A unified typology for small European rivers

Jonathan F. Jupke & Ralf B. Schäfer

iES Landau,
Institute for Environmental Sciences
University Koblenz-Landau, Landau, Germany

INSTITUT FÜR UMWELTWISSENSCHAFTEN LANDAU

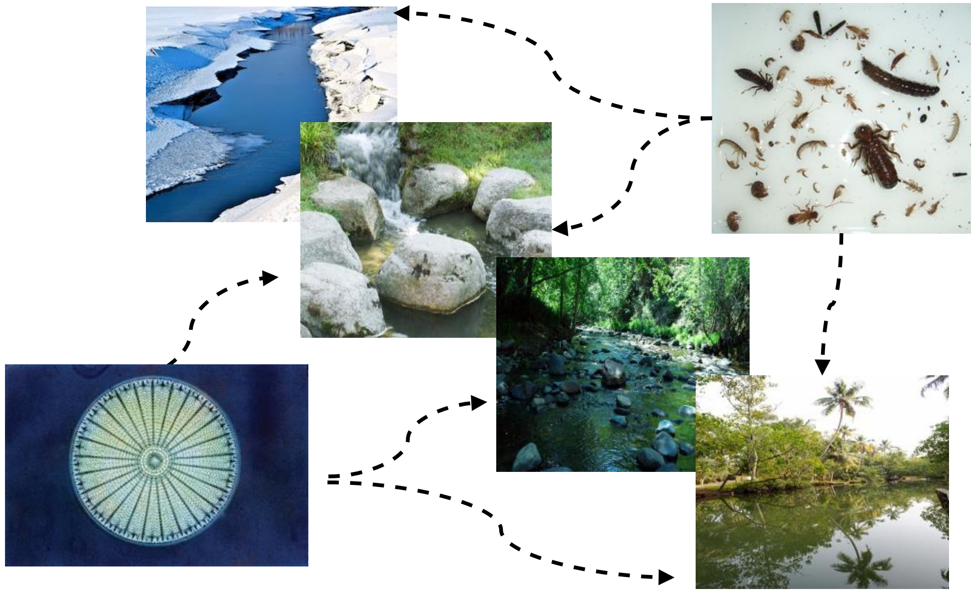
Adverse effects of Chemicals on Biodiversity

- In 2018, 84.1 Million tons of environmentally hazardous chemicals were produced in the EU (EUROSTAT, 2020)
- Their release into aquatic ecosystems can lead to adverse effects on biodiversity (Floehr *et al*. 2015; Schäfer *et al*. 2016; Vethaak *et al*. 2005)
- Ecological Risk Assessment (ERA) tries to prevent or limit such effects
- In Europe, ERA practices follow a one-size-fits-all approach – regulatory thresholds do not differ between recipient systems – But should they?

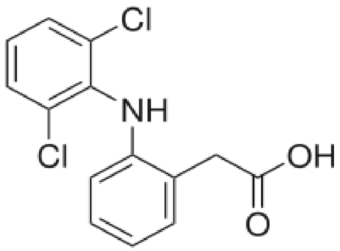
Spatial variation in sensitivity

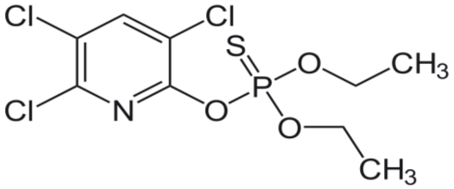
- How does chemical sensitivity vary across European streams?
- And how should this be considered in ERA?

Spatial variation



Variation in chemical sensitivity?





Get real!

To investigate these questions we started the GetReal project

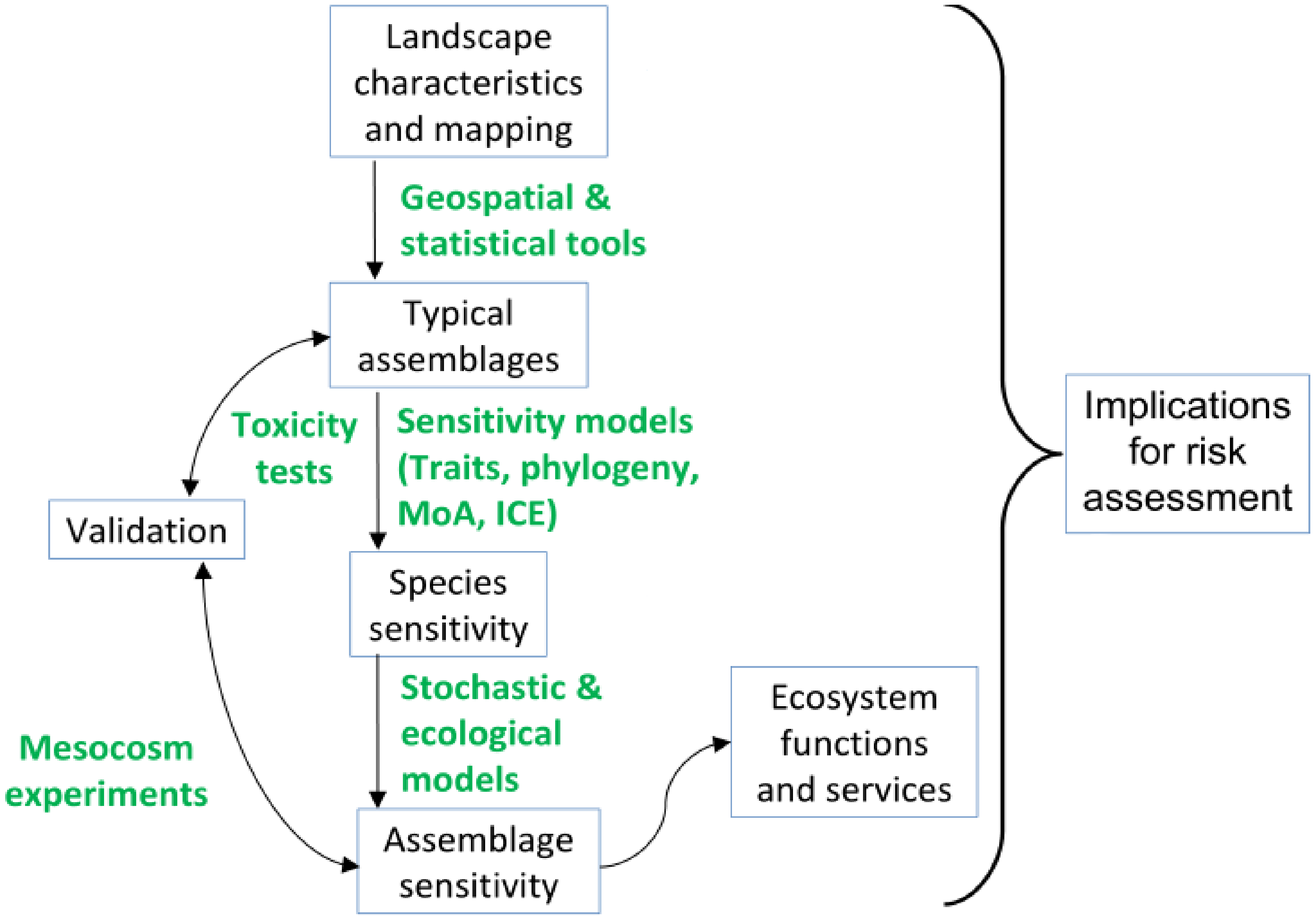


Figure 1: Research Program of GetReal

GetReal is a collaborative effort of: Ralf Schäfer, Martin Entling, Paul van den Brink, Lorraine Maltby, Frederik De Laender, Sebastian Scheu, Javier Jarillo, Sanne van den Berg, Tomás Duque and Jonathan Jupke

Get real!

To investigate these questions we started the GetReal project

In this presentation we will focus on the first step - a stream typology

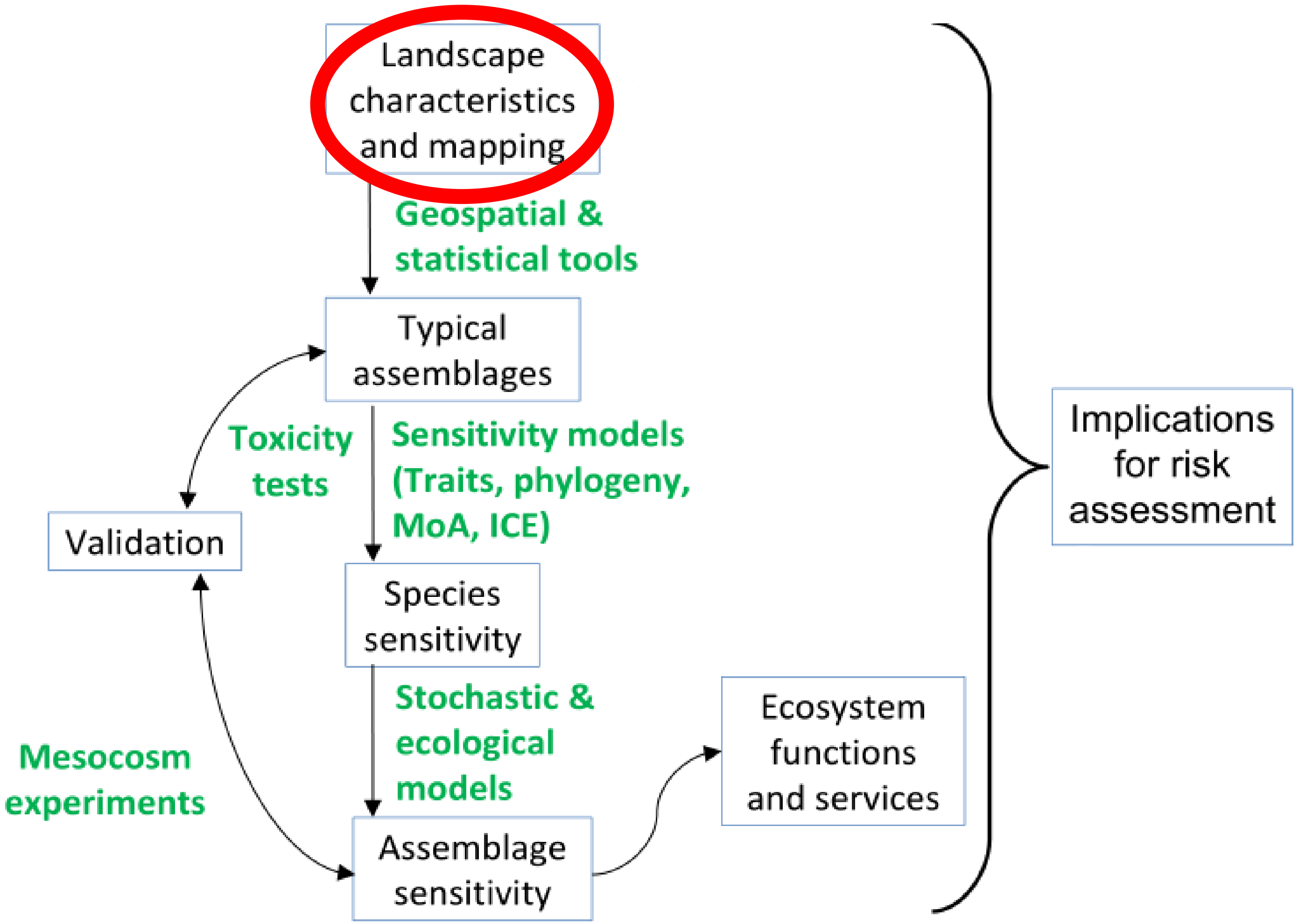


Figure 1: Research Program of GetReal

GetReal is a collaborative effort of: Ralf Schäfer, Martin Entling, Paul van den Brink, Lorraine Maltby, Frederik De Laender, Sebastian Scheu, Javier Jarillo, Sanne van den Berg, Tomás Duque and Jonathan Jupke

Rationale

- For this project, we selected eight countries, which represent macroecological gradients like climate and geology
- We further focused on small rivers which host a proportionally higher share of biodiversity and are at higher risk from chemical inputs (Lorenz *et al*. 2017; Link *et al*. 2017)
- A pan-European typology will be available later this year

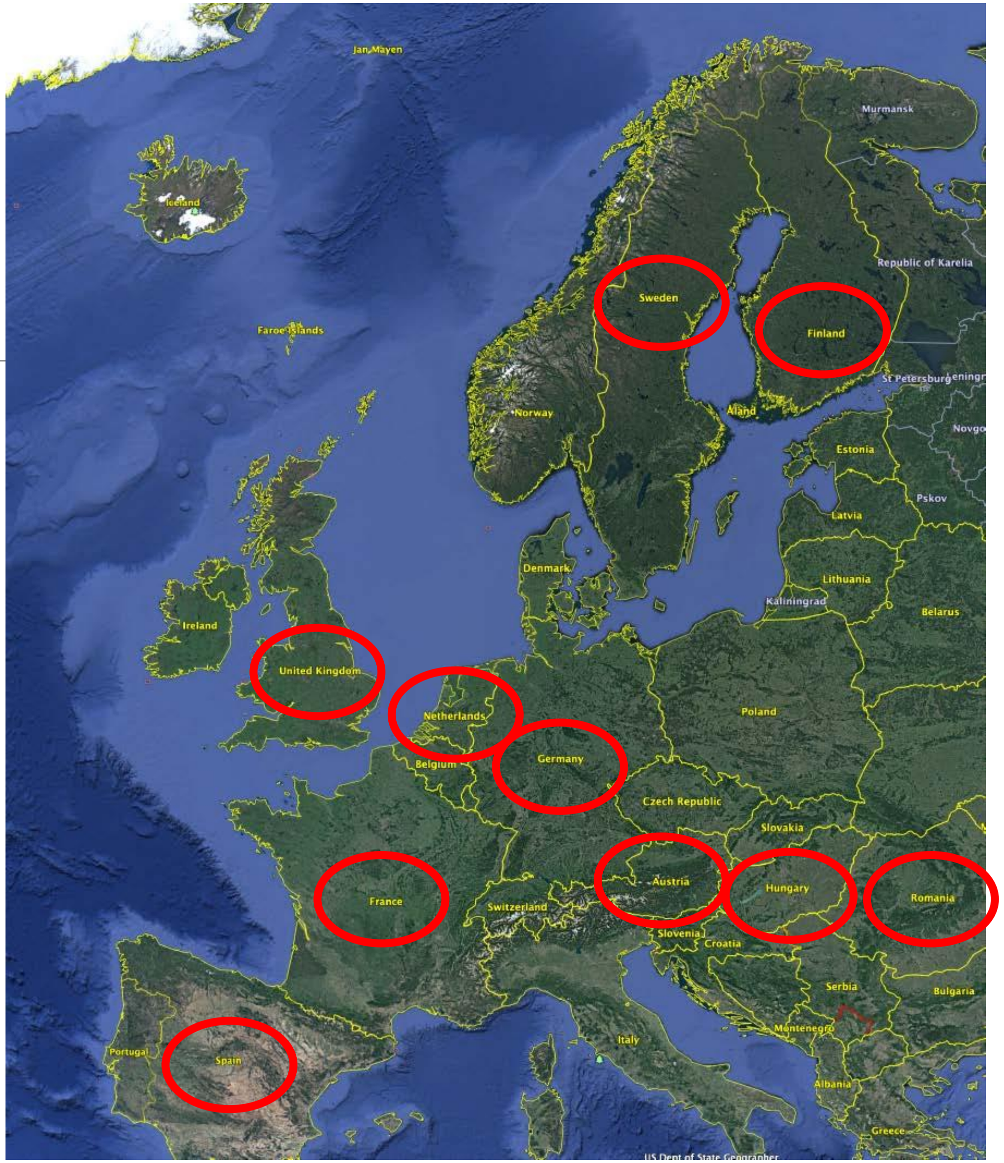


Figure 2: Map of selected Countries

Catchment characteristics and modelling

- The typology is based on CCM2 (Vogt *et al*. 2007)
- CCM2 is a Pan-European stream-network with associated catchments and environmental variables

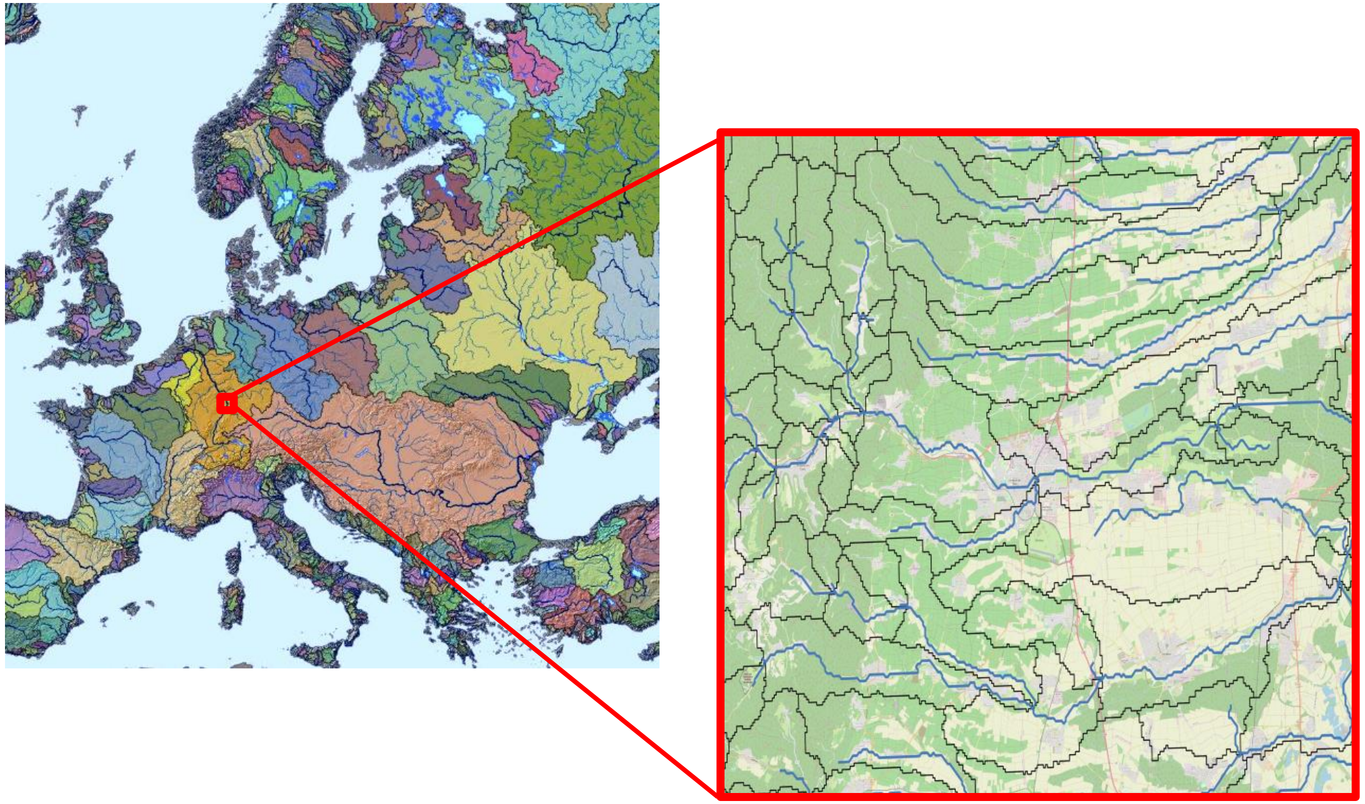


Figure 3: Extend of CCM2 with enlarged region showing streams (blue) and catchments (black)

Overview of variables

Abiotic Descriptor	Spatial Unit	Source
Catchment area	Catchments	CCM2.1 [7]
Upstream catchment area	Catchments	CCM2.1 [7]
Size class	River Segment	Derived from discharge
Sinuosity	River Segment	CCM2.1 [7]
Distance to Source	River Segments	CCM2.1 [7]
Elevation	Catchment	CCM2.1 [7]
Elevation heterogeneity	Catchment	CCM2.1 [7]
Maximal Elevation	Catchment	CCM2.1 [7]
Air temperature	Catchment	Worldclim [8]
Temperature range	Catchment	Worldclim [8]
Upstream catchment area	Catchment	CCM2.1 [7]
Share of Ice-covered Area	Catchment	Vegetation at last glacial maximum[9]
Share of steppe	Catchment	Vegetation at last glacial maximum [9]
Share of forest	Catchment	Vegetation at last glacial maximum [9]

Abiotic Descriptor	Spatial Unit	Source
Share of boreal vegetation	Catchment	Vegetation at last glacial maximum [9]
Precipitation	Catchment	CHELSA [10]
Precipitation Seasonality	Catchment	CHELSA [10]
Share of Calcareous/Silicate/Sediment	Catchment	IHME 1500 v1.1 [11]
Mean Discharge	River Segment	CHELSA [10]; GRDC [12]; NRFA [13]
Coefficient of Variation Discharge	River Segment	CHELSA [10]; GRDC [12]; NRFA [13]
Skewness discharge	River Segment	CHELSA [10]; GRDC [12]; NRFA [13]
Kurtosis discharge	River Segment	CHELSA [10]; GRDC [12]; NRFA [13]
the autoregressive lag-one correlation coefficient	River Segment	CHELSA [10]; GRDC [12]; NRFA [13]
Amplitude of seasonal signal	River Segment	CHELSA [10]; GRDC [12]; NRFA [13]
Phase of seasonal signal	River Segment	CHELSA [10]; GRDC [12]; NRFA [13]

Discharge Timeseries

- We modeled daily discharge for each stream segment from 01.01.2000 until 31.12.2013
- Following Irving *et al*. 2018, we used upstream precipitation as predictor variable
- Gauging data from Global Runoff data center and the National River Flow archive were used
- Resulting discharge time series were summarized as: mean, coefficient of variation, skewness, kurtosis, autoregressive lag-one correlation coefficient, amplitude, and phase of the seasonal signal (Archfield *et al*., 2014)
- All stream segments with mean discharge > 10 m^3/s were categorized as big and omitted from the subsequent clustering

Weighting Variables

- We use Generalized Dissimilarity Models (GDM) to weigh the variables prior to clustering (Ferrier *et al*., 2007)
- GDMs are GLMs in which a distance matrix between sites is modeled through the differences in I-spline smooth functions of the environmental variables. The splines are scaled by their explanatory power in the model. The smooth functions from the GDM can be applied to a data set with the same variables but additional sites
- Using a large database of aquatic invertebrate occurrences compiled for the second part of GetReal, we fit a GDM and transformed the variables for all stream segments

Clustering

- We dropped variables until no correlation was >= 0.7
- We computed tree-based dissimilarity (d4) between all stream segments (Buttrey &Whitaker, 2015) and clustered the resulting distance-matrix with CLARA (Kaufmann & Rousseeuw, 1986)
- CLARA is a non-hierarchical clustering approach, hence the number of clusters must be supplied by the researcher. We computed all clusterings with group sizes from 2 to 30 and selected the minimum group size that was within 5% silhouette with of the group with the absolute minimum silhouette width.
- For the unweighted variables this resulted in 15 stream types and for the weighted variables in 11.

Stream Typologies

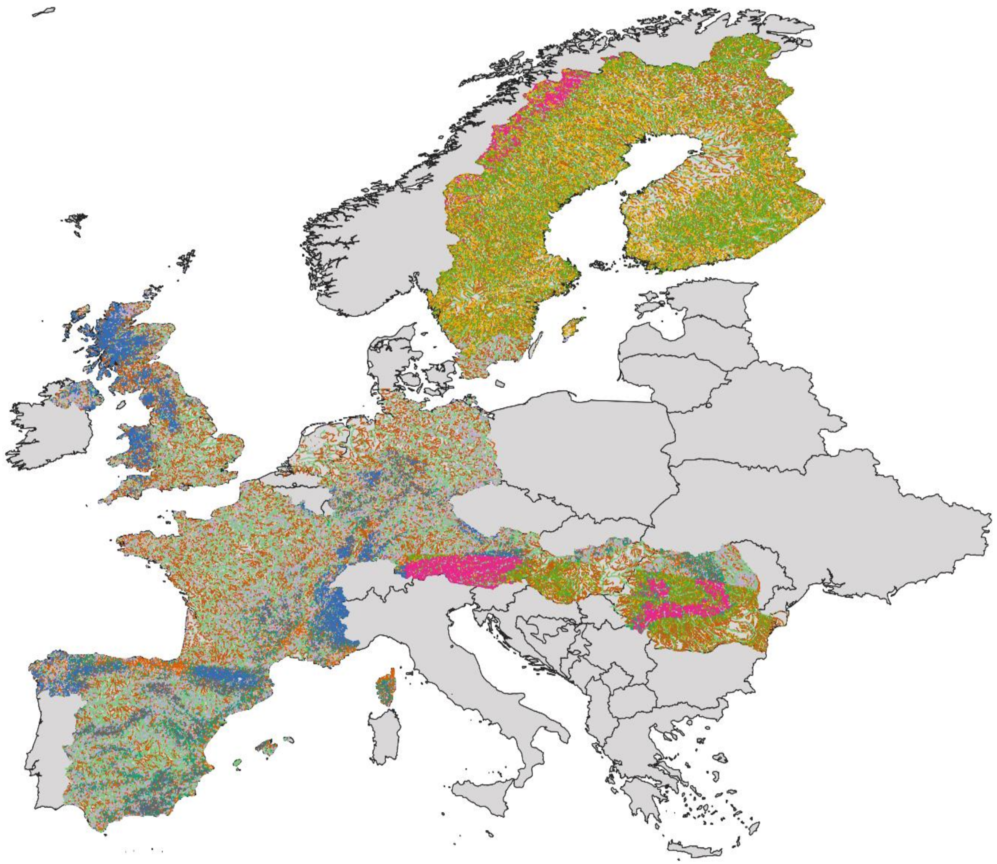


Figure 4: Unweighted stream typology

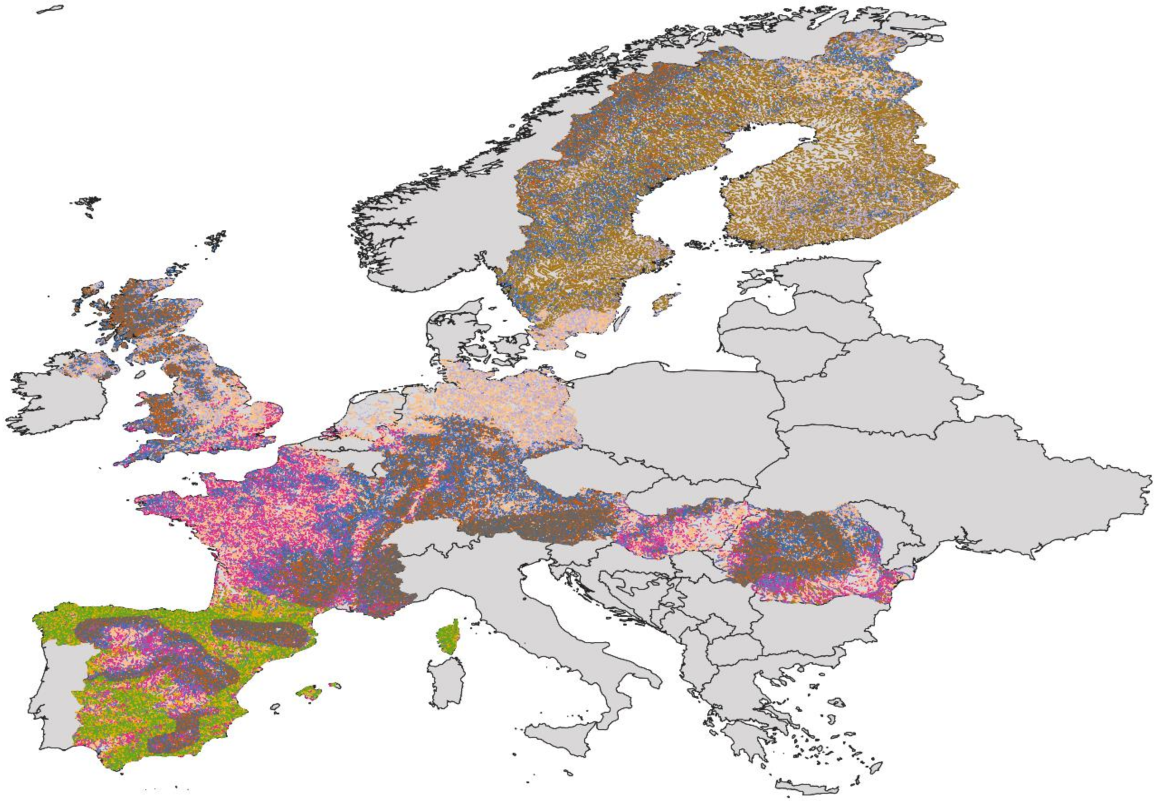
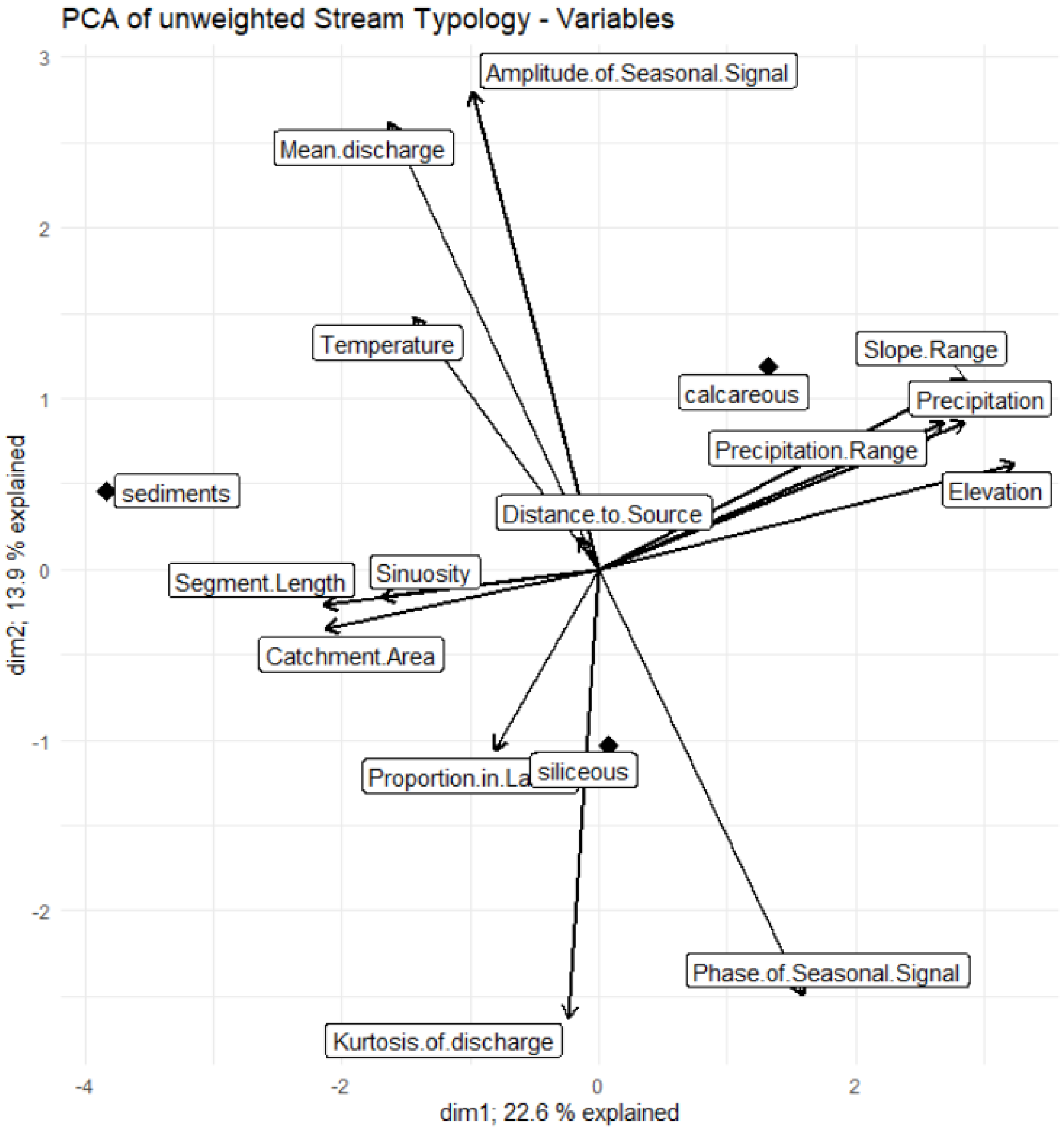
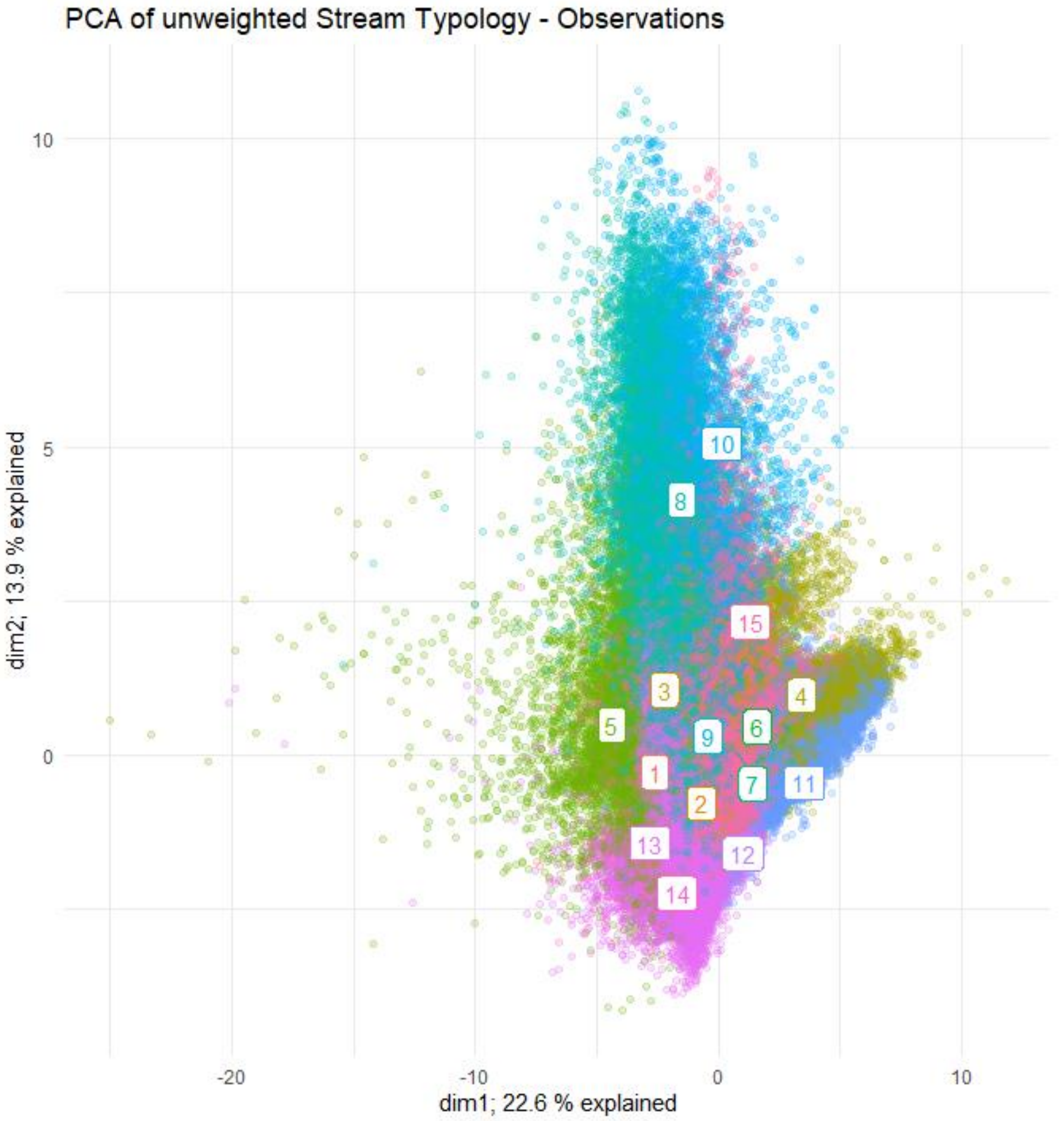


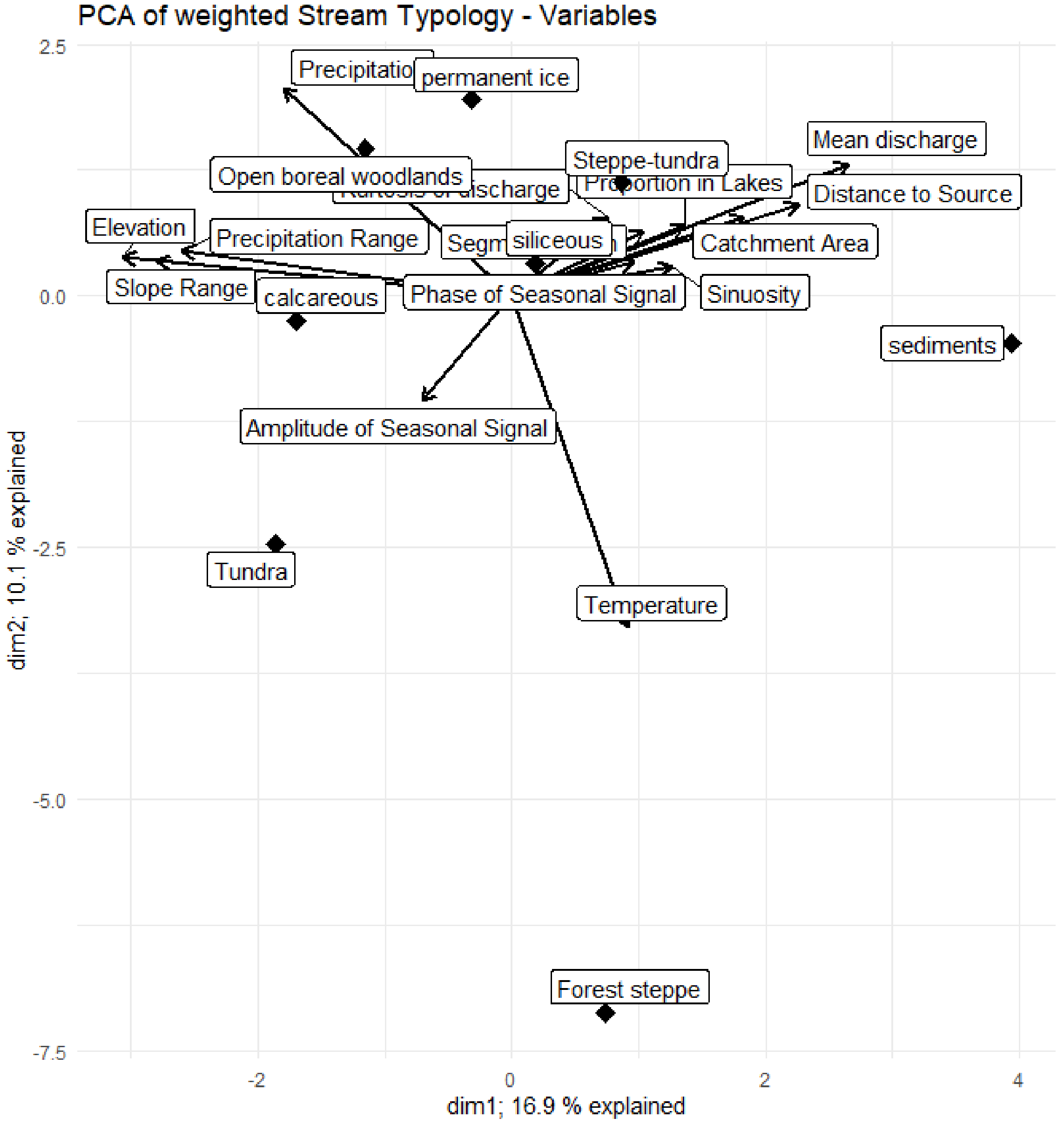
Figure 5: Weighted stream typology

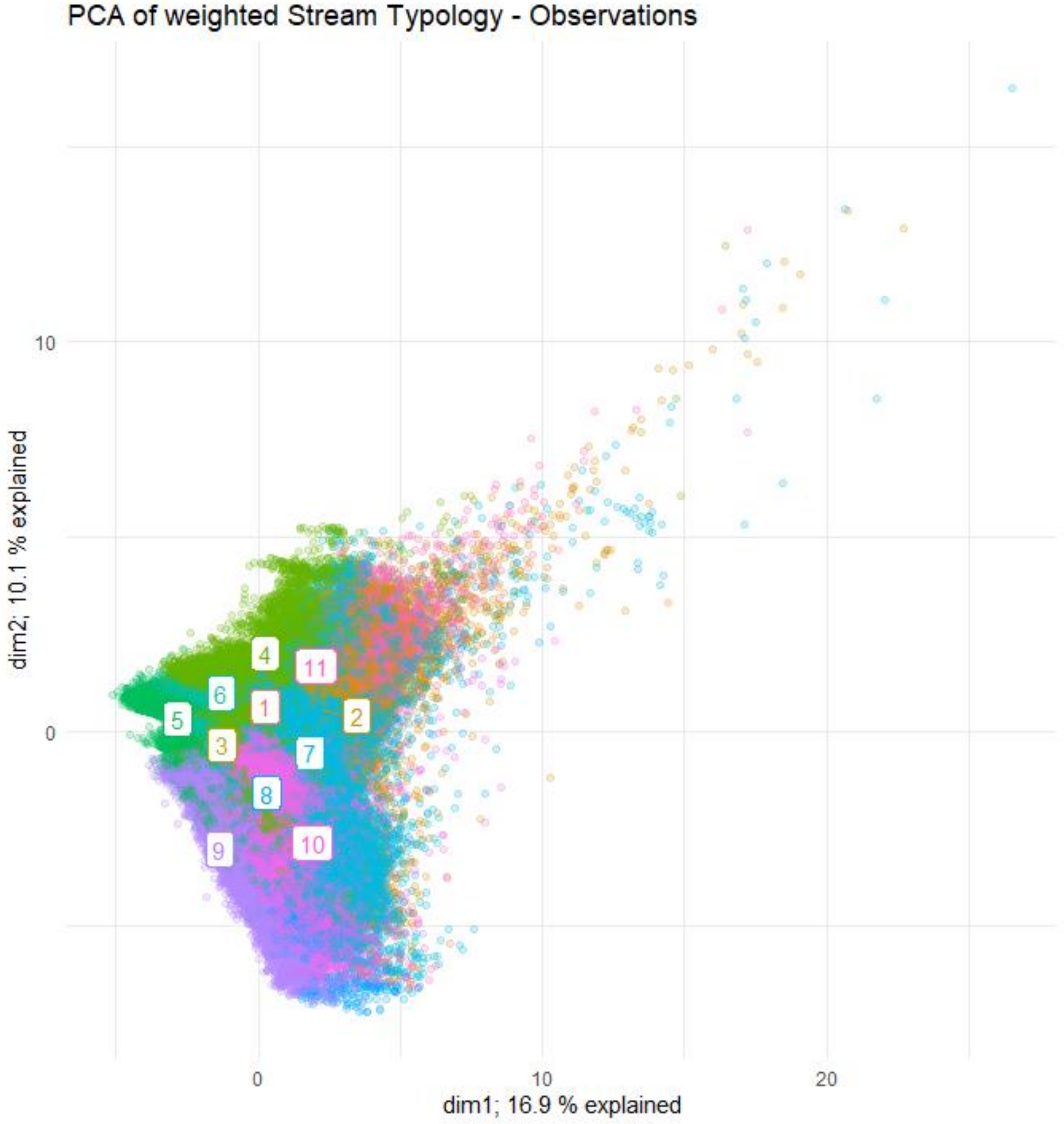
Ordination of input variables

- On the next two slides you will see biplots of first the unweighted and then the weighted typology
- To declutter these images we separated the variables (left) from the observations (right)
- Please note, that axis length differs between the variable and the observation plot and that the variable coordinates have been stretched by a factor of four to improve legibility
- The color of observations together with a label in the same color indicates stream type
- The ordination is a mixed ordination employing Principal Component Analysis for continuous variables and Multiple Correspondence Analysis for categorical variables. See PCAmix in Chavent *et al*. (2017) for details


PCA of unweighted Stream Typology - Variables
Amplitude.of.Seasonal.Signal
Mean.discharge
Temperature
calcareous
Slope.Range
Precipitation
Precipitation.Range
sediments
Elevation
Distance.to.Source
Segment.Length
Sinuosity
Catchment.Area
Proportion.in.La
siliceous
Phase.of.Seasonal.Signal
Kurtosis.of.discharge
dim2; 13.9 % explained
dim1; 22.6 % explained
3
2
1
0
-1
-2
-4
-2
0
2


PCA of unweighted Stream Typology - Observations
10
8
15
3
4
5
6
9
1
7
11
2
13
12
14
dim2; 13.9 % explained
dim1; 22.6 % explained
10
5
0
-20
-10
0
10


PCA of weighted Stream Typology - Variables
Precipitatio
permanent ice
Mean discharge
Steppe-tundra
Open boreal woodlands
Proportion in Lakes
Distance to Source
Elevation
Precipitation Range
siliceous
Catchment Area
Slope Range
calcareous
Phase of Seasonal Signal
Sinuosity
sediments
Amplitude of Seasonal Signal
Tundra
Temperature
Forest steppe
dim2; 10.1 % explained
dim1; 16.9 % explained
2.5
0.0
-2.5
-5.0
-7.5
-2
0
2
4


PCA of weighted Stream Typology - Observations
dim2; 10.1 % explained
dim1; 16.9 % explained
10
0
0
10
20
4
11
6
1
2
5
3
7
8
10
9

Potential use cases

- Besides our aim to predict typical assemblages our typology can be applied to:
  - develop stratification schemes for monitoring programs
  - analyze gaps in existing or planned conservation endeavors
  - Species Distribution Models as a predictor variable (McManamay *et al*. 2018)
  - guide restoration efforts by identifying reference conditions within each type

Thanks to

cefic

ecetoc

References

[1] EUROSTAT (2020) Production and consumption of chemicals by hazard class. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=env_chmhaz&lang=en

[2] Floehr T, Scholz-Starke B, Xiao H, Hercht H, Wu L, Hou J, Schmidt-Posthaus H, Segner H, Kammann U, Yuan X, Roß-Nickoll M, Schäffer A, Hollert H (2015) Linking Ah receptor mediated effects of sediments and impacts on fish to key pollutants in the Yangtze Three Gorges Reservoir, China — a comprehensive perspective. Sci Total Environ 538:191–211

[3] Schäfer RB, Kühn B, Malaj E, König A, Gergs R (2016) Contribution of organic toxicants to multiple stress in river ecosystems. Freshw Biol. https://doi.org/10.1111/fwb.12811

[4] Vethaak AD, Lahr J, Schrap SM, Belfroid AC, Rijs GBJ, Gerritsen A, de Boer J, Bulder AS, Grinwis GCM, Kuiper RV, Legler J, Murk TAJ, Peijnenburg W, Verhaar HJM, de Voogt P (2005) An integrated assessment of estrogenic contamination and biological effects in the aquatic environment of the Netherlands. Chemosphere 59:511–524

[5] Lorenz S., Rasmussen J.J., Süß A., Kalettka T., Golla B., Horney P., et al. (2017). Specifics and challenges of assessing exposure and effects of pesticides in small water bodies.Hydrobiologia 793, 213–224. https://doi.org/10.1007/s10750-016-2973-6

[6] Link M., Ohe P.C. von der, Voß K. & Schäfer R.B. (2017). Comparison of dilution factors for German wastewater treatment plant effluents in receiving streams to the fixed dilution factorfrom chemical risk assessment. Science of The Total Environment 598, 805–813.https://doi.org/10.1016/j.scitotenv.2017.04.180

References

[7] Vogt, J., Soille, P., De Jager, A., Rimavičiūtė, E., Mehl, W., Foisneau, S., … Bamps, C. (2007). A pan-European river and catchment database. In *JRC Reference Reports*. https://doi.org/10.2788/35907

[8] 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*(15), 1965–1978. https://doi.org/10.1002/joc.1276

[9] Ray, N., & Adams, J. (2001). A GIS-based Vegetation Map of the World at the Last Glacial Maximum (25,000-15,000 BP). *Internet Archaeology*, *11*(11), 0–15. https://doi.org/10.11141/ia.11.2

[10] Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., … Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, *4*, 1–20. https://doi.org/10.1038/sdata.2017.122

[11] Cornu, J.-F., Eme, D., & Malard, F. (2013). The distribution of groundwater habitats in Europe. *Hydrogeology Journal*, *21*(5), 949–960. https://doi.org/10.1007/s10040-013-0984-1

[12] Global Runoff Data Centre (2019): River Discharge Data/ *Global Runoff Data Centre. Koblenz, Federal Institute of Hydrology (BfG).*

[13] UK National River Flow Archive (2019). https://nrfa.ceh.ac.uk/. *Accessed via:* Vitolo, C., Fry, M. and Buytaert, W., (2016). rnrfa: an R package to retrieve, filter and visualize data from the UK National River Flow Archive. The R Journal, 8(2), pp.102-116.

References

[14] Irving, K., Kuemmerlen, M., Kiesel, J., Kakouei, K., Domisch, S., & Jähnig, S. C. (2018). A high-resolution streamflow and hydrological metrics dataset for ecological modeling using a regression model. *Scientific Data*, *5*, 1–14. https://doi.org/10.1038/sdata.2018.224

[15] Archfield, S. A., Kennen, J. G., Carlisle, D. M., & Wolock, D. M. (2014). An objective and parsimonious approach for classifying natural flow regimes at a continental scale. *River Research and Applications*, *30*(9), 1166–1183. https://doi.org/10.1002/rra

[16] Ferrier, S., Manion, G., Elith, J., & Richardson, K. (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions*, *13*(3), 252–264. https://doi.org/10.1111/j.1472-4642.2007.00341.x

[17] Buttrey, S. E., & Whitaker, L. R. (2015). treeClust: An R package for tree-based clustering dissimilarities. *R Journal*, *7*(2), 227–236.

[18] Kaufmann, L., & Rousseeuw, P. (1986). Clustering large data sets. In *Pattern Recognition in Practice* (pp. 425–437). Elsevier.

[19] Marie Chavent, Vanessa Kuentz, Amaury Labenne, Benoit Liquet and Jerome Saracco (2017). PCAmixdata: Multivariate Analysis of Mixed Data. R package version 3.1. https://CRAN.R-project.org/package=PCAmixdata

[20] McManamay, R. A., Troia, M. J., DeRolph, C. R., Sheldon, A. O., Barnett, A. R., Kao, S. C., & Anderson, M. G. (2018). A stream classification system to explore the physical habitat diversity and anthropogenic impacts in riverscapes of the eastern United States. In *PLoS ONE* (Vol. 13). https://doi.org/10.1371/journal.pone.0198439