellsten, S., Hjort, J., Hoyer, M.V., Ilg, C., Kolada, A., Kuoppala, M., Lauridsen, T., Li, E.-H., Lukács, B.A., Mjelde, M., Mikulyuk, A., Mormul, R.P., Nishihiro, J., Oertli, B., Rhazi, L., Rhazi, M., Sass, L., Schranz, C., Søndergaard, M., Yamanouchi, T., Yu, Q., Wang, H., Willby, N., Zhang, X.-K., Heino, J., 2017. Global variation in the beta diversity of lake macrophytes is driven by environmental heterogeneity rather than latitude. J. Biogeogr. 44, 1758-1769.

Bastolla, U., Lässig, M., Manrubia, S.C., Valleriani, A., 2005. Biodiversity in model ecosystems, I: coexistence conditions for competing species. J. Theor. Biol. 235, 521-530.

Brundu, G., 2015. Plant invaders in European and Mediterranean inland waters: Profiles, distribution, and threats. Hydrobiologia 746, 61-79.

Butcher, R.W., 1933. Studies on the ecology of rivers: I. On the distribution of macrophytic vegetation in the rivers of Britain. J. Ecol. 21, 58-91.

Camenish, M., Cook, C.D.K., 1996. *Wiesneria triandra* (Dalzell) Micheli (Alismataceae): a rare and unusual south Indian endemic. Aquat. Bot. 55, 115–131.

Carrara R., San Blas, G., Agrain, F., Roig-Juñent, S., 2017. Towards biodiversity hotspots effective for conserving mammals with small geographic ranges. Acta Oecol. 78, 7- 14.

Carvalho, P., Bini, L. M., Diniz-Filho, J.A.F., Murphy, K.J., 2009. A macroecological approach to study aquatic macrophyte distribution patterns. Acta Limnol. Bras. 21, 169-74.

Carvalho, P., Thomaz, S.M., Kobayashi, J.T., Bini, L.M., 2013. Species-richness increases the resilience of wetland plant communities in a tropical floodplain. Austral Ecol. 38, 592-598.

Chambers, P.A., Lacoul, P., Murphy, K.J., Thomaz, S.M., 2008. Global diversity of aquatic macrophytes in freshwater. Hydrobiologia 595, 9-26.

Chappuis, E., Ballesteros, E., Gacia, E., 2012. Distribution and richness of aquatic plants across Europe and Mediterranean countries: Patterns, environmental driving factors and comparison with total plant richness. J. Veg. Sci. 23, 985-997.

Cheek, M., van der Burgt, X., Mornoh, A., Lebbie, A., 2017a. *Ledermanniella yiben* sp. nov. (Podostemaceae), critically endangered at the proposed Yiben Reservoir, Sierra Leone. Kew Bull. 72, 31

Cheek, M., Feika, A., Lebbie, A., Goyder, D., Tchiengue, B., Sene, O., Tchouto, P., van der Burgt, X., 2017b. A synoptic revision of *Inversodicraea* (Podostemaceae). Blumea 62, 125-156.

Cheek, M., Lebbie, A., 2018. *Lebbiea* (Podostemaceae-Podostemoideae), a new, nearly extinct genus with foliose tepals, in Sierra Leone. PLoS One 13: e0203603. https://doi.org/10.1371/journal.pone.0203603.

Cirtwill, A.R., Stouffer, D.B., Romanuk, T.N., 2015. Latitudinal gradients in biotic niche-breadth vary across ecosystem types. Proc. R. Soc. Lond. B: Biol. 282, http://dx.doi.org/10.1098/rspb.2015.1589

Cook, C.D.K., 1983. Aquatic plants endemic to Europe and the Mediterranean. Bot. Jahrb. 103, 539–582.

Cook, C.D.K., 1985. Range extensions of aquatic vascular plant species. J. Aq. Plant Manage. 23, 1-6.

Coughlan, N.E., Kelly, T.C., Davenport, J., Jansen M.A.K., 2017. Up, up and away: Bird-mediated ectozoochorous dispersal between aquatic environments. Freshw. Biol. 62, 631-648.

Crow, G.E., 1993. Species diversity in aquatic angiosperms: latitudinal patterns. Aquat. Bot. 44, 229-258.

Danielson, J.J., Gesch, D.B., 2011. Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010). Open File Report 2001-1073. U.S. Geological Survey, Reston, Virginia. 26 p.

De'ath G., 2002. Multivariate regression trees: a new technique for modeling species-environment relationships. Ecology 83,1105-1117.

De'ath, G., 2007. Boosted trees for ecological modeling and prediction. Ecology 88, 243-251.

Elith, J., Leathwick, J. R., Hastie, T., 2008. A working guide to boosted regression trees. J. Animal Ecol. 77, 802-813.

Escalante, T., Rodríguez-Tapia, G., Szumik, C., Morrone, J.J., Rivas, M., 2010. Delimitation of the Nearctic region according to mammalian distributional patterns. J. Mammal. 91, 1381-1388.

Gaston, K.J., 1994. Measuring geographic range sizes. Ecography 17, 198-205.

Gessner, F., 1955. Hydrobotanik – Die physiologischen Grundlagen der Pflanzenverbreitung im Wasser. Vol. I., pp.160-161. VEB Deutscher Verlag der Wissenschaften, Berlin.

Grime, J.P., 1979. Plant Strategies and Vegetation Processes. Wiley, Chichester, UK.

Haslam, S.M., 1978. River Plants. Cambridge University Press, Cambridge, UK.

Hickey, R.J. (1981). A new *Isoetes* from Jamaica. Am. Fern J. 71, 69-74.

Hill, M.O., 1979. TWINSPAN — A FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. Section of Ecology and Systematics, Cornell University, Ithaca, NY, USA.

Hillebrand, H., 2004. On the generality of the latitudinal diversity gradient. Am. Nat. 163, 192-211.

Jacobsen, D., Terneus, E. (2001). Aquatic macrophytes in cool aseasonal and seasonal streams: A comparison between Ecuadorian highland and Danish lowland streams. Aquat.Bot. 71, 281–295.

Jacobs, S.W.L., Wilson, K.L., 1996. A biogeographical analysis of the freshwater plants of Australasia. Aust. Syst. Bot. 9, 169–183.

Jones, J.I., Li, W., Maberly, S.C., 2003. Area, altitude and aquatic plant diversity. Ecography 26, 411-420.

Kennedy, M.P., Lang, P., Tapia Grimaldo, J., Varandas Martins, S., Bruce, A., Hastie, A., Lowe, S., Ali, M.M., Briggs, J., Sichingabula, H., Murphy, K.J., 2015. Environmental drivers of aquatic macrophyte communities in southern tropical African rivers: Zambia as a case study. Aquat. Bot. 124, 19-28.

Kennedy, M.P., Lang, P., Tapia Grimaldo, J., Varandas Martins, S., Bruce, A., Moore, I., Taubert, R., Macleod-Nolan, C., McWaters, S., Briggs, J., Lowe, S., Saili, K., Sichingabula, H., Willems, F., Dallas, H., Morrison, S., Franceschini, C., Bottino, F., Murphy, K.J., 2017. Niche-breadth of freshwater macrophytes occurring in tropical southern African rivers predicts species global latitudinal range. Aquat. Bot. 136, 121-130.

Kerr, J.T., 1997. Species-richness, endemism, and the choice of areas for conservation. Conserv. Biol. 11, 1094-1100.

Kier, G., Barthlott, W., 2001. Measuring and mapping endemism and species-richness: A new methodological approach and its application on the flora of Africa. Biodiv. Conserv. 10, 1513-1529.

Klink, C.A., Machado, R.B., 2005. Conservation of the Brasilian cerrado. Conserv. Biol. 19, 707-713.

Lang, P., Murphy, K.J., 2011. Environmental drivers, life strategies and bioindicator capacity of aquatic bryophyte communities in high-latitude upland streams. Hydrobiologia 679, 1-17.

Les, D.H., Crawford, D.J., Kimball, R.T., Moody, M.L., Landolt, K.E., 2003. Biogeography of discontinuously distributed hydrophytes: A molecular appraisal of intercontinental disjunctions. Int. J. Plant Sci. 164, 917-932.

Lewis, N., Murphy, K.J., Vellozo da Costa Filho, L., Campos, J.B., Thomaz, S.M., Dickinson, G., Dibble, E, Macleod, H., Brown, M., Burns, F., Campbell, E., Downie, L., Home-Robertson, P., Livingstone, P., Mackintosh, F., McArthur, N., McBean, M., O'Halloran, K., Padial, A., Palmar, J., Rickard, I., 2008. Estrutura functional e diversidade da vegetação ripária de três rios do Brazil. Cadernos da Biodiversidade 5, 10-18.

Lukács, B.A., Sramkó, G., Molnár, V.A., 2013. Plant diversity and conservation value of continental temporary pools. Biol. Conserv. 158, 393-400.

Mikulyuk, A., Sharma, S., van Egeren, S., Erdmann, E., Nault, M.E., Hauxwell, J., 2011. The relative role of environmental, spatial, and land-use patterns in explaining aquatic macrophyte community composition. Can. J. Fish. Aquat. Sci. 68, 1778-1789.

Morandeira, N.S., Kandus, P., 2015. Multi-scale analysis of environmental constraints on macrophyte distribution, floristic groups and plant diversity in the Lower Paraná River floodplain. Aquat. Bot. 123, 13-25.

Morueta-Holme, N., Enquist, B.J., McGill, B.J., Boyle, B., Jørgensen, P.M., Ott, J.E., Peet, R.K., Símová, I., Sloat, L.L., Thiers, B., Violle, C., Wiser, S.K., Dolins, S., Donoghue, J.C., Kraft, N.J.B., Regetz, J., Schildhauer, M., Spencer, N., Svenning, J.-C., 2013. Habitat area and climate stability determine geographical variation in plant species range sizes. Ecol. Lett. 16, 1446-1454.

Moura Júnior, E.G. de, Paiva, R.M.S. de, Ferreira, A.C., Pacopahyba, L.D., Tavares, A.S., Ferreira, F.A., Pott, A. 2015. Updated checklist of aquatic macrophytes from Northern Brazil. Acta Amazonica 45, 111-132.

Murphy, K.J., Eaton, J.W., 1983. The effects of pleasure-boat traffic on macrophyte growth in canals. J. Appl. Ecol. 20, 713-729.

Nies, G., Reusch, T., 2005. Evolutionary divergence and possible incipient speciation in post-glacial populations of a cosmopolitan aquatic plant. J. Evol. Biol. 18, 19-26.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R., 2001. Terrestrial ecoregions of the world: a new map of life on Earth. Bioscience 51, 933-938.

Pezeshki, S.R., 2001. Wetland plant responses to soil flooding. Environ. Exp. Bot. 46, 299-312.

Pieterse, A.H., Murphy, K.J., eds., 1993. Aquatic Weeds, 2nd ed. Oxford University Press, Oxford. 593 pp.

Pott, V.J., Pott, A., 2000. Plantas aquáticas do Pantanal. Embrapa, Centro de Pesquisa Agropecuária do Pantanal, Communicaçao para Transferência de Tecnologia, Brasília, Brasil.

Pulido, C., Riera, J.L., Ballesteros, E., Chappuis, E., Gacia, E., 2015. Predicting aquatic macrophyte occurrence in soft-water oligotrophic lakes (Pyrenees mountain range). J. Limnol. 74, 143-154.

Renner, S.S., Zhang, L., 2004. Biogeography of the *Pistia* clade (Araceae): Based on chloroplast and mitochondrial DNA sequences and Bayesian divergence time inference. Systematic Biol. 53, 422-432.

Rørslett, B., 1991. Principal determinants of aquatic macrophyte richness in northern European lakes. Aquat. Bot. 39, 173-193.

Sandel, B., Arge, L., Dalsgaard, B., Davies, R.G., Gaston, K.J., Sutherland, W.J., Svenning, J.-C., 2011. The influence of late Quaternary climate-change velocity on species endemism. Science www.sciencexpress.org / 6 October 2011 / Page 3 / 10.1126/science.1210173

Santamaría, L., 2002. Why are most aquatic plants widely distributed? Dispersal, clonal growth and small-scale heterogeneity in a stressful environment. Acta Oecol. 23, 137-154.

Santos, A.M., Thomaz, S.M., 2007. Aquatic macrophyte diversity in lagoons of a tropical floodplain: the role of connectivity and water level. Austr. Ecol. 32, 177-190.

Slatyer, R.A., Hirst, M., Sexton, J.P., 2013. Niche breadth predicts geographical range size: a general ecological pattern. Ecol. Lett. 16, 1104-1114.

Stevens, G.C., 1989. The latitudinal gradient in geographical range: how so many species coexist in the tropics. Am. Nat., 133, 240-256.

Stuckey, R.L., 1993. Phytogeographical outline of aquatic and wetland angiosperms in continental eastern North America. Aquat. Bot. 44, 259–301.

Tapia Grimaldo, J., Bini, L.M., Landeiro, V., O'Hare, M.T., Kennedy, M.P., Caffrey, J., Spink, A., Varandas Martins, S., Kennedy, M.P., Murphy, K.J., 2016. Spatial and environmental drivers of macrophyte diversity and community composition in temperate and tropical calcareous rivers. Aquat. Bot. 132, 49-61.

Tapia Grimaldo, J., O'Hare, M., Kennedy, M., Davidson. T.A., Bonilla-Barbosa, J., Santamaría, B., Gettys, L., Varandas Martins, S., Thomaz, S., Murphy, K., 2017. Environmental drivers of freshwater macrophyte diversity and community composition in calcareous warm-water rivers of America and Africa. Freshw. Biol. 62, 1511-1527.

van Geest, G.J., Wolters, H., Roozen, F.C.J.M., Coops, H., Roijackers, R.M.M., Buijse, A.D., Scheffer, M., 2005. Water-level fluctuations affect macrophyte richness in floodplain lakes. Hydrobiologia 539, 239-248.

Varandas Martins, S., Milne, J., Thomaz, S.M., McWaters, S., Mormul, R.P., Kennedy, M., Murphy, K., 2013. Human and natural drivers of changing macrophyte community dynamics over twelve years in a Neotropical riverine floodplain system. Aquat. Conserv. Mar. Freshwat. Ecosyst. 23, 678-697.

Vestergaard, O., Sand-Jensen, K., 2000. Aquatic macrophyte richness in Danish lakes in relation to alkalinity, transparency, and lake area. Can. J. Fish. Aquat. Sci. 57, 2022-2031.

Volkova, P.A., Arutynyan, N.G., Shanzer, I.A., Chemeris, E.V., Bobrov, A.A., 2018. Genetic variability of Eurasian *Nuphar* species unravels possible routes in which freshwater plants could fill their wide areas. Aquat. Bot. 145, 49-57.

Zhang, M., García Molinos, J., Zhang, X., Xu, J. 2018. Functional and taxonomic differentiation of macrophyte assemblages across the Yangtze River floodplain under human impacts. Front. Plant Sci. doi.org/10.3389/fpls.2018.00387.

Zhang, Y., Jeppesen, E., Liu, X., Qin, B., Shi, K., Zhou, Y., Thomaz, S.M., Deng, J., 2017. Global loss of aquatic vegetation in lakes. Earth Sci. Rev. 173, 259-265.

List of Figures

List of Appendices

Appendix A8. List of most “at-risk” macrophyte species.

Appendix A9. TWINSPAN assemblage and sample group data.

Appendix A10: Supplementary to Results and Discussion. Additional results diagrams and information.

Table 1. Macrophyte species extent of occurrence (EOO: as % of total world gridcell area (182.47 x 10^6 km^2) within 248 10x10° gridcells).

Range category	**% world gridcell extent of occurrence (EOO) category (total world gridcell area per category in brackets)**	**Number of species**	**% of 3457 world macrophyte species**
Broad	>50.00 (>91.0 x 10^6 km^2)	42	1.21
Intermediate	40.00-49.99 (72.8-91.0 x 10^6 km^2)	53	1.53
Intermediate	30.00-39.99 (54.6-72.7 x 10^6 km^2)	87	2.52
Intermediate	20.00-29.99 (36.4-54.5 x 10^6 km^2)	175	5.06
Intermediate	10.00-19.99 (18.2-36.3 x 10^6 km^2)	403	11.66
Narrow	5.00-9.99 (9.1-18.1 x 10^6 km^2)	554	16.03
Narrow	<5.00 (<9.1 x 10^6 km^2)	2143	61.99

Table 2. Numbers of genera and species per family for macrophyte families with ≥100 macrophyte species worldwide (excluding species endemic to Oceania or Antarctica)

Family	**Total number of genera containing macrophyte species per family**	**Total number of macrophyte species per family**	**Ecozone: number of macrophyte species per family**					
	World	**World**	**Nearctic**	**Neotropics**	**Australasia**	**Palaearctic**	**Afrotropics**	**Orient**
Poaceae	69	254	87	137	75	85	76	81
Podostemaceae	56	336	1	162	4	8	96	74
Cyperaceae	34	540	177	312	154	214	159	171
Araceae	25	159	26	61	20	24	21	90
Alismataceae	17	108	35	60	23	29	14	23
Plantaginaceae	15	153	42	58	37	49	40	43
Hydrocharitaceae	13	110	21	23	30	40	44	42
Lythraceae	8	109	17	34	15	26	43	34
Eriocaulaceae	7	132	9	76	10	12	21	32
Potamogetonaceae	6	113	38	33	30	61	16	31
Lentibulariaceae	2	102	18	46	37	9	20	13
Isoetaceae	1	125	28	52	12	25	14	7

Table 3. Macrophyte families most at risk (n = 29): number of threatened species (assigned NT/ VU/ EN/ CR status in IUCN Red List, together with an estimate of AOO <20,000 km^2) per family.

Family	Threatened macrophyte species	Total macrophyte species per family	% threatened species per family
Podostemaceae	66	325	20.3
Isoetaceae	26	125	20.8
Eriocaulaceae	16	132	12.1
Hydrostachyaceae	12	20	60.0
Lythraceae	11	109	10.1
Plantaginaceae	9	153	5.9
Alismataceae	7	108	6.5
Cyperaceae	8	540	1.5
Marsileaceae	6	45	13.3
Aponogetonaceae	6	56	10.7
Menyanthaceae	5	69	7.2
Araceae	4	159	2.5
Apiaceae	4	64	6.2
Juncaceae	3	50	6.0
Compositae	3	61	4.9
Hydrocharitaceae	3	108	2.8
Elatinaceae	2	29	6.9
Haloragaceae	2	68	2.9
Potamogetonaceae	2	113	1.6
Lentibulariaceae	2	102	2.0
Poaceae	2	254	0.8
Droseraceae	1	1	100.0
Oxalidaceae	1	2	50.0
Amaryllidaceae	1	7	14.3
Commelinaceae	1	23	4.3
Brassicaceae	1	27	3.7
Onagraceae	1	59	1.7
Ranunculaceae	1	70	1.4
Nymphaeaceae	1	75	1.3

Table 4. Species assemblage presence in gridcell-groups identified by TWINSPAN analysis of 246 world gridcells x 257 broad-range macrophyte species (each with extent of occurrence >25% of total world gridcell area). Two gridcells had no common species present so are excluded from the classification. Values are % occurrence of species comprising each assemblage in gridcells making up each sample group.

		Gridcell-group (number of gridcells per group)								
Assemblage (number of species per assemblage)	**Eigenvalue for assemblage production**	**A (31)**	**B (30)**	**C (21)**	**D (26)**	**E (22)**	**F (23)**	**G (27)**	**H (30)**	**I (36)**
I (51)	0.059	0.1	52.5	60.5	73.4	15.9	11.4	5.1	0.5	0
II (32)	0.059	0	47.4	56.2	69.0	38.2	18.9	17.2	3.0	0
III (39)	0.164	0.3	36.7	44.7	73.5	55.8	17.6	41.0	11.8	0.4
IV (21)	0.308	0.2	22.2	41.9	40.3	50.0	16.4	59.2	36.5	8.2
V (25)	0.277	2.8	22.3	31.8	35.7	53.6	20.9	81.9	70.0	29.1
VI (34)	0.155	11.9	4.1	15.7	15.3	38.8	8.4	80.0	79.3	46.2
VII (29)	0.085	44.0	0.3	3.1	4.0	13.0	5.2	58.5	76.4	86.7
VIII (26)	0.085	13.4	0	3.4	1.2	8.6	4.3	58.7	79.6	52.2

Table 5. Comparison between TWINSPAN gridcell-groups A-I for seven variables. Data are means (± standard error) for group gridcell values for (a) land cover, LC (% of gridcell occupied by land); (b) altitude, ALT (% land area in gridcell >1000 m a.s.l.; (c) inland waterbody presence, WB (% inland waterbody occurrence in land area within gridcell); (d) mid-unit latitude, LAT; (e) total species α-diversity per gridcell (S_T); (f) ecozone-endemic species α-diversity per gridcell (S_E); and (g) "endemic species diversity index" per gridcell (IH). Means labelled with a superscript letter in common, per variable, do not significantly differ (p>0.05): tests used: Kruskal-Wallis (KW) with post-hoc Dunn's test, for non-normal datasets which could not be normalised by transformation; ANOVA (AOV) with post-hoc Tukeys mean separation test, for datasets shown to be normal by Ryan-Joiner testing. Outcomes are considered significant at $p< 0.05$.

	TWINSPAN gridcell-group (mean ± standard error)										
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	***p***	**Test**
LC	44.12^{ab} ± 4.31	65.53^{abc} ± 6.38	34.43^{a} ± 6.10	57.1^{ab} ± 7.27	42.67^{ab} ± 7.03	59.20^{abc} ± 5.02±	64.62^{bc} ± 6.37	86.18^{c} ± 4.57	72.07^{bc} ± 5.68	<0.01	KW
ALT	1.44^{a} ± 0.59	17.52^{b} ± 4.60	7.74^{b} ± 1.50	13.92^{b} ± 3.69	15.54^{b} ± 4.29	18.07^{b} ± 3.81	26.81^{b} ± 5.04	24.61^{b} ± 3.80	9.15^{ab} ± 2.27	<0.001	KW
WB	5.49^{cd} ± 0.57	1.69^{b} ± 0.29	2.80^{bc} ± 0.46	2.36^{bc} ± 0.28	1.75^{ab} ± 0.33	0.35^{a} ± 0.09	2.97^{bc} ± 0.42	3.63^{bc} ± 0.92	4.78^{bc} ± 0.65	<0.001	KW
LAT	67.87^{e} ± 2.27	12.63^{a} ± 1.35	12.48^{a} ± 1.80	15.00^{ab} ± 1.72	38.91^{cd} ± 1.20	24.48^{bc} ± 1.59	42.04^{cd} ± 1.26	50.33^{de} ± 1.39	60.56^{de} ± 0.92	<0.001	KW
S_T	34.42^{a} ± 3.85	182.23^{cd} ± 16.11	228.14^{de} ± 24.24	423.23^{f} ± 23.50	325.74^{e} ± 23.85	71.26^{ab} ± 21.94	260.22^{e} ± 19.15	216.37^{de} ± 9.24	136.97^{bc} ± 4.94	<0.001	AOV
S_E	2.23^{a} ± 0.48	54.73^{cd} ± 5.72	32.48^{abcd} ± 6.07	164.92^{e} ± 17.38	62.38^{d} ± 4.81	9.96^{abd} ± 2.64	37.81^{bcd} ± 4.29	25.57^{abd} ± 1.54	11.00^{abd} ± 0.85	<0.001	AOV
IH	1.03^{a}± 0.18	21.83^{c} ± 2.71	22.93^{c} ± 3.90	41.34^{d}± 4.18	30.54^{c} ± 2.59	5.15^{a} ± 1.71	15.55^{bc} ± 1.69	9.60^{ab} ± 0.81	4.38^{a} ± 0.26	<0.001	AOV