A Journal of Macroecology

Global Ecology and Biogeography

RESEARCH PAPER

A high-resolution bioclimate map of the world: a unifying framework for global biodiversity research and monitoring

Marc J. Metzger[1]*, Robert G. H. Bunce[2], Rob H. G. Jongman[3], Roger Sayre[4], Antonio Trabucco[5,6] and Robert Zomer[7]

[1]*School of GeoSciences, University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, Scotland, UK,* [2]*Estonian University of Life Sciences, 51041 Tartu, Estonia,* [3]*Alterra, Wageningen University and Research Centre, PO Box 47, 6700 AA Wageningen, The Netherlands,* [4]*Climate and Land Use Change Program, United States Geological Survey, 12201 Sunrise Valley Drive, MS 519, Reston, VA 20192, USA,* [5]*Euro-Mediterranean Centre for Climate Change, via de Nicola 25, I-07100 Sassari, Italy,* [6]*Division Forest, Nature and Landscape, KU Leuven, Celestijnenlaan 200E, 3001 Leuven, Belgium,* [7]*Centre for Mountain Ecosystems Studies, Kunming Insitute of Botany, China*

*Correspondence: Marc J. Metzger, School of GeoSciences, University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, Scotland, UK.
E-mail: marc.metzger@ed.ac.uk

ABSTRACT

Aim To develop a novel global spatial framework for the integration and analysis of ecological and environmental data.

Location The global land surface excluding Antarctica.

Methods A broad set of climate-related variables were considered for inclusion in a quantitative model, which partitions geographic space into bioclimate regions. Statistical screening produced a subset of relevant bioclimate variables, which were further compacted into fewer independent dimensions using principal components analysis (PCA). An ISODATA clustering routine was then used to classify the principal components into relatively homogeneous environmental strata. The strata were aggregated into global environmental zones based on the attribute distances between strata to provide structure and support a consistent nomenclature.

Results The global environmental stratification (GEnS) consists of 125 strata, which have been aggregated into 18 global environmental zones. The stratification has a 30 arcsec resolution (equivalent to 0.86 km^2 at the equator). Aggregations of the strata were compared with nine existing global, continental and national bioclimate and ecosystem classifications using the Kappa statistic. Values range between 0.54 and 0.72, indicating good agreement in bioclimate and ecosystem patterns between existing maps and the GEnS.

Main conclusions The GEnS provides a robust spatial analytical framework for the aggregation of local observations, identification of gaps in current monitoring efforts and systematic design of complementary and new monitoring and research. The dataset is available for non-commercial use through the GEO portal (http://www.geoportal.org).

Keywords
Bioclimate indicators, biodiversity monitoring, climate, environmental stratification, global climate classification.

INTRODUCTION

Global climate classifications were first developed by the ancient Greeks (Sanderson, 1999), and were at the heart of the inception of biogeography in the late 19th and first half of the 20th century (von Humboldt, 1867; Köppen, 1900; Holdridge, 1947; Thornthwaite, 1948). More recently, bioclimate biome classifications have been used to underpin dynamic global vegetation models (Prentice *et al.*, 1992; Sitch *et al.*, 2003). However, these existing classifications provide limited regional detail by distinguishing only 10–30 classes globally, and with generally coarse spatial resolutions. More detailed approaches to describe global ecoregions (Olson *et al.*, 2001) rely heavily on expert judgement for interpreting class divisions, making it difficult to ensure reliability across the world and limiting their use in scientific analysis (Lugo *et al.*, 1999).

By contrast, statistically derived classifications, or stratifications[1], of land into relatively homogeneous strata provide useful spatial frameworks for comparison and analysis of ecological and environmental data across large heterogeneous areas (Paruelo *et al.*, 1995; McMahon *et al.*, 2001). Environmental stratifications do not aim to classify recognizable biomes, habitats or landscape units, but partition variation in independent variables to support analysis or statistical inference. The basic principle is that the relation between biodiversity and the environment can be expressed in a multiple orthogonal regression model (Jongman *et al.*, 2006), i.e. there is a probability of occurrence that can be explained by the environment. Because field-scale biodiversity is determined by many dependent explanatory variables, stratification of the wider environment should be based on the most important independent factors, which are climatic for the global and continental context (Klijn & de Haes, 1994). Multivariate clustering of climate data has proved successful in creating stratifications in many parts of the world (Bunce *et al.*, 1996; Leathwick *et al.*, 2003; Tappan *et al.*, 2004; Metzger *et al.*, 2005) using transparent, reproducible methods that are, as far as possible, independent of personal bias. Such stratifications have been used for stratified random sampling of ecological resources, the selection of representative study sites and summary reporting of trends and impacts bias (Jongman *et al.*, 2006). Nevertheless, no high-resolution global bioclimate classification derived from multivariate statistical clustering has been constructed until now.

This paper presents a novel global environmental stratification (GEnS), based on statistical clustering of bioclimate data so that subjective choices are explicit, their implications are understood and the strata can be seen in the global context. It provides a framework for coordination and analysis of global biodiversity observation efforts (Scholes *et al.*, 2008, 2012) and research, e.g. for targeting research and monitoring efforts, aggregating observations and for the comparison of trends within similar environments, and will be publicly available to support global ecosystem research and monitoring.

METHODS

Bioclimate indicators and data

Bioclimate indicators are directly related to plant physiological processes determining primary productivity, and therefore provide a useful basis for stratification (Leathwick *et al.*, 2003). A suite of bioclimate indicators has been developed since Köppen used observed vegetation patterns to subdivide five global climate zones into 30 classes based on various temperature- and precipitation-related indicators. Thornthwaite (1943) stressed the importance of including better measures to represent seasonality and plant available moisture, developing a classification based on humidity and aridity indices

[1] When classes are not meant as descriptive units, but specifically designed to divide gradients into relatively homogeneous subpopulations we prefer to use the statistical term *stratification*.

(Thornthwaite, 1948). Meanwhile Holdridge (1947) devised a life zone system using a three-dimensional bioclimate classification based on biotemperature, precipitation and an aridity index, and Emberger (1930) developed a tailored pluviothermic indicator for distinguishing climate zones in the Mediterranean. Although there have been several more recent classifications using bioclimate indicators to model terrestrial ecosystem distributions (Sayre *et al.*, 2009), they are now mainly used in modelling the impacts of climate change on vegetation (e.g. Sitch *et al.*, 2003; Thuiller *et al.*, 2005).

For this paper, several of the most important and contrasting methods have been reviewed to identify relevant bioclimate indicators. The resulting list (Appendix S1 in Supporting Information) is not exhaustive, but provides a wide range of relevant indicators that can be calculated using the WorldClim global climate dataset (Hijmans *et al.*, 2005), including several moisture availability indicators (CGIAR; Trabucco *et al.*, 2008; Zomer *et al.*, 2008). WorldClim has the greatest spatial resolution (30 arcsec; equivalent to 0.86 km^2 at the equator) of the available global climate datasets, enabling representation of regional environmental gradients. Appendix S1 provides an overview of the 42 variables, including an explanation of how they were calculated. To avoid negative numbers in subsequent calculations, all temperature variables were converted to kelvin (K).

Constructing the stratification

The construction of the stratification consisted of three stages. Firstly, the initial pool of variables was screened to remove those variables with very high correlations and select a subset of variables that represent the dominant global gradients. The second stage entailed the actual statistical clustering. Finally, post-processing improved accessibility of the dataset. Unless stated differently, all calculations were performed using ESRI ArcGIS 9.2 software.

Screening of the variables

High correlation is likely between many of the 42 variables (Appendix S1). To prevent the classification being weighted to the most frequently used or correlated variables, a subset was used in the clustering procedure. Firstly, a correlation matrix was calculated to identify highly correlated variables. For pairs of variables with a Pearson's correlation coefficient of 1.00 a single variable was selected and any remaining variables were omitted from further analysis. Principal components analysis (PCA), based on the covariance matrix, was performed on the remaining list to identify those variables that did not represent dominant trends in the data. Variables with eigenvector loadings > 0.1 in the first three principal components were retained for further analysis. The eigenmatrix was calculated using ERDAS IMAGINE 10.0.

Clustering

The classification followed the approach used by Metzger *et al.* (2005) in constructing the environmental stratification for

Europe. PCA was used once more to reduce the subset of input variables into a set of fewer dimensions that are non-correlated and independent and are more readily interpretable than the source data (Jensen, 1996). The first three principal components were subsequently used in the statistical clustering algorithm. Formal stopping rules (Gordon, 1996) were considered to avoid judgement in the procedure, but could not be implemented in ArcGIS and do not necessarily result a better dataset (Bunce *et al.*, 1996; Manning *et al.*, 2008). It was decided to distinguish 125 strata in the data, an arbitrary choice that still permits characterization and interpretation of the strata, whilst providing far greater detail than existing approaches.

The iterative self-organizing data analysis technique (ISODATA) (Tou & Conzalez, 1974) was used to cluster the principal components into environmental strata. ISODATA is iterative in that it repeatedly performs an entire classification and recalculates statistics. Self-organizing refers to the way in which it locates clusters with minimum user input. The ISODATA method uses minimum Euclidean distance in the multidimensional feature space of the principal components to assign a class to each candidate grid cell.

Post-processing

To provide structure and support the development of a consistent nomenclature, as well as to facilitate summarizing and reporting, it is useful to consistently aggregate the strata to a limited set of environmental zones (Bunce *et al.*, 1996; Leathwick *et al.*, 2003; Metzger *et al.*, 2005). The dendrogram tool in ArcGIS was used to derive a hierarchical diagram showing the Euclidean distance between cluster means of the strata, thus illustrating the order in which the dataset progressively combines similar environments into larger groups. The dendrogram was then used to determine the aggregation of the 125 strata into 15 to 20 global environmental zones (GEnZs), a similar number to existing global biome classifications (e.g. Prentice *et al.*, 1992) and the stratifications listed before.

The GEnZs were ordered based on the mean values of their principal component scores using the dendrogram and assigned letters starting with 'A' for the zone with the lowest value. Likewise, within each GenZ the strata were numbered by mean first principal component (PC1) score, assigning '1' to the lowest value. The strata were then assigned a unique code based on the combination of the letter (GEnZ) and number (e.g. A1 or D6). In addition, consistent descriptive names were attributed to each GEnZ based on the dominant classification variables, as detailed in the Results section. Finally, a legend was developed for the strata based on the mean scores of first three components in each stratum following Leathwick *et al.* (2003).

Comparison with existing classifications

The reliability of the patterns derived by the statistical clustering can be tested by comparing them with other datasets. This is not straightforward, because comparable datasets may not exist or may have been created in a more subjective manner (Lugo *et al.*,

1999; Metzger *et al.*, 2005). Differences between datasets could therefore reflect differences in methodology and objectives, rather than illustrating the strength or weakness of any new classification (Hazeu *et al.*, 2011). Nevertheless, it is important to demonstrate that the GEnS distinguishes recognized environmental divisions as evidenced by high correlations with independent datasets. The strength of agreement between the GEnS and nine global, continental and national climate classifications was therefore determined by calculating Kappa statistics (Monserud & Leemans, 1992).

For the Kappa analysis, the datasets that are compared must have the same spatial resolution and distinguish the same classes. To meet these requirements, the classifications were resampled and projected to the Mollweide equal area projection, and the two classifications were clipped to the largest overlapping extent. A contingency matrix was calculated to determine the aggregation, based on the greatest overlap between the GEnS strata and the classes in the alternative dataset. Kappa could then be calculated using the Map Comparison Kit (Visser & de Nijs, 2006). The alternative classifications used in this comparison were: the biomes used to underpin the World Wildlife Fund (WWF) ecoregions (Olson *et al.*, 2001); a recently updated Köppen map of the world (Peel *et al.*, 2007); the European Environmental Stratification (Metzger *et al.*, 2005); isoclimate maps for the United States (Sayre *et al.*, 2009), South America (Sayre *et al.*, 2008) and Africa (Sayre, 2011); the ecoregions map of the United States (CEC, 1997); the land classification of Great Britain (Bunce *et al.*, 1996); and a geoclimate stratification of Spain (Regato *et al.*, 1999).

RESULTS

The correlation matrix of the 42 variables, listed in Appendix S2, confirmed that there were high correlations globally among many variables. There were 10 variables with a correlation coefficient of 1.00 with at least one other variable. From these variables a subset of four readily interpretable variables was chosen for inclusion in the further analysis.

The subsequent PCA of the remaining 36 variables revealed that the first three components, explaining 99.9% of the total variation, were determined by only four variables: growing degree-days on a 0 °C base[2] (GDD), reflecting latitudinal and altitudinal temperature gradients; the aridity index (Trabucco *et al.*, 2008), which forms an expression of plant available moisture; and temperature and potential evapotranspiration seasonality, which express both seasonality and continentality (Appendix S3). These four variables were used as the input to the actual clustering.

The PCA of the four clustering variables shows that each component mainly relates to one variable, although the other variables also display some influence (Table 1). PC1, mainly determined by the GDD, and PC2, which expresses the aridity index, explain the majority of the variation.

---

[2] Reflects the annual sum of daily temperatures above 0 °C, a standard variable in vegetation and crop models to determine germination.

Table 1 (a) Eigenvalues and (b) eigenvectors for four principal components (PC) of the final clustering variables.

		PC1	PC2	PC3	PC4
(a)					
Eigenvalues		3.6×10^8	8.7×10^7	2.6×10^6	4.8×10^5
% explained		80.1%	19.2%	0.6%	0.1%
Cumulative		80.1%	99.3%	99.9%	100.0%
(b)					
Variable					
var_4	Temp. seasonality	−0.11	−0.09	0.95	−0.28
var_12	Growing degree-days on 0 °C base	0.98	0.18	0.13	−0.02
var_36	Aridity index	−0.19	0.98	0.07	0.02
var_37	PET seasonality	−0.01	−0.05	0.27	0.96

PET, potential evapotranspiration.
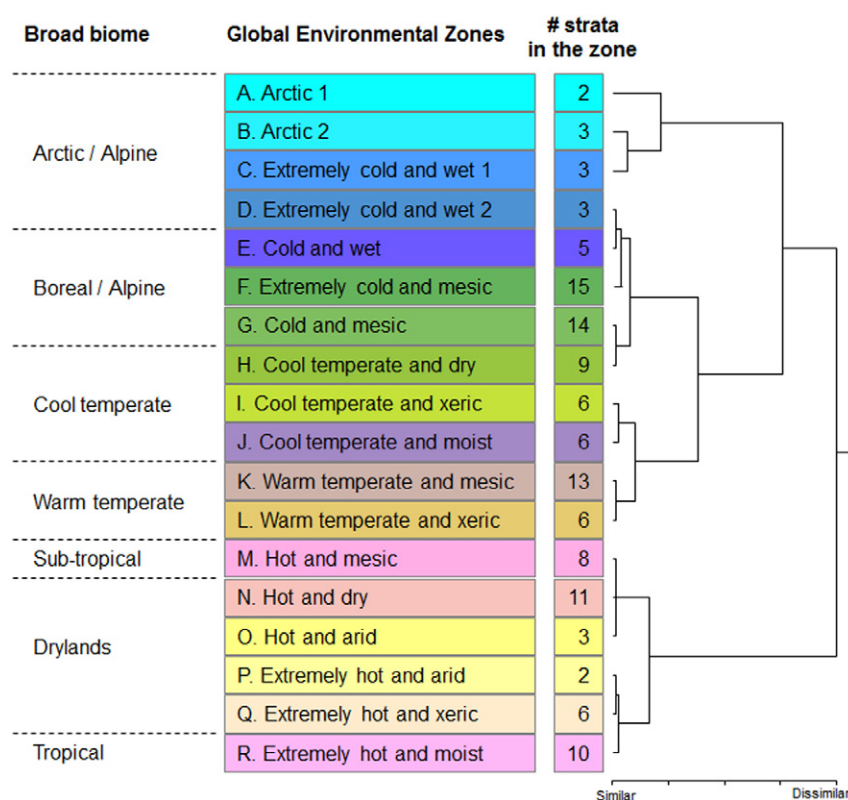
Figure 1 Dendrogram based on Euclidean distance between cluster means illustrating the relation between the 18 aggregated global environmental zones and the number of strata per zone. Colours are based on the mean principal components analysis scores of the zones, following the same method used to colour the individual strata in Fig. 2.

The 125 global environmental strata were aggregated to 18 GEnZs (labelled A to R) based on the dendrogram (Fig. 1). The GEnZs and the strata were assigned consistent codes, as described above. In practice this means that cooler strata in a GEnZ will have a lower number. In addition, the zones were given a descriptive label based primarily on mean statistics for GDD and the aridity index based on the classification in Table 2. Two zones fell within the 'Arctic' and 'Extremely cold and wet' descriptors. To provide unique names an additional numerical distinction was added, e.g. 'Arctic 1' and 'Arctic 2'.

A map legend was constructed using the mean values per stratum of the first three principal components to define the red–green–blue colour scheme. PC1 (annual temperature gra-

dient) was used to define the amount of red, PC2 (aridity) the blue coloration and PC3 (seasonality) the green coloration. The resulting legend produces a map that clearly distinguishes well-known climate zones, as well as more detailed divisions within these zones (Fig. 2; see Appendix S4 for a high-resolution map with labels in the strata). Shades of green reflect highly seasonal continental climates ranging from bright green and petrol for cool boreal to warmer chartreuse coloured regions; yellow reflects hot and arid conditions; pink reflects hot and wet tropical conditions; and purple represents wet oceanic environments.

Table 3 shows that the Kappa values for the comparison of the GEnS with existing climate classifications range between 0.54 and 0.72 indicating 'good' and 'very good' comparisons.

Table 2 The global environmental zones (GEnZs) were given descriptive names based on the mean values of (a) the growing degree-days on a 0 °C base (GDD) and (b) the aridity index for the strata. An exception was made for Arctic temperatures in which case only the label 'Arctic' was used.

(a)

GDD	Label
[0, 1000)	Extremely cold
[1000, 2500)	Cold
[2500, 4500)	Cold temperate
[4500, 7000)	Warm temperate
[7000, 9000)	Hot
[9000, ∞)	Extremely hot

(b)

Aridity index	Label
[0, 0.1)	Arid
[0.1, 0.3)	Xeric
[0.3, 0.6)	Dry
[0.6, 1.0)	Mesic
[1.0, 1.5)	Moist
[1.5, ∞)	Wet

DISCUSSION

Discussion of the results

The GEnS represents the first global high-resolution quantitative stratification distinguishing more than the basic biome divisions in the 20 to 30 classes. Major advantages of quantitative approaches, argued for by Lugo *et al.* (1999) and summarized by Leathwick *et al.* (2003), include: the much greater objectivity, consistency and spatial accuracy of the classification process; their ability to define hierarchical classifications that can be used at varying degrees of detail; and their open nature, which allows the ready incorporation of new or improved data.

The GEnS recognizes known environmental similarities, e.g. K5 identifies similar Mediterranean climates in Europe, Australia, Chile, South Africa and California; R9 links tropical parts of northern Australia to Papua New Guinea, Indochina and beyond; and G13 connects the cold mesic climates of Scotland (UK) with the south-eastern hill country of the South Island (New Zealand). Nevertheless, some heterogeneity remains when the global variation in bioclimate is partitioned in 125 strata. For example, regionally important gradients in precipitation, which can mark significant regional differences in dry ecosystems, are not always reflected sufficiently. In Israel the northern Negev Desert and the city of Tel Aviv both fall in the hot and dry stratum N6, while the latter is considered Mediterranean with greater precipitation, concentrated in the winter (Levin & Shmida, 2007). Such cases will occur more generally as a result from both artefacts in the underpinning climate data and the

number of strata that are distinguished (both discussed below). However, the results showed good comparisons with existing classifications (Table 3) and confirm recognized climate patterns. The Kappa values are similar to those reported in earlier comparisons of European classifications (Bunce *et al.*, 2002; Ortega *et al.*, 2012) and although the details of the classifications differed, there were broad similarities reflecting important divisions along major environmental gradients.

### Limitations and uncertainties

Validation challenges

Traditionally, spatial classifications are validated by comparison with *in situ* observations of the same phenomenon. However, since environmental stratifications do not aim to describe observable entities, but to partition the variation in underlying independent variables, such approaches are unsuitable here. It would be possible to test the statistical stratification efficiency for reducing errors in sample estimates (de Gruijter *et al.*, 2006), but in fact the construction of a stratification would be the first step in the design of such a sampling framework for studies covering large environmental gradients, as it will help ensure that sample units are geographically spread across the domain and that a representative sample of the diversity of environmental conditions is covered (Bunce *et al.*, 1996; de Gruijter *et al.*, 2006; Ståhl *et al.*, 2011). Alternatively, stratification efficiency can be calculated if the full population of the target parameters is known (de Gruijter *et al.*, 2006), but existing global spatial ecological and environmental data (e.g. global land cover maps) are thematically too coarse to be useful. This leaves the presented comparison with existing climate or ecosystem classification (Table 3), which provides reassurance, but should be interpreted with caution because of differences in objectives and methods.

The aim here was to develop a spatial framework for the integration and analysis of ecological and environmental data. Optimizing the stratification for any quality criteria or benchmark dataset is problematic give the current state of consistent global ecological data, and is furthermore besides the point for a generic framework. The true validity of the approach can therefore only be demonstrated through application. It is nevertheless important to be aware of specific limitations and uncertainties in the current approach, which are discussed below.

Variable selection

A wide range of bioclimate indicators were considered, but others could have been included in the analysis. Some arbitrary choices were inevitable in the statistical screening that could have influenced the outcome (e.g. eigenvector loadings > 0.1 in the first three principal components). The results nevertheless show that there is a distinct division between the dominant variables above the threshold (eigenvector loadings 0.98, 0.98, 0.94 and 0.27; Appendix S3), and the remaining variables (eigenvector loadings of 0.09 and lower). Furthermore, the final four variables represent bioclimate characteristics that are
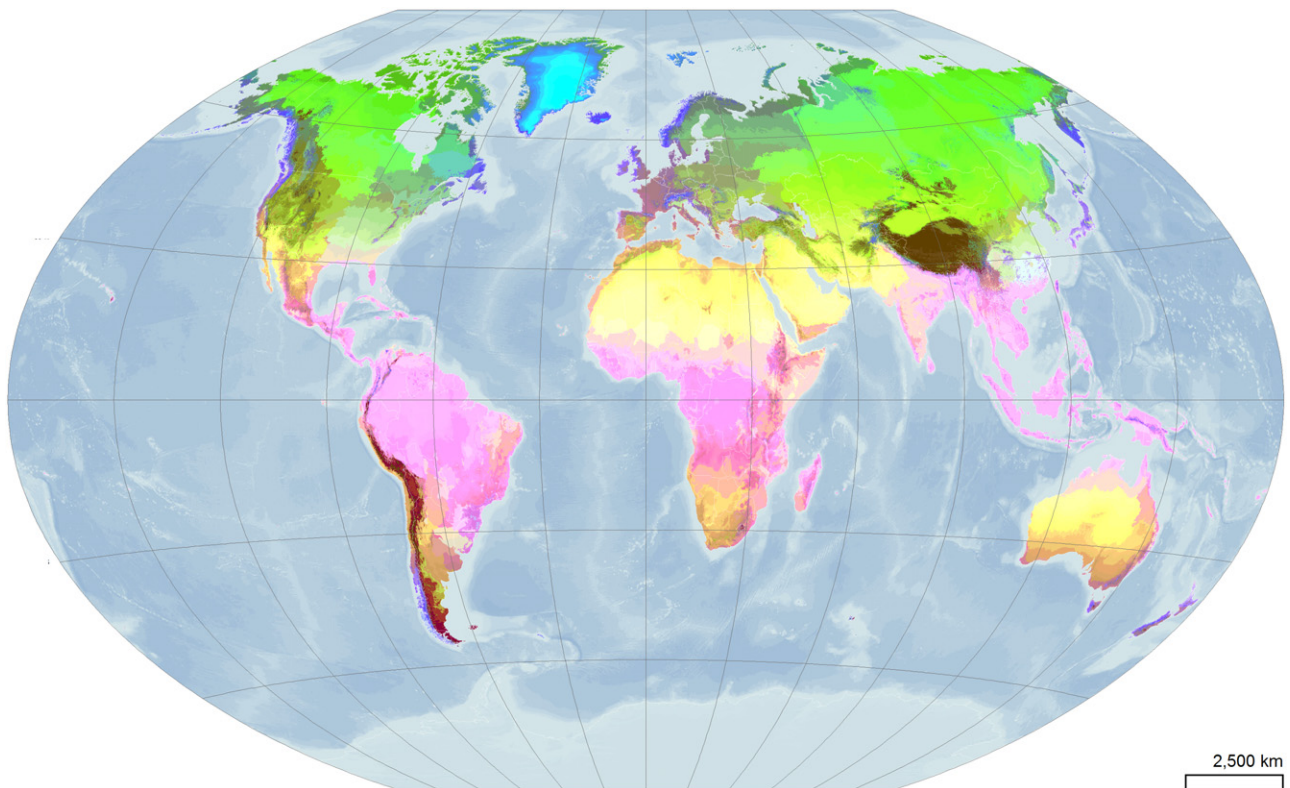
Figure 2 Map of the global environmental stratification, depicting 125 strata at a 30 arcsec (approximately 1 km^2 at the equator) spatial resolution in the Winkel Tripel projection. The legend provides a visual combination of the three main climate gradients incorporated in the clustering. Shades of green reflect highly seasonal continental climates ranging from bright green and petrol for cool boreal to warmer chartreuse coloured regions; yellow reflects hot and arid conditions; pink reflects hot and wet tropical conditions; and purple represents wet oceanic environments. A high-resolution map is available as Appendix S4.

included in most existing classifications. Thus the screening provides statistical rules for the selection of the condensed subset of variables.

A priori selection of 125 strata

Conceptually it would have been desirable to use statistical stopping rules (Gordon, 1996) to determine the optimal number of strata, reducing judgement by determining a mathematically optimal number of divisions in multivariate parameter space. Such cut-off rules balance high intra-cluster similarity and low inter-cluster similarity, but such internal criteria for the quality of a clustering do not necessarily translate into good effectiveness in an application (Manning *et al.*, 2008). Outcomes would therefore require further testing through application, as suggested above, or comparisons with external criteria of cluster quality. However, the latter would require some form of benchmark for comparison, which does not exist. Existing statistical environmental stratifications have therefore been developed for an arbitrary number of strata that was deemed practical whilst providing sufficient detail (Bunce *et al.*, 1996; Regato *et al.*, 1999; Leathwick *et al.*, 2003; Metzger *et al.*, 2005). Bunce *et al.* (1996) provide a further discussion of statistical stopping rules in environmental stratification and concluded that accepting an arbitrary number was appropriate.

The number of strata for the GEnS was therefore based on the objectives and envisaged application of the dataset, i.e. to provide greater detail than previous global approaches, and to provide a suitable framework for global biodiversity research and monitoring. Previous experience (see examples in Table 3), and consultation within the GEO BON community (http://www.earthobservations.org/geobon.shtml), indicated that between 100 and 150 strata would be an appropriate balance between increased detail and complexity. The choice for 125 strata was the middle point of the envisaged range, and initial tests indicated that additional divisions result mainly in increased altitudinal and latitudinal bands, which in our opinion did not justify added complexity. The chosen number of strata is comparable to the environmental stratification of Europe (Metzger *et al.*, 2005) and the isoclimates of the United States (Sayre *et al.*, 2009) (Table 3).

Classification method

The selected classification method was successfully used before to construct a European stratification (Metzger *et al.*, 2005), but other algorithms exist that could have been tested and compared through performance indices, e.g. the silhouette index (Rousseeuw, 1987). As discussed before, deciding on the best stratification is difficult without external quality criteria.

Table 3 Strength of agreement, expressed by the Kappa statistic, between the global environmental stratification (GenS) and nine other climate or ecosystem classifications. Monserud & Leemans (1992) give an indication of the strength of agreement for different ranges of Kappa, which are noted here.

Climate classification	Extent	No. of classes	No. of GEnS strata	Kappa	Strength of agreement
Köppen[1]	Global	30	125	0.57	Good
WWF biomes[2]	Global	14	125	0.65	Good
EnS[3]	Europe	84	67	0.64	Good
Ecoregions of North America[4]	North America	183	121	0.65	Good
USGS isoclimates[5]	USA	125	86	0.68	Good
USGS isoclimates[6]	South America	10	78	0.62	Good
USGS isoclimates[7]	Africa	156	87	0.72	Very good
ITE land classes[8]	Great Britain	41	13	0.54	Good
CLATERES[9]	Spain	218	40	0.62	Good

WWF, World Wildlife Fund; EnS, environmental stratification; USGS, United States Geological Survey; ITE, Institute of Terrestrial Ecology; CLATERES, Clasificación Biogeoclimática Territorial de España.
[1]Peel *et al.* (2007).
[2]Olson *et al.* (2001).
[3]Metzger *et al.* (2005).
[4]CEC (1997); http://www.epa.gov/wed/pages/ecoregions/na_eco.htm
[5]Sayre *et al.* (2009); http://rmgsc.cr.usgs.gov/ecosystems/dataviewer.shtml
[6]Sayre *et al.* (2008); http://rmgsc.cr.usgs.gov/ecosystems/dataviewer.shtml
[7]Sayre (2011); http://rmgsc.cr.usgs.gov/ecosystems/dataviewer.shtml
[8]Bunce *et al.* (1996).
[9]Regato *et al.* (1999).

However, statistical environmental stratifications have much in common, even when clustering algorithms differ (Bunce *et al.*, 2002; Ortega *et al.*, 2012). Differences occur mainly in the subdivision of long continuous gradients without clear boundaries. In such cases the precise location of the division will have relatively little influence on the efficiency of the stratification.

Climate data

There are limitations to the climate data used to construct the GEnS, which will affect its quality. Hijmans *et al.* (2005) discuss how the quality of the surfaces is spatially variable and depends on the local climate variability in an area, the quality and density of the observations and the degree to which a spline can be fitted through it. Locally important climate drivers, e.g. those caused by aspect in mountain areas or the formation of sea fog along coastal ranges, are also poorly represented. Despite these limitations, WorldClim provides sufficient spatial detail to distinguish and partition steep environmental gradients.

A further consideration is the stability of the strata under climate change. The GEnS is based on climate observations from 1950–2000, a period of comparatively stable climate. Under climate change, the current distribution of the strata will shift polewards and upwards. Climate change scenarios can be used to visualize projected change, which can provide insight into broad trends as well as regional impacts (Metzger *et al.*, 2008). However, for analytical purposes, such as the design of an observation network, a stratification should ideally be stable. Given current projected shift in biomes (Loarie *et al.*, 2009) this seems unlikely. Nevertheless, the distribution of the GEnS strata, which broadly follow attitudinal and latitudinal gradients, is likely to remain appropriate for the coming decades.

CONCLUSION

The GEnS provides a robust spatial analytical framework for the aggregation of local observations, identification of gaps in current monitoring efforts and systematic design of complementary and new monitoring and research. The dataset can support global environmental assessment, and has been identified as a focal geospatial data resource for tasks of the Group on Earth Observation Biodiversity Observation Network (GEO BON; Scholes *et al.*, 2012). The dataset is available for non-commercial use through the GEO portal: http://www.geoportal.org.

ACKNOWLEDGEMENTS

REFERENCES

Bunce, R.G.H., Barr, C.J., Clarke, R.T., Howard, D.C. & Lane, A.M.J. (1996) Land classification for strategic ecological survey. *Journal of Environmental Management*, **47**, 37–60.

Bunce, R.G.H., Carey, P.D., Elena-Roselló, R., Orr, J., Watkins, J.W. & Fuller, R. (2002) A comparison of different biogeographical classifications of Europe, Great Britain and Spain. *Journal of Environmental Management*, **65**, 121–134.

CEC (1997) *Ecological regions of North America, toward a common perspective*. Commission for Environmental Cooperation, Montréal, Canada.

Emberger, L. (1930) Sur une formule climatique applicable en géographie botanique. *Comptes Rendus Hebdomadaires des Séances de l'Academie des Sciences*, **181**, 389–391 (in French).

Gordon, A. (1996) Null models in cluster validation. *From data to knowledge* (ed. by W. Gaul and D. Pfeifer), pp. 32–44. Springer, New York.

de Gruijter, J.J., Brus, D.J., Bierkens, M.F.P. & Knotters, M. (2006) *Sampling for natural resource monitoring*. Springer, Berlin.

Hazeu, G.W., Metzger, M.J., Mucher, C.A., Perez-Soba, M., Renetzeder, C. & Andersen, E. (2011) European environmental stratifications and typologies: an overview. *Agriculture, Ecosystems and Environment*, **142**, 29–39.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Holdridge, L.R. (1947) Determination of world plant formations from simple climatic data. *Science*, **105**, 367–368.

von Humboldt, A. (1867) *Ideeen zu einem Geographie der Pflanzen nebst einem naturgemälde der Tropenländer*. F. G. Cotta, Tübingen.

Jensen, J.R. (1996) *Introductory digital image processing: a remote sensing perspective*. Prentice-Hall, Englewood Cliffs, NJ.

Jongman, R.H.G., Bunce, R.G.H., Metzger, M.J., Mücher, C.A., Howard, D.C. & Mateus, V.L. (2006) Objectives and applications of a statistical environmental stratification of Europe. *Landscape Ecology*, **21**, 409–419.

Klijn, F. & de Haes, H.A.U. (1994) A hierarchical approach to ecosystems and its implications for ecological land classification. *Landscape Ecology*, **9**, 89–104.

Köppen, W. (1900) Versuch einer Klassifikation der Klimat, Vorsuchsweize nach ihren Beziehungen zur Pflanzenwelt. *Geographische Zeitschrift*, **6**, 593–611, 657–679 (in German).

Leathwick, J.R., Overton, J.M. & McLeod, M. (2003) An environmental domain classification of New Zealand and its use as a tool for biodiversity management. *Conservation Biology*, **17**, 1612–1623.

Levin, N. & Shmida, A. (2007) Determining conservation hotspots across biogeographic regions using rainfall belts: Israel as a case study. *Israel Journal of Ecology and Evolution*, **53**, 33–58.

Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B. & Ackerly, D.D. (2009) The velocity of climate change. *Nature*, **462**, 1052–1055.

Lugo, A.E., Brown, S.L., Dodson, R., Smith, T.S. & Shugart, H.H. (1999) The Holdridge life zones of the conterminous United States in relation to ecosystem mapping. *Journal of Biogeography*, **26**, 1025–1038.

McMahon, G., Gregonis, S.M., Waltman, S.W., Omernik, J.M., Thorson, T.D., Freeouf, J.A., Rorick, A.H. & Keys, J.E. (2001) Developing a spatial framework of common ecological regions for the conterminous United States. *Environmental Management*, **28**, 293–316.

Manning, C.D., Prabhakar, R. & Schutze, H. (2008) *Introduction to information retrieval*. Cambridge University Press, Cambridge.

Metzger, M.J., Bunce, R.G.H., Jongman, R.H.G., Mücher, C.A. & Watkins, J.W. (2005) A climatic stratification of the environment of Europe. *Global Ecology and Biogeography*, **14**, 549–563.

Metzger, M.J., Bunce, R.G.H., Leemans, R. & Viner, D. (2008) Projected environmental shifts under climate change: European trends and regional impacts. *Environmental Conservation*, **35**, 64–75.

Monserud, R.A. & Leemans, R. (1992) Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling*, **62**, 275–293.

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. (2001) Terrestrial ecoregions of the world: a new map of life on earth. *BioScience*, **51**, 933–938.

Ortega, M., Metzger, M.J., Bunce, R.G.H., Wrbka, T., Allard, A., Jongman, R.H.G. & Elena-Roselló, R. (2012) The potential for integration of environmental data from regional stratifications into a European monitoring framework. *Journal of Environmental Planning and Management*, **55**, 39–57.

Paruelo, J.M., Lauenroth, W.K., Epstein, H.E., Burke, I.C., Aguiar, M.R. & Sala, O.E. (1995) Regional climatic similarities in the temperate zones of North and South America. *Journal of Biogeography*, **22**, 915–925.

Peel, M.C., Finlayson, B.L. & McMahon, T.A. (2007) Updated world map of the Koppen–Geiger climate classification. *Hydrology and Earth System Sciences*, **11**, 1633–1644.

Prentice, I.C., Cramer, W., Harrison, S.P., Leemans, R., Monserud, R.A. & Solomon, A.M. (1992) A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography*, **19**, 117–134.

Regato, P., Castejón, M., Tella, G., Giménez, S., Barrera, I. & Elena-Roselló, R. (1999) Cambios recientes en los paisajes de los sistemas forestales de España. *Investigación Agraria. Sistemas y Recursos Forestales*, **1**, 383–398 (in Spanish).

Rousseeuw, P. (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational Applied Mathematics*, **20**, 53–65.

Sanderson, M. (1999) The classification of climates from Pythagoras to Koeppen. *Bulletin of the American Meteorological Society*, **80**, 669–673.

Sayre, R. (2011) The GEOSS era: towards operational environmental monitoring. *34th International Symposium on Remote Sensing of Environment* (ed. by C. Yamashita-Gill), pp.

278–281. International Symposium for Remote Sensing of the Environment, Sydney, Australia.

Sayre, R., Bow, J., Josse, C., Sotomayor, L. & Touval, J. (2008) Terrestrial ecosystems of South America. *North America land cover summit-a special issue of the Association of American Geographers* (ed. by J.C. Campbell, K.B. Jones and J.H. Smith), pp. 131–152. Association of American Geographers, Washington, DC.

Sayre, R., Comer, P., Warner, H. & Cress, J. (2009) *A new map of standardized terrestrial ecosystems of the conterminous United States:* US Geological Survey Professional Paper 1768. US Geological Survey, Reston, VA.

Scholes, R.J., Mace, G.M., Turner, W., Geller, G.N., Jurgens, N., Larigauderie, A., Muchoney, D., Walther, B.A. & Mooney, H.A. (2008) ECOLOGY: toward a global biodiversity observing system. *Science*, **321**, 1044–1045.

Scholes, R.J., Walters, M., Turak, E., Saarenmaa, H., Heip, C.H.R., Tuama, É.Ó., Faith, D.P., Mooney, H.A., Ferrier, S., Jongman, R.H.G., Harrison, I.J., Yahara, T., Pereira, H.M., Larigauderie, A. & Geller, G. (2012) Building a global observing system for biodiversity. *Current Opinion in Environmental Sustainability*, **4**, 139–146.

Sitch, S., Smith, B., Prentice, I.C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J.O., Levis, S., Lucht, W., Sykes, M.T., Thonicke, K. & Venevsky, S. (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic vegetation model. *Global Change Biology*, **9**, 161–185.

Ståhl, G., Allard, A., Esseen, P.-A., Glimskär, A., Ringvall, A., Svensson, J., Sundquist, S., Christensen, P., Torell, Å., Högström, M., Lagerqvist, K., Marklund, L., Nilsson, B. & Inghe, O. (2011) National Inventory of Landscapes in Sweden (NILS) – scope, design, and experiences from establishing a multiscale biodiversity monitoring system. *Environmental Monitoring and Assessment*, **173**, 579–595.

Tappan, G.G., Sall, M., Wood, E.C. & Cushing, M. (2004) Ecoregions and land cover trends in Senegal. *Journal of Arid Environments*, **59**, 427–462.

Thornthwaite, C.W. (1943) Problems in the classification of climates. *Geographical Review*, **33**, 233–255.

Thornthwaite, C.W. (1948) An approach toward a rational classification of climate. *Geographical Review*, **38**, 55–94.

Thuiller, W., Lavorel, S., Araujo, M.B., Sykes, M.T. & Prentice, I.C. (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences USA*, **102**, 8245–8250.

Tou, J.T. & Conzalez, R.C. (1974) *Pattern recognition principles*. Addison-Wesley, Reading, MA.

Trabucco, A., Zomer, R.J., Bossio, D.A., van Straaten, O. & Verchot, L.V. (2008) Climate change mitigation through afforestation/reforestation: a global analysis of hydrologic impacts with four case studies. *Agriculture, Ecosystems and Environment*, **126**, 81–97.

Visser, H. & de Nijs, T. (2006) The map comparison kit. *Environmental Modelling and Software*, **21**, 346–358.

Zomer, R.J., Trabucco, A., Bossio, D.A. & Verchot, L.V. (2008) Climate change mitigation: a spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agriculture, Ecosystems and Environment*, **126**, 67–80.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Overview of the 42 variables considered in the screening.
Appendix S2 Pearson correlation matrix for the full set of 42 variables.
Appendix S3 Eigenvalues and eigenvectors for the remaining 36 variables.
Appendix S4 High-resolution map of the global environmental stratification.

BIOSKETCH

Marc Metzger is a lecturer in the School of GeoSciences, University of Edinburgh. His research interests include environmental stratification and biodiversity monitoring, global change foresight analysis and the development of assessment techniques for assessing global change impacts, across scales. He is a member of the Group on Earth Observation Biodiversity Observation Network (GEO BON) working group on terrestrial ecosystem change, which focuses on national, regional and continental monitoring of changes in the distribution, extent and conditions of ecosystems.

Editor: Martin Sykes