GREENeR: An R Package to Estimate and Visualize Nutrients Pressures on Surface Waters

by Angel Udías, Bruna Grizzetti, Olga Vigiak, Alberto Aloe, Cesar Alfaro, and Javier Gomez

Abstract Nutrient pollution affects fresh and coastal waters around the globe. Planning mitigating actions requires tools to assess fluxes of nutrient emissions to waters and expected restoration impacts. Conceptual river basin models take advantage of data on nutrient emissions and concentrations at monitoring stations, providing a physical interpretation of monitored conditions, and enabling scenario analysis. The GREENeR package streamlines water quality model in a region of interest, considering nutrient pathways and the hydrological structure of the river network. The package merges data sources, analyzes local conditions, calibrate the model, and assesses yearly nutrient levels along the river network, determining contributions of load in freshwaters from diffuse and point sources. The package is enriched with functions to perform thorough parameter sensitivity analysis and for mapping nutrient sources and fluxes. The functionalities of the package are demonstrated using datasets from the Vistula river basin.

1 Introduction

Nitrogen and phosphorus are key nutrients that heavily impact aquatic ecosystems. Detecting primary sources of nutrient pollution and their downstream spread, alongside assessing achievable reductions through restoration policies, are crucial for effective natural resource management. They aid in identifying priority intervention areas and planning actions to restore the ecological balance of receiving waters.

Modelling tools can be useful to assess the impacts of future scenarios, policy measures, and climate changes at the regional and continental scale (Arheimer, Dahné, and Donnelly 2012; Bartosova et al. 2019; Beusen et al. 2022; Bouraoui et al. 2014; Ludwig et al. 2010; Seitzinger et al. 2005), and to check the coherence of different policy targets, for instance between water and agricultural policies. The assessment of policy scenarios requires a flexible and spatially detailed analysis to account for climatic, hydrological, and socio-economic gradients (Bruna Grizzetti et al. 2021).

Several types of models can be applied to predict the transport of nutrients in river basins Fu et al. (2019). Among them, statistical or conceptual models have the advantage of being readily applied in large watersheds. These model rely on calibration of few parameters for establishing links between emissions at sources and fate in the stream network. They take full advantage of nutrient emissions and concentrations data at monitoring stations, which are now accessible with increasing spatial and temporal resolution.

A classic example of river basin conceptual model is SPARROW (Smith, Schwarz, and Alexander 1997; Schwarz et al. 2006), which is widely applied in the U.S. to assess water quality over large regions. SPARROW has inspired the Geospatial Regression Equation for European Nutrient (GREEN) losses model (B. Grizzetti et al. 2005; Bruna Grizzetti, Bouraoui, and Aloe 2012; Bruna Grizzetti et al. 2021) which is adapted to European conditions. GREEN has been used for assessing the nutrient loads to the European seas (Bruna Grizzetti, Bouraoui, and Aloe 2012; Bruna Grizzetti et al. 2021), nitrogen retention in European freshwaters (Bruna Grizzetti, Bouraoui, and Aloe 2012; La Notte et al. 2017), and for policy scenario analysis (La Notte et al. 2017; Bouraoui et al. 2014; Leip et al. 2015; Malagó et al. 2019; Bruna Grizzetti et al. 2021). Despite their usefulness, this type of river basin models are seldom compiled as dedicated R packages, for example SPARROW has some scripts available to process information for and analysis of models results in the R environment.

The GREEN model application comprises several key steps, including data extraction and organization, data harmonization and integration, examination and validation of input data sets, model calibration and parameter selection, model run, and result visualization. All of these features are now integrated into an R package (GREENeR) that comprises functions to streamline the process, evaluate and visualize all the steps, thus strengthening the robustness of model application.

1.1 About water surface nutrients estimation with GREEN

GREEN (B. Grizzetti et al. 2005; Bruna Grizzetti, Bouraoui, and Aloe 2012; B. Grizzetti, Bouraoui, and De Marsily 2008; Bruna Grizzetti et al. 2021) is a conceptual model to assess total nitrogen TN and

total phosphorous TP from a region of interest (usually a river basin), accounting for both diffuse and point sources. The package allows the analysis of different scenarios of nutrient input in the region of interest, where "scenario" indicates a combination of annual time-series of inputs, such as nitrogen or phosphorus and the topological structure of the region. The model comprises several nutrient sources and pathways:

1. agricultural diffuse sources include nutrients from mineral fertilisers and manure application, nitrogen crop and soil fixation. These sources undergo retention in the land phase (basin retention that account for e.g. crop uptake and volatilization losses) before reaching the stream;
2. other diffuse emissions, such as from scattered dwellings (i.e., isolated houses and small agglomerations that are not connected to sewerage systems), and atmospheric nitrogen deposition (for nitrogen module) or background losses (for phosphorus module), are also reduced, e.g. due to soil processes, before reaching the stream network;
3. point sources consist of urban and industrial wastewater discharges that are discharged into surface waters directly.

Once in the river network, all nutrient loads are reduced by in-stream retention in rivers and lakes.

Basin retention of agriculture sources is a decay function proportional to the inverse of the total annual precipitation in the catchment. Conversely, river retention is a decay function proportional to the river length, considered as a proxy for water residence time. Finally, lake retention is simulated as a function of the lakes residence time and average depth.

The basin is divided into spatial subunits (called catchments), with a given area, a river reach, an inlet node, and an outlet node. The catchments are topologically connected from the headwaters to the outlet in a cascading sequence. The sequence of nutrient load accumulation through the stream network is defined by Shreve's order (Shreve 1966). Nutrient input from the different sources, basin and river retention are simulated in each catchment and routed through the river network. For each catchment i in the basin, the GREEN nutrient load L_i is estimated by the general equation:

$$L_{i,y} = (1 - Lret_i) \cdot (DSA_{i,y} \cdot (1 - Bret_{i,y}) + DSB_{i,y} + PS_{i,y} + U_{i,y}) \cdot (1 - Rret_i) \tag{1}$$

where $L_{i,y}$ is the nutrient load at the catchment outlet-node (ton/yr) in y year; and the other variables represent different sources and sinks of nutrients.

Sources of nutrients are:

- $DSA_{i,y}$. Annual nutrient diffuse sources on agricultural land in the catchment (ton/yr): mineral and manure fertilization, atmospheric nitrogen deposition, plant and soil fixation for TN; mineral and manure fertilization, and background losses for TP.

- $DSB_{i,y}$. Annual nutrient diffuse sources in the catchment related with scatter dwellings and atmospheric nitrogen deposition on non-agricultural land for TN ($DSB_{i,y} = 0.38 \cdot FF_{i,y} \cdot AtmN_{i,y} + sd_{coeff} \cdot SdN_{i,y}$, where $FF_{i,y}$ is the fraction of non-agricultural land cover in the catchment, and $AtmN_{i,y}$ is the annual atmospheric nitrogen deposition on the catchment (ton/yr)); nutrient diffuse sources in the catchment related with scatter dwellings and background losses on non-agricultural areas for TP ($DSB_{i,y} = sd_{coeff} \cdot SdP_{i,y}$).

- $PS_{i,y}$. Nutrient point sources in the catchment (ton/yr).

- $U_{i,y}$. Nutrient load from upstream catchments (ton/yr).

Sinks of nutrients are:

- $Lret_i$ denotes the lake retention (fraction) of the i catchment. $Lret$ is currently defined according to Kronvang et al. (2004), but limited to a 10% maximum reduction:

$$Lret = \max \left(0.1, 1 - \frac{1}{1 + (dref/z) \cdot RT} \right) \tag{2}$$

where z represents the average lake depth (m); RT is the hydraulic residence time (yr); and $dref$ denotes a nutrient-related coefficient ($dref = 7.3$ for TN and $dref = 26$ for TP).

- $Bret_{i,y}$ is the fraction of basin retention of the i catchment in y year:

$$Bret_{i,y} = 1 - \exp(-alpha_P \cdot NrmInvRain_{i,y}) \tag{3}$$

where $NrmInvRain_{i,y}$ is the inverse of annual precipitation (mm) of the i catchment in y year, normalized by its maximum (Frank and Todeschini 1994). To keep basin retention coefficients comparable

across regions, the minimum precipitation *minPrec* is set to 50 mm/y, and thus $NrmInvRain_{i,y}$ is defined as

$$NrmInvRain_{i,y} = \frac{1/\max(50, precipitation_{i,y})}{(1/minPrec)}$$

- $Rret_i$ is the fraction of river retention of the i catchment:

$$Rret_i = 1 - \exp(-alpha_L \cdot NrmLengthKm_i) \tag{4}$$

where $NrmLengthKm_i$ is the length (km) of the catchment reach, normalized by the maximum in the dataset:

$$NrmLengthKm_i = \frac{\text{catchment length } i \text{ reach, in km}}{\max(\text{Reach length in the region, in km})}$$

Since the maximum reach length depends on the region and its subdivision of network reaches, the calibrated coefficient $alpha_L$ cannot be compared across regions that adopt different discretization. Note that basin retention varies from year to year according to annual precipitation, whereas the fraction of river and lake retention for a given catchment is constant in time.

Equation (1) is applied sequentially from the most upstream nodes to the basin outlet. The model parameters are:

1. the basin retention coefficient ($alpha_P$), which together with annual precipitation regulates nutrient retention of diffuse agricultural sources (Equation (3));
2. the river retention coefficient ($alpha_L$: Equation (4)), which regulates the river retention per river km.
3. the fraction of domestic diffuse sources that reaches the stream network (sd_{coeff}).

2 About the GREENeR package

2.1 Package organization

This work presents a new efficient and enhanced implementation of the model GREEN developed as an R package named **GREENeR**. **GREENeR** is developed in the R statistical software and provides tools and methods for applying GREEN to an area of interest and assessing annual time series of nutrient loads in a river network and at the basin outlet, plus contributions of nutrient sources to these loads. Some of the key features of the package comprise: a) functions for the creation of scenarios from data sources; b) computational efficiency; c) parallel-capable ; d) an extended suite of fine tuning options to control model calibration; e) built-in parameter sensitivity analysis; and f) functions to perform customised post-process analyses.

The functions of **GREENeR** package are arranged in three groups: i) functions to perform graphical summaries of model inputs; ii) functions to perform model parameters calibration and sensitivity analysis; iii) functions to compute, analyze and visualize through graphs and maps the model outputs, i.e. total loads and contributions by source. The package supports parallel processing, which helps reduce the computational load in handling large basins.

The time-series can represent either historic (current or past) conditions that can be associated to observations for model calibration, or hypothetical conditions foreseen under theoretical changes (e.g. to forecast results of nutrient management plans, or under climatic change).

A scenario contains:

1. The geospatial geometry of the catchments of the region, which currently is defined according to the Catchment Characterisation and Modelling Database for European Rivers v2 (CCM2) (De Jager and Vogt 2007).
2. The annual time series (1990-2018) of nutrient inputs per catchment.
3. Additional catchment information (annual precipitation, annual forest fraction, lake retention, reach length).
4. Observed total load (TN or TP) from monitoring stations.

Two historical scenarios, one for TN and one for TP, of the Lay basin (France) are provided with the **GREENeR** package. The Lay basin has an area of 1971 km^2, and is divided into 189 CCM2

catchments. The mean catchment area is 10.4 km^2. Nutrient observations comprised 22 TN entries from six monitoring stations and 58 TP data entries from eight monitoring stations (data from WISE Waterbase (EEA 2021)).

Further, in this article we present examples drawn for the TN historic scenario of the transboundary Vistula (Wisla) river basin, one of the largest river basins of Europe. The Vistula scenario comprises 15465 CCM2 catchments that compose the 193.894 km^2 river basin, TN inputs for 1990-2018, as well as 1364 observed TN loads from 412 monitoring stations. The TN input dataset is part of the larger dataset created to assess nutrient concentration in the European freshwaters (Vigiak et al. 2023).

2.2 Key functions in the GREENeR package

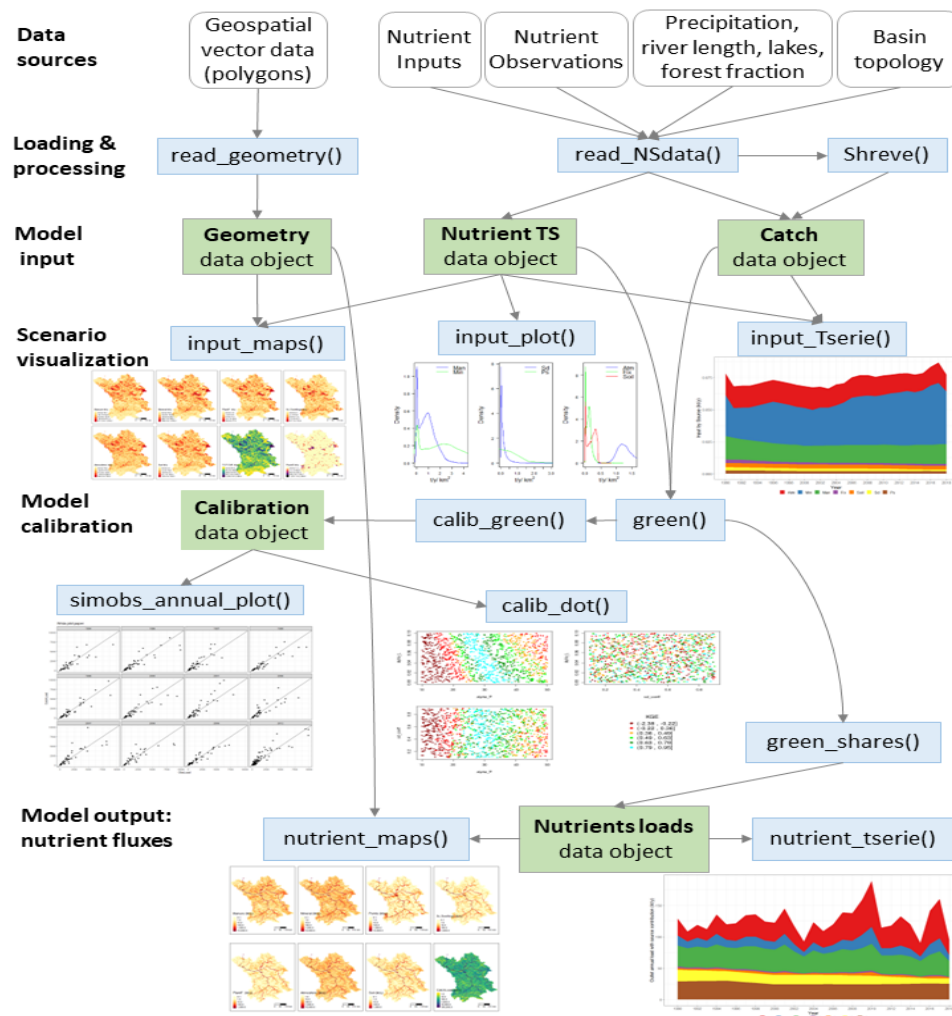A brief overview of the package and its main functions is given in Figure 1.

Figure 1: Schematic diagram of the procedure of GREENeR, including functions, data inputs and outputs, to estimate the nutrient loads. The green boxes represent data objects and the blue boxes represent the functions.

The key functions included in the package are:

- `read_geometry()`: imports the geospatial vector format (shapefile or ESRI shapefile) with the spatial information of the catchments in R.

- `read_NSdata()`: imports the annual time series of nutrient inputs per catchment and type of nutrient source (manure, mineral fertilization, point sources, scatter dwellings), plus the forest

fraction, and the observed annual nutrient loads from monitoring station data. The original data are stored in several comma-separated tables (CSV files).

- `input_maps()`: creates a map showing the mean nutrient load input by source.
- `input_plot()`: creates either a grouped barplot representing the average input load by source for the whole basin, or three density plots showing the distribution of nutrient sources.
- `input_Tserie()`: creates a time series plot showing basin inputs by source.
- `shreve()`: returns the Shreve order of the catchments, a useful indicator of stream size, discharge, and drainage area (Strahler 1957) calculated based on the sum of the orders of up of upstream tributaries (Shreve 1966). Commercial GIS software usually provides a Shreve calculation function, but in other software, it is harder to find. In GREEN, the Shreve's order defines the cascade of upstream-downstream catchments.
- `calib_green()`: conducts sensitivity analysis of model calibration utilizing a Latin Hypercube Sampling (LHS) scheme (Manache and Melching 2004) for three model parameters. It evaluates model performance by calculating several "goodness-of-fit" (GoF) metrics (Refsgaard and Henriksen 2004) against available observations during the specified simulation period (years).
- `select_params()`: extracts the best parameter set according to a selected GoF metric from the object generated by the calibration function, `calib_green()`.
- `calib_boxplot()`: creates a figure of boxplots that show the relationships between the best parameter sets determined by the specific GoF metric in each boxplot title, and six other commonly used hydrological metrics (Althoff and Rodrigues 2021; Mauricio Zambrano-Bigiarini 2014). In the lower panel, the figure also highlights the distribution of model parameters; the value of the best parameter set is marked as a red dot in each boxplot.
- `region_nut_balance()`: runs the GREEN model with the selected parameter set, and returns the mean annual mass balance of nutrient fluxes for the whole simulation period of the region. The results of this function can be visualized using a Sankey diagram via the `N4_sankey()` function.
- `green()`: runs the GREEN model with the selected parameter set and returns nutrient load time series for the simulation period. It generates less information than the `green_shares()` function, but its execution is faster, so it is used as a function for calibration iterations and is embedded in the `calib_green()` function.
- `green_shares()`: runs the GREEN model with the selected parameter set and returns time series of nutrient loads and the contributions of each nutrient source in the simulation period. The results of the model can be examined by `nutrient_tserie()` and `nutrient_maps()` functions.
- `scatter_plot()`: generates dot plots correlating parameters realizations in the calibration data with GoF metric, visualizing the impact of each parameter on model outcomes. The plot vary based on selected GoF metric from `green_calib()` function.
- `simobs_annual_plot()`: generates scatter plots comparing model load predictions (PredictLoad) with observations (ObsLoad) for each year of the stimulation period.

2.3 Input data requirements

GREENeR requires information on the nutrient inputs, the topology, and the geospatial geometry of the region of interest. The spatio-temporal input data must include all nutrient source fields, and differs in the two nutrient scenarios (TN or TP). In the case of TN (e.g. Figure 2), fields are (Equation (5) in Appendix 1):

- *Atmospheric*: Annual amount of atmospheric nitrogen deposition (ton/yr).
- *Mineral*: Annual amount of nitrogen from mineral fertilisers (ton/yr).
- *Manure*: Annual amount of nitrogen in manure fertilisers (ton/yr).
- *Fix*: Annual amount of nitrogen fixation by leguminous crops and fodder (ton/yr).
- *Soil*: Annual amount of nitrogen fixation in soils (ton/yr).
- *Sc.Dwellings*: Nitrogen input from scattered dwellings (ton/yr).
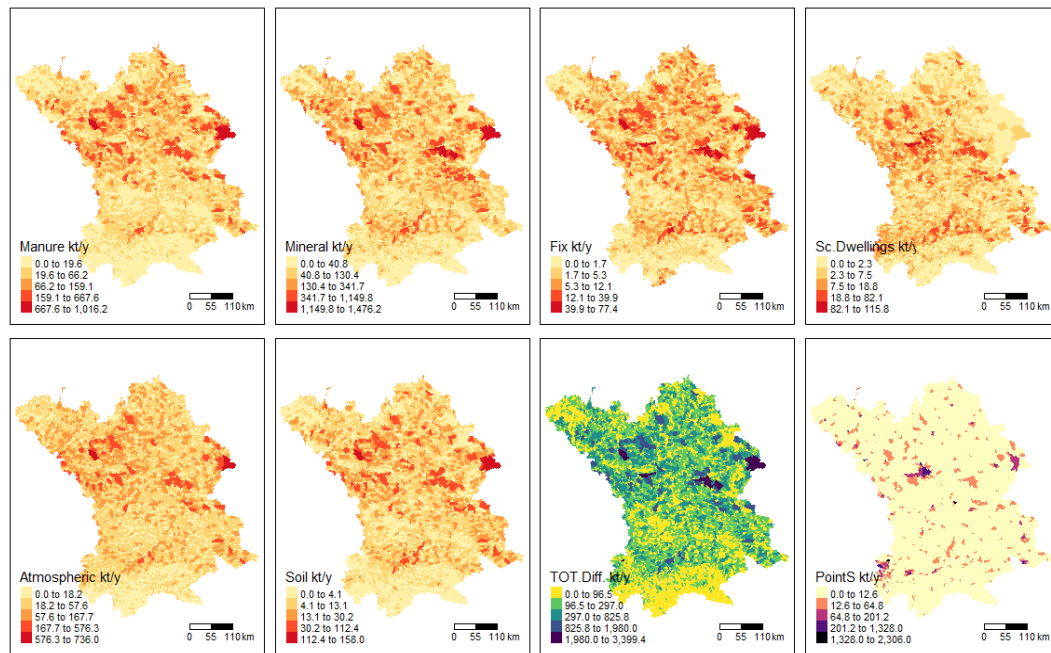- *PointS*: Nitrogen input from point sources (ton/yr).

Figure 2: Maps showing the mean annual TN inputs in the Vistula river basin in 1990-2018. The figure was generated with GREENeR input_maps() function. TOT.Diff = sum of diffuse inputs.

In the case of TP, fields are (Equation (6) in Appendix 1):

- *Background*: annual amount of phosphorus from background losses (ton/yr).
- *Mineral*: annual amount of phosphorus from mineral fertilisers (ton/yr).
- *Manure*: annual amount of phosphorus in manure fertilisers (ton/yr).
- *Sc.Dwellings*: phosphorus input from scattered dwellings (ton/yr).
- *PointS*: phosphorus input from point sources (ton/yr).

European datasets for 1990-2018 generated to assess historic and current nutrient fluxes (Vigiak et al. 2023) are available upon request. The model and the package are compatible with an external dataset and can be customized with local information as long as the data structure is respected.

In both nutrient cases, **GREENeR** needs additional annual catchment information:

- *ForestFraction*: annual fraction (0-1) of non-agricultural land cover area in the catchment (*FF* in Equations (5) and (6), in the Appendix 1).
- *Rain*: annual precipitation (mm, Equation (3)).
- *LakeFrRet*: lake retention fraction (0-1) (*Lret* in Equation (1), in the Appendix 1). Note, this differs for TN or TP scenarios (Equation (2)). At European scale, average lake depth and hydraulic residence time can be obtained from HydroLAKES database (https://www.hydrosheds.org/pages/hydrolakes, Messager et al. (2016)).
- Length: is the length (km) of the catchment reach (Equation (4)).

The catchment topology outlines the hydrological network configuration within the region. Each catchment must have a unique numerical identifier (HydroID)), to establish the network structure via a source HydroID and destination HydroID table. It is important to note that any outlet of the basin will be given as destination HydroID identifier "-1". Complementing the topology of the hydrologic network, the length of catchment reach should also be included (Equation (4)).

Finally, in order to calibrate the model, it is necessary to have some observed nutrient loads from monitoring stations associated to any catchment of the region. Ultimately, the quality, size, and spatial distribution of observed loads determine the robustness of the calibration process. In Europe, a large dataset of annual concentrations is publicly available (WISE Waterbase, EEA 2021), but annual flow must be derived from other sources (Bruna Grizzetti et al. 2021; Vigiak et al. 2023).

2.4 Performance and memory use

The **GREENeR** package performance and memory usage depends on the size of the region (number of catchments) and the number of years of the simulation. As a reference, the Danube basin (i.e. the largest European basin, with 138013 catchments) TN dataset for 1990-2018 requires approximately 389 Mb of memory to store the annual data. Table 1 shows the memory occupied by the scenarios for 6 European basins. In addition, it shows the computation times required in the execution of some key functions of the package. All executions were conducted on two computer configurations:

- CPU1: Desktop running window 10 with one Intel(R) Core i5-8259U CPU (4 cores, two threads per core) CPU at 2.30 GHz and 16 GB of RAM memory
- CPU2: Workstation running Linux with one AMD EPYC 7352 (24 cores, two threads per core) CPU at 2.3 GHz and 128 GB of RAM memory

The time required for the execution of the different functions increases linearly with the number of catchments (or Shreve order) of the region (Table 1). The calib_green() function is the most computationally demanding. However, the parallel implementation of this function drastically reduces the time required for this process. Calibration can be carried out in about 6 hours for practical implementation in large basins.

Table 1: Computational requirements to run some **GREENeR** functions in six European basins scenarios of 29 years. The columns Shreve order, Area and Number of catchments characterize the size of each basin. Mem is the amount of memory occupied by the GREEN model scenario generated by **GREENeR** package. The green(), green_shares() and calib_green() columns show the computation time required to execute the corresponding **GREENeR** functions for each scenarios under CPU1 and CPU2 computer configurations. All runs of calib_green() have been performed with 200 iterations. The reported times are the average of 5 runs of each process.

Basin name	Shreve order	Area (Km^2)	Number of catchments	Mem (MB)	green() (sec)		green_shares() (sec)		calib_green() (sec)	
					CPU1	CPU2	CPU1	CPU2	CPU1	CPU2
Lay	95	1971	189	0.5	6	4	15	12	316	31
Miño	2803	16985	5572	15.7	57	45	147	90	4110	394
Seine	2902	75989	5793	16.3	77	55	177	106	6320	452
Ebro	9351	85611	18568	52.3	160	123	425	224	10876	995
Vistula	7757	193894	15465	43.6	132	100	314	187	9004	814
Danube	69505	802032	138013	388.6	1120	766	2855	1108	94682	5638

3 Estimating nutrient loads using the GREENeR package

The entire procedure is summarized on Figure 1. Once the input data have been uploaded, the region scenario (Nutrient data and Catch data) is generated. The calib_green() function explores the parameter set ranges with LHC scheme, calculating GoF of parameter sets. Sensitivity analysis of its results is conducted to determine the best parameter set. Finally, the green_shares() function is used to estimate the nutrient loads and source apportionment (i.e. contribution to loads by source), per year and per catchment.

3.1 Scenario preparation

Assembling input data for running the GREEN model is time consuming. To facilitate the process, the read_NSdata() function assembles and organises annual information to generate a list of two objects: Nutrient Time Series Data Object and Catch Data Object of GREEN scenario. It needs four CSV files with specific data, namely:

- Time-series of annual nutrient inputs per catchment.
- Time-series of annual observed loads at monitoring stations.
- Basin topology and lake properties.
- Other: precipitation, forest fraction, reach length.

The spatial identifiers (HydroID) and temporal (year) units must be coherent in all the files. Besides organizing the input data, the read_NSdata() function also calculates the Shreve order of each catchment, normalises the precipitation (calculating $NrmInvRain_{i,y}$) and the reach length ($NrmLengthKm_i$).

```
csv_path <- "data/csv/"
scen <- read_NSdata(csv_path,"nutrients.csv","monitoring.csv", "forestFr.csv",
                    "precipitation.csv","topology.csv", "lakeProperties.csv",
                    "length.csv")
nutrient <- scen[[1]]
catch    <- scen[[2]]
```

The `shreve()` function can calculate the Shreve order independently based on the topology of the basin:

```
shreve_order <- shreve(catch)
```

Finally, the geospatial geometry of the catchments should be uploaded to enable the visualisation functionalities in map form. The `read_geometry()` function loads the geometry information file, i.e. a geospatial vector with the geometry of the catchment region:

```
geometry <- read_geometry(file = "data/shapes/Wisla.shp")
```

The geo-reference information, geometry and attributes of the spatial entities must be in a shapefile format (.shp), editable with ArcGIS or similar software. The identifiers of each catchment (HydroID) must be consistent with those in the CSV files of the scenario dataset.

Once the scenario has been generated, the library includes functions to examine the nutrient sources in a basin, and explore their distribution over time or in space. The `input_plot()` function provides annual average nutrient loads for the whole period and density plots of nutrient loads. The `input_Tserie()` function allows to examine the temporal evolution of the inputs, whereas the `input_maps()` function generates maps of nutrient inputs distribution in the basin (Figure 2).

```
input_maps(nutient, catch, geometry, "Wisla", "gr1")
```

3.2 Calibration procedure and sensitivity analysis

The calibration process is essential to find a suitable set of parameters for the model. In **GREENeR** package, the core function `calib_green()` enables to run a parameter sensitivity analysis and concurrently assessing the model performance according to several GoF metrics.

Sensitivity Analysis (SA) investigates how the variation in the output of a numerical model can be attributed to variations of its input factors. SA aims at identifying the most influential inputs or parameters, and quantifying how much they contribute to the variability/uncertainty of the model outputs. SA provides information on how much of the output variance is controlled by each parameter, or combination of parameters. SA is increasingly being used in environmental modelling for a variety of purposes, including uncertainty assessment, model calibration and diagnostic evaluation, dominant control analysis, and robust decision-making (Pianosi et al. 2016; Saltelli, Annoni, et al. 2011; Butler et al. 2014; Norton 2015; Mannina et al. 2018). This is achieved by running the model for many samples of the parameter space to determine their impact on the model outputs. SA allows identification of the parameters and input variables that strongly influence the model response (model output). Conversely, it may be of interest to the modeller to verify that although some model parameters may not be very well established they do not significantly contribute to output uncertainty. Saltelli, Tarantola, and Campolongo (2000) single out three main classes of SA methods: screening, local, and global methods. Local sensitivity analysis methods focus on assessing the impact of small variations in the input values of a model on the output results. Global sensitivity methods seek to understand how variations in the full range of input values affect the results. Puy et al. (2022) provide a comprehensive overview of several R packages for performing global sensitivity analysis.

Screening methods are economical and qualitative methods. Only screening methods are included in this implementation of the library as they provide a quick assessment of the relative importance of variables. This is particularly useful in exploratory studies and in cases where an initial subset of relevant variables needs to be identified.

Adequate sampling of parameter space is very important in model calibration. Several studies of uncertainty analysis in water resources problems (Melching and Bauwens 2001) concluded that Monte Carlo simulation and Latin Hypercube Sampling (LHS) (McKay, Beckman, and Conover 1979; Carnell 2012) methods are very powerful, robust, and flexible. Other approaches are possible (Mauricio Zambrano-Bigiarini 2014), but the LHS has the advantage that it is easily parallelisable, it explores the full range of parameter sets, it does not direct the search depending on the values of previous iterations like other optimisation methods (gradients, genetic algorithms, etc.), and is therefore independent of

the GoF metric. This allows performing a posteriori sensitivity analysis for several metrics without having to repeat the process.

The `calib_green()` function computes 15 GoF metrics (Althoff and Rodrigues 2021; Mauricio Zambrano-Bigiarini 2014) (listed in Appendix 2), and returns an object with all parameter sets generated by the LHS and the corresponding GoF scores. Running `calib_green()` requires the following settings:

1. The expected ranges for each parameter. The ranges are defined by two vectors of three values, one for the lower limits and one for the upper limits of the three parameters. The values correspond to each model parameter in sequence: $alpha_P$ (Equation (3)), $alpha_L$ (Equation (4)), and sd_{coeff} (Equations (5) and (6)).

2. The number of iterations to be performed during the calibration process. The higher the number of iterations, the more likely it is to achieve good results, but the longer the computation time. Depending on the basin and parameter range width, it is recommended to run at least 200 iterations to have enough information to continue the calibration process.

3. The years to be included in the calibration process.

```
n_iter <- 2000
low <- c(10, 0.000, 0.1)
upp <- c(50, 0.1, 0.9)
years <- c(1990:2018)

calibration <- calib_green(nutrient, catch, n_iter, low, upp, years)
```

The calibration process is automatically parallelised by the package and uses all computer cores except one. The computation time depends on the computer, the number of catchments in the basin, and the number of iterations (Table 1). The `calib_green()` function returns a data frame with a row for each iteration with parameter values and the resulting 15 GoF metric scores. An example of four parameter sets is shown below (cut to the first five GoF metrics).

```
#>     alpha_P    alpha_L   sd_coeff     NSE    mNSE    rNSE     KGE PBIAS %
#> s1 39.31203 0.07628558 0.6325056  0.5623  0.6452  0.9595  0.2512   -44.1
#> s2 35.06049 0.05510985 0.1577470  0.6733  0.6948  0.9563  0.3763   -36.3
#> s3 28.34947 0.04418381 0.7209566  0.8690  0.7737  0.8998  0.7926    -1.7
#> s4 40.73533 0.01351723 0.6195734  0.8370  0.7734  0.9607  0.6292   -22.6
```

The selection of the appropriate GoF metric(s) for calibration and evaluation of hydrological models can be challenging even for very experienced hydrologists. Choosing the right GoF metric for model calibration largely depends on the overall study scope, which defines the main interest (e.g. high or low load, upper or lower catchment area), and on the available observation dataset (size and quality) (Kumarasamy and Belmont 2018). Automated calibration typically relies, often exclusively, on a single GoF metric, with the Nash-Suttclife efficiency (NSE) been the most frequently used metric in hydrological models (Hoshin V. Gupta et al. 2009; Westerberg et al. 2011; Wöhling, Samaniego, and Kumar 2013). As described in Singh and Frevert (2002), a single criterion is inherently predisposed to bias towards certain components of the data time series. Automated procedures may be improved by using more than one criteria as discussed by Hoshin Vijai Gupta, Sorooshian, and Yapo (1998). Although automation can help the calibration process become more objective, efficient and practical, it is not a substitute for expert hydrologic intuition and understanding. Whether automated or manual calibration is used, a common approach is to adjust the parameters that display the highest sensitivity (Madsen 2003; Doherty and Skahill 2006).

GREENeR includes several functions to assist in selecting the best parameter set. The `calib_boxplot()` function shows (Figure 3) relationships between best parameter sets chosen according to one GoF parameter (title of each boxplot) in relation to six most frequently used metrics. Additionally, in the lower panel, the figure shows the distribution of model parameters in the most performing subset of parameter sets. The best parameter set according to each GoF metrics is marked as a red dot in each boxplot (Figure 3).

```
calib_boxplot(calibration, rate_bs = 5)
```
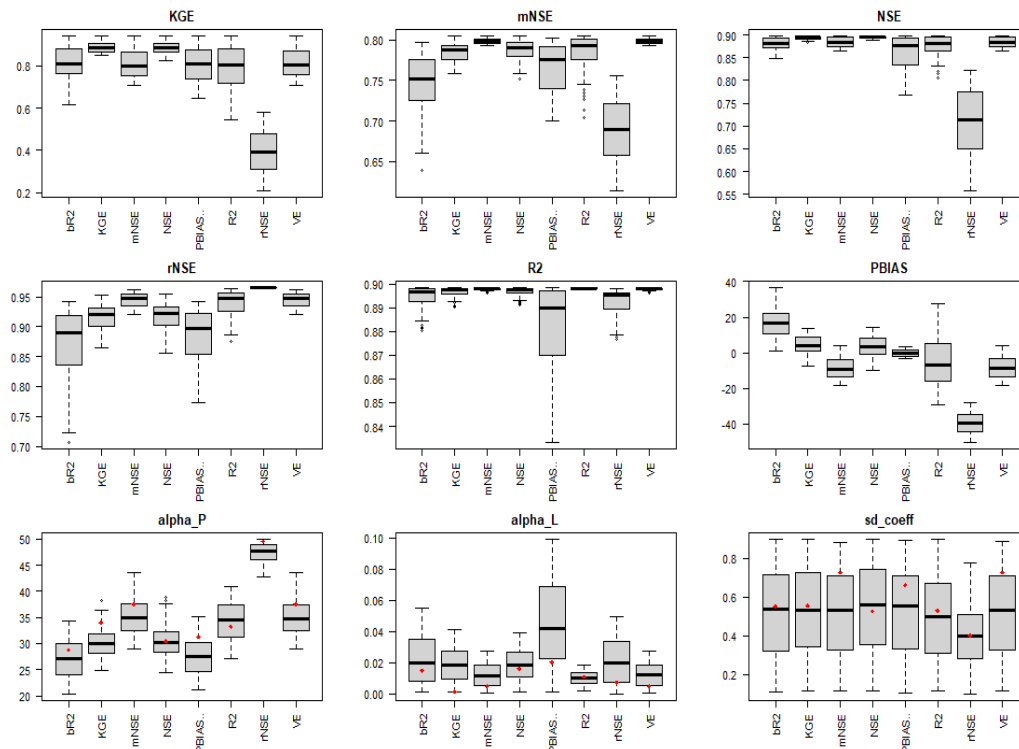
Figure 3: Calibration results for the Vistula river basin (TN Scenario). The figure displays the output of the calib_boxplot() function, and compare several GoF metrics (bR2, KGE, mNSE, NSE, PBIAS, R2, rNSE, VE) with each other and with the model parameters distributions. The top two rows show boxplots of each metric (specified in the title) with others (in the x labels). The lowest row shows the parameter value distributions for the top 5% (or another threshold indicated in rate_bs parameter of the function) parameter sets ranked according to the GoF metrics in the x label. The red dots represent the best parameter values for each boxplot.

The select_params() function extracts the parameter set that scored the best GoF metric of choice from the calibration result object.

```
best_params <- select_params(calibration, param = "NSE")

alpha_P  <- best_params$alpha_P
alpha_L  <- best_params$alpha_L
sd_coeff <- best_params$sd_coeff
```

It is not recommended to use the parameters extracted by the select_params() function without having performed an analysis of all the calibration results. Instead, alternative parameter sets (according to different GoF) should be compared before making the final selection.

Screening and SA of model parameters can be done via the scatter_plot() and calib_dot() functions of **GREENeR** package. The scatter_plot() function shows all parameter realizations in the LHS dataset against a selected GoF metric to visualize the influence of each parameter on the metric scores. The result depends on the GoF metric of choice (any metric calculated with calib_green() can be selected) (Figure 4).
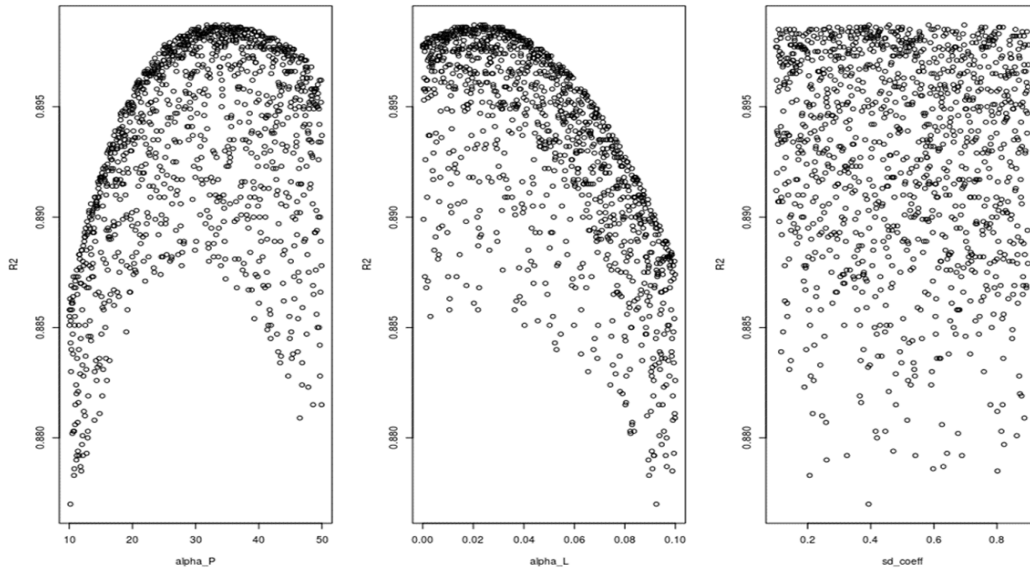
```
scatter_plot(calibration, param = "R2")
```

Figure 4: Scatter plots of model parameters against R2 metric generated by the scatter_plot() function for the Vistula river basin (TN).

Figure 4 example shows that the highest R2 are achieved for values of the parameter $alpha_P$ around 33.3 and for the parameter $alpha_L$ around 0.025, whereas the model is insensitive to the sd_{coef} parameter. Further, scatter_plot() function helps check if the parameter ranges defined in the search were suitable, or if the number of iterations was sufficient. In Figure 4, the best values of the parameters $alpha_P$ and $alpha_L$ are in the central part of the plots, thus the search parameter ranges were appropriate, and the distribution of dots is sufficient to visualize the effect of each parameter, indicating that the number of iterations was adequate.

The calib_dot() function shows the distribution of parameters in relation to each other for a chosen GoF metric, and highlights potential correlations between parameters (Figure 5).
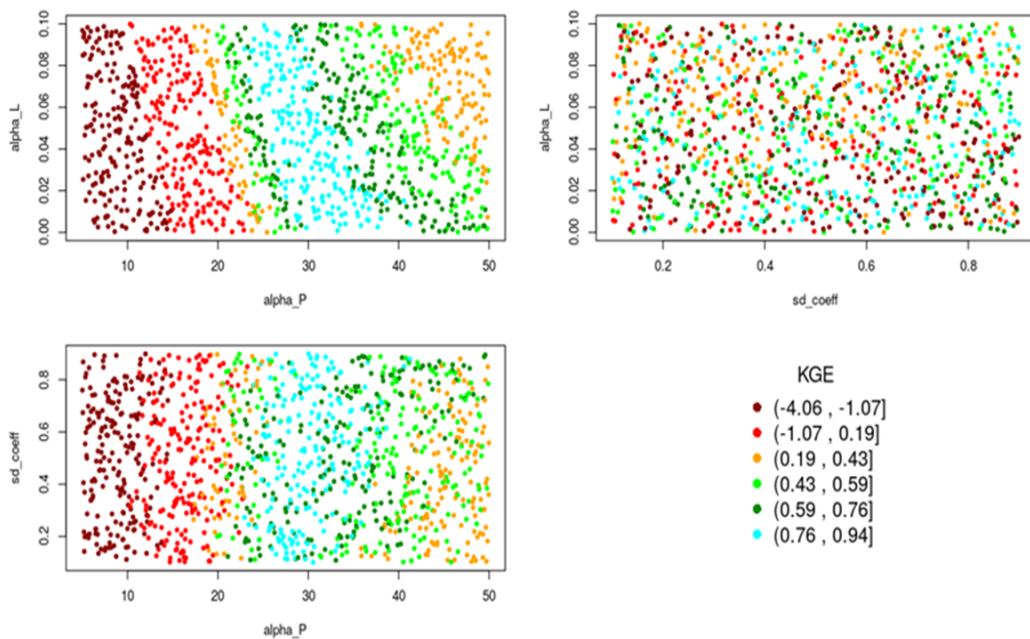
```
calib_dot(calibration, param = "KGE")
```


Figure 5: Dot plots of parameter pairs for the Kling-Gupta efficiency (KGE) generated by the calib_dot() function for the Vistula river basin (TN).

In Figure 5, the Kling-Gupta Efficiency metric (KGE, Althoff and Rodrigues 2021) shows interactions between $alpha_P$ and $alpha_L$, i.e. a higher $alpha_P$ should be associated to a lower $alpha_L$ for achieving similar KGE. The best $alpha_P$ values identified for both R2 and KGE metrics are similar, whereas $alpha_L$ values vary considerably (see also Figure 3).

The **GREENeR** package includes two useful model calibration functions: `compare_calib()` and `simobs_annual_plot()`. `compare_calib()` shows a scatter plot that compares observed data points with corresponding values predicted by two parameter sets. Figure 6 displays a plot generated by the `compare_calib()` for the Vistula (TN) basin scenario when using the best parameter sets based on NSE and rNSE metrics. The plot reveals that when rNSE parameter set result in an underestimation of loads in the upper range.

```
metrics <- c("NSE","rNSE")
compare_calib(nutrient, catch, alpha_p1 = 31, alpha_l1 = 0.02, sd_coef1 = 0.6,
              alpha_p2 = 50, alpha_l2 = 0.01, sd_coef2 = 0.6, years = c(1990:2018),
              name_basin = "Wisla", metrics)
```
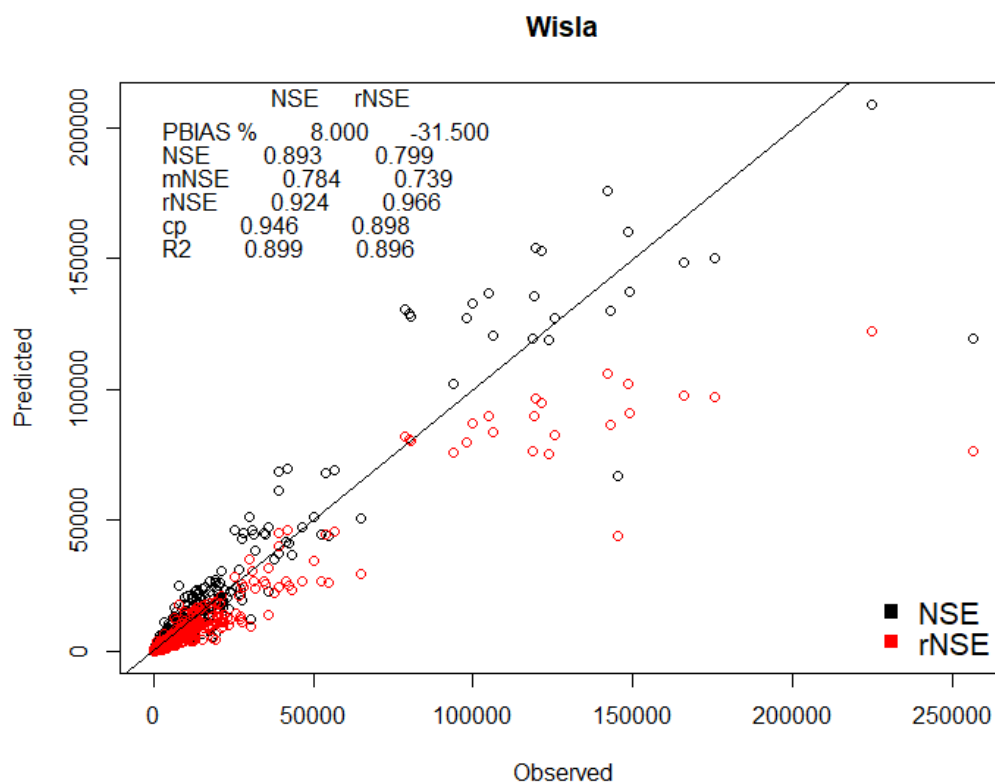

Figure 6: Result of compare_calib() for the Vistula river basin (TN), comparing the best set of parameters selected according to the NSE and rNSE metrics.

The `simobs_annual_plot()` function compares observed and predicted values over the years (Figure 7), which identifies erroneous model performance or data distribucion problems over time.

```
year_range <-  c(1994, 1996:2001, 2006:2009, 2012)
simobs_annual_plot(nutrient, catch, alpha_P, alpha_L, sd_coeff, year_range,
                   name_basin = "Wisla", max_value = 10000)
```
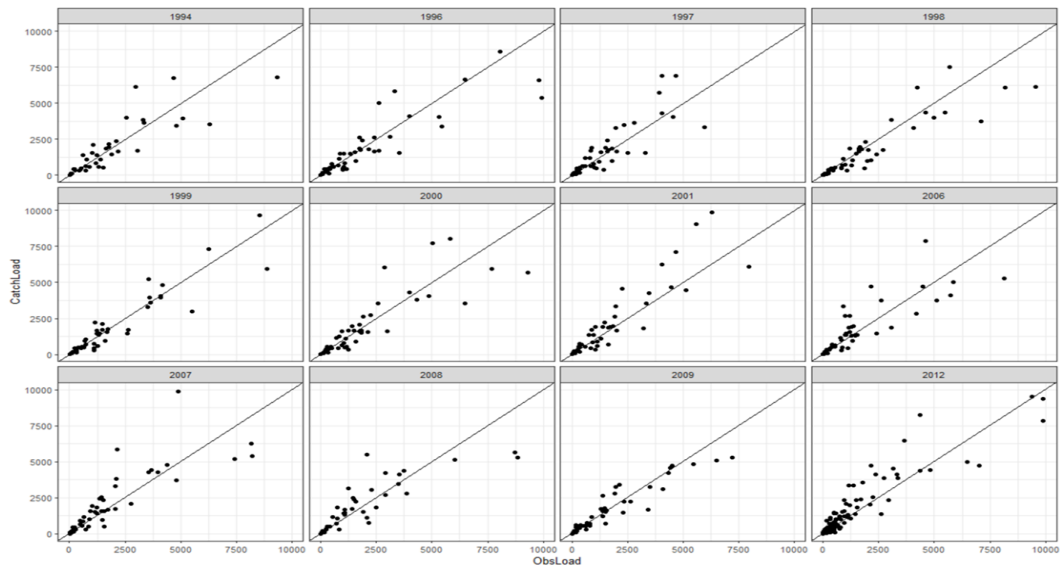
Figure 7: Result of the simobs_annual_plot() for the Vistula river basin (TN), with the parameter set selected for highest NSE and limiting the plots to the interval between 0 and 10000

3.3 Estimation of catchment nutrient loads and contribution in the basin

Once the most appropriate parameter set has been selected, it is possible to run the model to estimate nutrient loads in the region. The green_shares() function runs GREEN with the selected parameter set and returns catchment nutrient loads and the contributions by source in the simulation period.

The nutrient_tserie() function shows the temporal evolution of the total load at the basin outlet in the simulation period with contributions by sources (Figure 8). Other options of the function show nutrient loads in different zones of the basin (upper, central and lower part).

```
nutrient_load <- green_shares(nutrient, catch, alpha_P, alpha_L, sd_coeff, years)
nutrient_tserie(nutrient_load, geometry, "Wisla basin", "gr3")
```
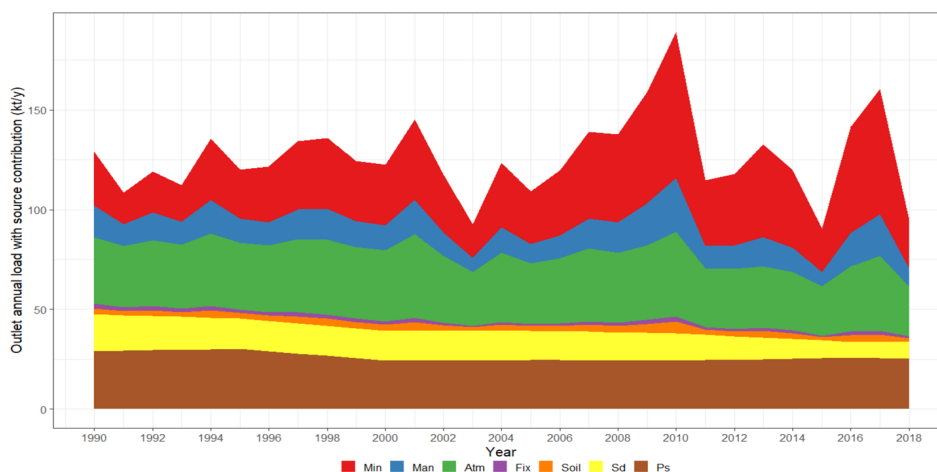

Figure 8: TN load at the Vistula river basin outlet from 1990 to 2018. Colors indicate the contribution of different sources (Min = mineral, Man = manure, Atm = atmospheric deposition, Fix = crop fixation, Soil = soil fixation, Sd = scattered dwellings and Ps = point sources).

The nutrient_maps() function uses the object returned by green_shares() to generate maps of the distribution of nutrient loads by source for a given year or as the mean for several years (Figure 9). The results are shown in logarithmic scale to improve the visualization as nutrient loads in a region may vary over several orders of magnitude. Alternative options of the function show the total load at the outlets of the catchment or the specific loads (i.e. load per catchment area; in $kt/km^2/y$).

```
map_title <- "Output Loads for the Vistula basin"
nutrient_maps(nutrient_load, geometry, map_title, "gr1", legend_position = 1)
```
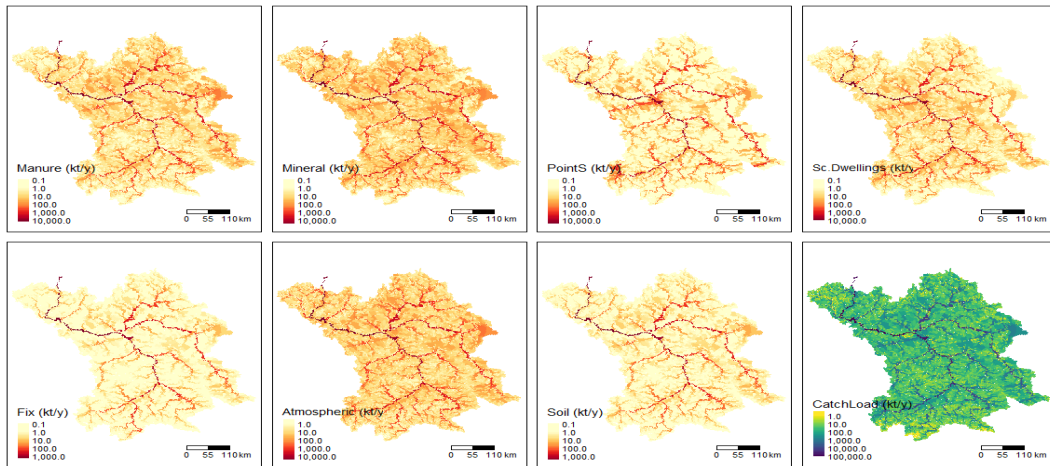

Figure 9: Maps of TN loads in the Vistula river basin by different sources and in total (the sum of all other maps, bottom right), in logarithm scale as generated with nutrient_maps() function.

3.4 Estimation of nutrient fluxes and sources contribution in the basin

The region_nut_balance() function runs GREEN with the selected parameter set, and returns the nutrient mass balance and fluxes of the region (mean value for the simulation period). The results of this function can be visualized in a Sankey diagram with the function N4_sankey() (Figure 10) .

```
nut_bal <- region_nut_balance(nutrient, catch, alpha_P, alpha_L, sd_coeff,
                               years)
sank <- N4_sankey(nut_bal)
```
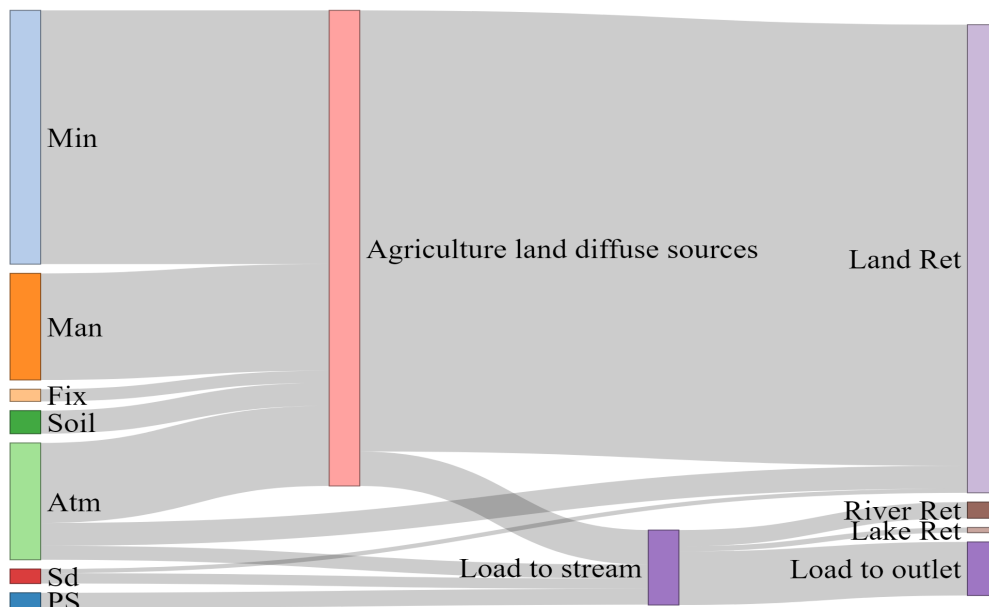

Figure 10: Sankey plot for the Vistula river basin for TN scenario, average for period 1990-2018. The plot represents nitrogen input sources on the left (Min=mineral, Man=manure, Atm=atmospheric deposition, Fix=crop fixation, Soil=soil fixation, Sd=scattered dwellings and Ps=point sources), and nitrogen sinks (land, river and lake retention) and outlet discharge (load to outlet) on the right. The bars in the middle visualize nitrogen agricultural diffuse sources and loads to the stream network.

4 Conclusion

The **GREENeR** package provides several functions to assess nutrient pressures in fresh and coastal waters based on the GREEN model (Bruna Grizzetti, Bouraoui, and Aloe 2012; Bruna Grizzetti et al. 2021). It allows assessing annual nutrient loads and concentrations throughout the river network and at river outlets, as well as the contributions of diffuse and point sources to the total load. **GREENeR** includes functions to route nutrient sources in a region, considering different pathways and the hydrological structure of the river network. The package is enriched by functions that aid selecting the set of parameters that best suits the study scope. Several functions assist in the preparation of scenarios by assembling input data from the appropriate sources, and in visualising inputs and nutrient fluxes in space and time. The version of the GREEN model implemented in the package is computationally efficient and includes parallel running capabilities for the calibration process, greatly reducing the time required for large basins or regional applications, e.g. at continental scale as in (Vigiak et al. 2023).

5 Acknowledgements

6 Appendix

6.1 Appendix 1

Diffuse and point sources are defined differently for each nutrient module, i.e. nitrogen or phosphorus. Equation (5) formulates the general GREEN function stated in equation (1) in the case of nitrogen. In GREEN nitrogen model, the total nitrogen load L_i of catchment i is estimated as:

$$L_i = (1 - Lret_i) \cdot ((MinN_i + ManN_i + FixN_i + SoilN_i + (1 - FF_i) \cdot AtmN_i) \cdot$$
$$(1 - Bret_i) + 0.38 \cdot FF_i \cdot AtmN_i + sd_{coeff} \cdot SdN_i + PsN_i + U_i) \cdot (1 - Rret_i) \tag{5}$$

where $MinN_i$ is the annual amount of nitrogen from mineral fertilisers (ton/yr); $ManN_i$ is the annual amount of nitrogen in manure fertilisers (ton/yr); $FixN_i$ is the annual amount of nitrogen fixation by leguminous crops and fodder (ton/yr); $SoilN_i$ is the annual amount of nitrogen fixation by bacteria in soils (ton/yr); $AtmN_i$ is the annual nitrogen deposition from atmosphere (ton/yr); FF_i is the non-agricultural land cover in the catchment (fraction); SdN_i is the nitrogen input from scattered dwellings (ton/yr); PsN_i is the nitrogen input from point sources (ton/yr); and, finally, U_i is the nitrogen load from upstream catchments (ton/yr).

Note that nitrogen atmospheric deposition losses are split into two parts, i.e. inputs to agricultural land, which undergo the basin retention (such as crop uptake) that depends on annual precipitation, while in forest areas they are reduced by a fixed rate, before entering into the stream. The contribution of atmospheric nitrogen deposition in non-agricultural land is thus estimated as $0.38 \cdot FF \cdot AtmN$. For an atmospheric deposition of 10 kgN/ha this corresponds to a background of 3.8 kgN/ha (in line with the values reported by HELCOM (2003)).

Phosphorous background losses are split into two parts, with the inputs to agricultural land undergoing basin retention, while in forest areas they are considered entering into the stream. Background losses for phosphorus are estimated at 0.15 kgP/ha (in line with the values reported by (HELCOM 2003).

In GREEN phosphorus model, the total annual phosphorus load L_i of catchment i the equation estimates:

$$L_i = (1 - Lret_i) \cdot ((MinP_i + ManP_i + (1 - FF_i) \cdot BgP_i) \cdot (1 - Bret_i) +$$
$$FF_i \cdot BgP_i + sd_{coeff} \cdot SdP_i + PsP_i + Ui_i) \cdot (1 - Rret_i) \tag{6}$$

where $MinP_i$ is the annual amount of phosphorus mineral fertilisers (ton/yr); $ManP_i$ is the annual amount of phosphorus in manure fertilisers (ton/yr); FF_i is the non-agricultural land cover in the catchment (fraction); BgP_i is the annual amount of phosphorus background losses (ton/yr); SdP_i is the phosphorus input from scattered dwellings (ton/yr); PsP_i is the phosphorus input from point sources (ton/yr); and Ui_i is the phosphorus load from upstream catchments (ton/yr).

6.2 Appendix 2

The `calib_green()` function applies the following GoF metrics (Althoff and Rodrigues 2021; Mauricio Zambrano-Bigiarini 2014) (NSE, rNSE, mNSE, R2, PBIAS), where:

- NSE: Nash-Sutcliffe efficiency.
- rNSE: Relative Nash-Sutcliffe efficiency.
- mNSE: Modified Nash-Sutcliffe efficiency.
- cp: Coefficient of Persistence.
- VE: Volumetric Efficiency.
- KGE : Kling-Gupta efficiency.
- d : Index of Agreement.
- md: Modified index of agreement.
- rd: Relative Index of Agreement.
- r: Linear correlation coefficient.
- R2: Coefficient of determination.
- PBIAS: Percent bias.
- MAE: mean absolute error.
- RMSE: Root mean square error.
- ME: Mean error.
- MSE: Mean square error.
- NRMSE: Normalized Root Mean Square Error.

See (Papacharalampous, Tyralis, and Koutsoyiannis 2019; Nash and Sutcliffe 1970; Kitanidis and Bras 1980; Yapo, Gupta, and Sorooshian 1996; Krause, Boyle, and Bäse 2005; Criss and Winston 2008; Hoshin Vijai Gupta, Sorooshian, and Yapo 1998; Mauricio Zambrano-Bigiarini 2014) for further details.

References

Althoff, Daniel, and Lineu Neiva Rodrigues. 2021. "Goodness-of-Fit Criteria for Hydrological Models: Model Calibration and Performance Assessment." *Journal of Hydrology* 600: 126674. https://doi.org/10.1016/j.jhydrol.2021.126674.

Arheimer, Berit, Joel Dahné, and Chantal Donnelly. 2012. "Climate Change Impact on Riverine Nutrient Load and Land-Based Remedial Measures of the Baltic Sea Action Plan." *Ambio* 41 (6): 600–612. https://doi.org/10.1007/s13280-012-0323-0.

Bartosova, Alena, René Capell, Jørgen E Olesen, Mohamed Jabloun, Jens Christian Refsgaard, Chantal Donnelly, Kari Hyytiäinen, Sampo Pihlainen, Marianne Zandersen, and Berit Arheimer. 2019. "Future Socioeconomic Conditions May Have a Larger Impact Than Climate Change on Nutrient Loads to the Baltic Sea." *Ambio* 48 (11): 1325–36. https://doi.org/10.1007/s13280-019-01243-5.

Beusen, A. H. W., J. C. Doelman, L. P. H. Van Beek, P. J. T. M. Van Puijenbroek, J. M. Mogollón, H. J. M. Van Grinsven, E. Stehfest, D. P. Van Vuuren, and A. F. Bouwman. 2022. "Exploring River Nitrogen and Phosphorus Loading and Export to Global Coastal Waters in the Shared Socio-Economic Pathways." *Global Environmental Change* 72: 102426. https://doi.org/https://doi.org/10.1016/j.gloenvcha.2021.102426.

Bouraoui, F, V Thieu, B Grizzetti, W Britz, and G Bidoglio. 2014. "Scenario Analysis for Nutrient Emission Reduction in the European Inland Waters." *Environmental Research Letters* 9 (12): 125007. https://doi.org/10.1088/1748-9326/9/12/125007.

Butler, Martha P, Patrick M Reed, Karen Fisher-Vanden, Klaus Keller, and Thorsten Wagener. 2014. "Identifying Parametric Controls and Dependencies in Integrated Assessment Models Using Global Sensitivity Analysis." *Environmental Modelling & Software* 59: 10–29. https://doi.org/10.1016/j.envsoft.2014.05.001.

Carnell, Rob. 2012. *Lhs: Latin Hypercube Samples*. https://CRAN.R-project.org/package=lhs.

Criss, Robert E, and William E Winston. 2008. "Do Nash Values Have Value? Discussion and Alternate Proposals." *Hydrological Processes: An International Journal* 22 (14): 2723–25. https://doi.org/10.1002/hyp.7072.

De Jager, A, and J Vogt. 2007. "Rivers and Catchments of Europe-Catchment Characterisation Model (CCM)." *European Commission, Joint Research Centre (JRC), Ispra, Italy.*

Doherty, John, and Brian E Skahill. 2006. "An Advanced Regularization Methodology for Use in Watershed Model Calibration." *Journal of Hydrology* 327 (3-4): 564–77. https://doi.org/10.1016/j.jhydrol.2005.11.058.

EEA. 2021. "European Environment Agency (EEA)." *WISE Water Framework Directive Database*. https://www.eea.europa.eu/data-and-maps/data/wise-wfd-4.

Frank, Ildiko E, and Roberto Todeschini. 1994. *The Data Analysis Handbook.* Elsevier.

Fu, Baihua, Wendy S Merritt, Barry FW Croke, Tony R Weber, and Anthony J Jakeman. 2019. "A Review of Catchment-Scale Water Quality and Erosion Models and a Synthesis of Future Prospects." *Environmental Modelling & Software* 114: 75–97. https://doi.org/10.1016/j.envsoft.2018.12.008.

Grizzetti, B, F Bouraoui, and G De Marsily. 2008. "Assessing Nitrogen Pressures on European Surface Water." *Global Biogeochemical Cycles* 22 (4). https://doi.org/10.1029/2007GB003085.

Grizzetti, B, F Bouraoui, G De Marsily, and G Bidoglio. 2005. "A Statistical Method for Source Apportionment of Riverine Nitrogen Loads." *Journal of Hydrology* 304 (1-4): 302–15. https://doi.org/10.1016/j.jhydrol.2004.07.036.

Grizzetti, Bruna, Fayçal Bouraoui, and Alberto Aloe. 2012. "Changes of Nitrogen and Phosphorus Loads to European Seas." *Global Change Biology* 18 (2): 769–82. https://doi.org/10.1111/j.1365-2486.2011.02576.x.

Grizzetti, Bruna, Olga Vigiak, Angel Udias, Alberto Aloe, Michela Zanni, Faycal Bouraoui, Alberto Pistocchi, et al. 2021. "How EU Policies Could Reduce Nutrient Pollution in European Inland and Coastal Waters." *Global Environmental Change* 69: 102281. https://doi.org/10.1016/j.gloenvcha.2021.102281.

Gupta, Hoshin Vijai, Soroosh Sorooshian, and Patrice Ogou Yapo. 1998. "Toward Improved Calibration of Hydrologic Models: Multiple and Noncommensurable Measures of Information." *Water Resources Research* 34 (4): 751–63. https://doi.org/10.1029/97WR03495.

Gupta, Hoshin V, Harald Kling, Koray K Yilmaz, and Guillermo F Martinez. 2009. "Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling." *Journal of Hydrology* 377 (1-2): 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.

HELCOM, Helsinki Commission. 2003. "Fourth Baltic Sea Load Compilation (PLC-4)." 93. *Fourth Baltic Sea Load Compilation*.

Kitanidis, Peter K, and Rafael L Bras. 1980. "Real-Time Forecasting with a Conceptual Hydrologic Model: 2. Applications and Results." *Water Resources Research* 16 (6): 1034–44. https://doi.org/10.1029/WR016i006p01034.

Krause, Peter, DP Boyle, and Frank Bäse. 2005. "Comparison of Different Efficiency Criteria for Hydrological Model Assessment." *Advances in Geosciences* 5: 89–97. https://doi.org/10.5194/adgeo-5-89-2005.

Kronvang, Brian, Josef Hezlar, Paul Boers, Jens P Jensen, Horst Behrendt, Tom Anderson, Berit Arheimer, Marcus Venohr, Carl Christian Hoffmann, and CB Nielsen. 2004. "Nutrient Retention Handbook." *Software Manual for EUROHARP-NUTRET and Scientific Review on Nutrient Retention. EUROHARP Report*, 9–2004.

Kumarasamy, Karthik, and Patrick Belmont. 2018. "Calibration Parameter Selection and Watershed Hydrology Model Evaluation in Time and Frequency Domains." *Water* 10 (6): 710. https://doi.org/10.3390/w10060710.

La Notte, Alessandra, Joachim Maes, Silvana Dalmazzone, Neville D Crossman, Bruna Grizzetti, and Giovanni Bidoglio. 2017. "Physical and Monetary Ecosystem Service Accounts for Europe: A Case Study for in-Stream Nitrogen Retention." *Ecosystem Services* 23: 18–29. https://doi.org/10.1016/j.ecoser.2016.11.002.

Leip, Adrian, Gilles Billen, Josette Garnier, Bruna Grizzetti, Luis Lassaletta, Stefan Reis, David Simpson, et al. 2015. "Impacts of European Livestock Production: Nitrogen, Sulphur, Phosphorus and Greenhouse Gas Emissions, Land-Use, Water Eutrophication and Biodiversity." *Environmental Research Letters* 10 (11): 115004. https://doi.org/10.1088/1748-9326/10/11/115004.

Ludwig, W, AF Bouwman, E Dumont, and F Lespinas. 2010. "Water and Nutrient Fluxes from Major Mediterranean and Black Sea Rivers: Past and Future Trends and Their Implications for the Basin-Scale Budgets." *Global Biogeochemical Cycles* 24 (4). https://doi.org/10.1029/2009GB003594.

Madsen, Henrik. 2003. "Parameter Estimation in Distributed Hydrological Catchment Modelling Using Automatic Calibration with Multiple Objectives." *Advances in Water Resources* 26 (2): 205–16. https://doi.org/10.1016/S0309-1708(02)00092-1.

Malagó, Anna, Fayçal Bouraoui, Bruna Grizzetti, and Ad De Roo. 2019. "Modelling Nutrient Fluxes into the Mediterranean Sea." *Journal of Hydrology: Regional Studies* 22: 100592. https://doi.org/10.1016/j.ejrh.2019.01.004.

Manache, Gemma, and Charles S Melching. 2004. "Sensitivity Analysis of a Water-Quality Model Using Latin Hypercube Sampling." *Journal of Water Resources Planning and Management* 130 (3): 232–42. https://doi.org/10.1061/(ASCE)0733-9496(2004)130:3(232).

Mannina, G, A Cosenza, MB Neumann, and PA Vanrolleghem. 2018. "Global Sensitivity Analysis in Wastewater Treatment Modelling." *Advances in Wastewater Treatment* 1: 32.

Mauricio Zambrano-Bigiarini. 2014. *hydroGOF: Goodness-of-Fit Functions for Comparison of Simulated and Observed Hydrological Time Series*. https://doi.org/10.5281/zenodo.839854.

McKay, MD, RJ Beckman, and WJ Conover. 1979. "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code." *Technometrics* 21 (2): 239–45. https://doi.org/10.2307/1268522.

Melching, Charles S, and Willy Bauwens. 2001. "Uncertainty in Coupled Nonpoint Source and Stream Water-Quality Models." *Journal of Water Resources Planning and Management* 127 (6): 403–13. https://doi.org/10.1061/(ASCE)0733-9496(2001)127:6(403).

Messager, Mathis Loïc, Bernhard Lehner, Günther Grill, Irena Nedeva, and Oliver Schmitt. 2016. "Estimating the Volume and Age of Water Stored in Global Lakes Using a Geo-Statistical Approach." *Nature Communications* 7 (1): 1–11. https://doi.org/10.1038/ncomms13603.

Nash, J Eamonn, and Jonh V Sutcliffe. 1970. "River Flow Forecasting Through Conceptual Models Part i-a Discussion of Principles." *Journal of Hydrology* 10 (3): 282–90. https://doi.org/10.1016/0022-1694(70)90255-6.

Norton, John. 2015. "An Introduction to Sensitivity Assessment of Simulation Models." *Environmental Modelling & Software* 69: 166–74. https://doi.org/10.1016/j.envsoft.2015.03.020.

Papacharalampous, Georgia, Hristos Tyralis, and Demetris Koutsoyiannis. 2019. "Comparison of Stochastic and Machine Learning Methods for Multi-Step Ahead Forecasting of Hydrological Processes." *Stochastic Environmental Research and Risk Assessment* 33 (2): 481–514. https://doi.org/10.1007/s00477-018-1638-6.

Pianosi, Francesca, Keith Beven, Jim Freer, Jim W Hall, Jonathan Rougier, David B Stephenson, and Thorsten Wagener. 2016. "Sensitivity Analysis of Environmental Models: A Systematic Review with Practical Workflow." *Environmental Modelling & Software* 79: 214–32. https://doi.org/10.1016/j.envsoft.2016.02.008.

Puy, Arnald, Samuele Lo Piano, Andrea Saltelli, and Simon A. Levin. 2022. "Sensobol: An r Package to Compute Variance-Based Sensitivity Indices." *Journal of Statistical Software* 102 (5): 1–37. https://doi.org/10.18637/jss.v102.i05.

Refsgaard, Jens Christian, and Hans Jørgen Henriksen. 2004. "Modelling Guidelines—-Terminology and Guiding Principles." *Advances in Water Resources* 27 (1): 71–82. https://doi.org/10.1016/j.advwatres.2003.08.006.

Saltelli, Andrea, Paola Annoni, et al. 2011. "Sensitivity Analysis."

Saltelli, Andrea, Stefano Tarantola, and Francesca Campolongo. 2000. "Sensitivity Analysis as an Ingredient of Modeling." *Statistical Science*, 377–95. http://www.jstor.org/stable/2676831.

Schwarz, Gregory E, Anne B Hoos, RB Alexander, and RA Smith. 2006. "The SPARROW Surface Water-Quality Model: Theory, Application and User Documentation."

Seitzinger, SP, JA Harrison, Egon Dumont, Arthur HW Beusen, and AF Bouwman. 2005. "Sources and Delivery of Carbon, Nitrogen, and Phosphorus to the Coastal Zone: An Overview of Global Nutrient Export from Watersheds (NEWS) Models and Their Application." *Global Biogeochemical Cycles* 19 (4). https://doi.org/10.1029/2005GB002606.

Shreve, Ronald L. 1966. "Statistical Law of Stream Numbers." *The Journal of Geology* 74 (1): 17–37. http://www.jstor.org/stable/30075174.

Singh, Vijay P, and Donald K Frevert. 2002. *Mathematical Models of Large Watershed Hydrology*. Water Resources Publication.

Smith, Richard A, Gregory E Schwarz, and Richard B Alexander. 1997. "Regional Interpretation of Water-Quality Monitoring Data." *Water Resources Research* 33 (12): 2781–98. https://doi.org/10.1029/97WR02171.

Strahler, Arthur N. 1957. "Quantitative Analysis of Watershed Geomorphology." *Eos, Transactions American Geophysical Union* 38 (6): 913–20. https://doi.org/10.1029/TR038i006p00913.

Vigiak, Olga, Angel Udías, Bruna Grizzetti, Michela Zanni, Alberto Aloe, Franz Weiss, Jordan Hristov, Berny Bisselink, Ad de Roo, and Alberto Pistocchi. 2023. "Recent Regional Changes in Nutrient Fluxes of European Surface Waters." *Science of The Total Environment* 858: 160063. https://data.jrc.ec.europa.eu/collection/wpi.

Westerberg, IK, J-L Guerrero, PM Younger, KJ Beven, Jan Seibert, Sven Halldin, JE Freer, and C-Y Xu. 2011. "Calibration of Hydrological Models Using Flow-Duration Curves." *Hydrology and Earth System Sciences* 15 (7): 2205–27. https://doi.org/10.5194/hess-15-2205-2011.

Wöhling, Thomas, Luis Samaniego, and Rohini Kumar. 2013. "Evaluating Multiple Performance Criteria to Calibrate the Distributed Hydrological Model of the Upper Neckar Catchment." *Environmental Earth Sciences* 69 (2): 453–68. https://doi.org/10.1007/s12665-013-2306-2.

Yapo, Patrice O, Hoshin Vijai Gupta, and Soroosh Sorooshian. 1996. "Automatic Calibration of Conceptual Rainfall-Runoff Models: Sensitivity to Calibration Data." *Journal of Hydrology* 181 (1-4): 23–48. https://doi.org/10.1016/0022-1694(95)02918-4.

Angel Udías
European Commission, Joint Research Centre (JRC)
Via Enrico Fermi 2749, I-21027, Ispra
Italy
Departament of Computer Science and Statistics, Rey Juan Carlos University
Calle Tulipán S/N, 28933, Móstoles, Madrid
Spain
ORCiD: 0000-0003-1219-0465
angel.luis.udias@urjc.es

Bruna Grizzetti
European Commission, Joint Research Centre (JRC)
Via Enrico Fermi 2749, I-21027, Ispra
Italy
bruna.grizzetti@jrc.it

Olga Vigiak
European Commission, Joint Research Centre (JRC)
Via Enrico Fermi 2749, I-21027, Ispra
Italy
Olga.VIGIAK@ec.europa.eu

Alberto Aloe
European Commission, Joint Research Centre (JRC)
Via Enrico Fermi 2749, I-21027, Ispra
Italy
alberto.aloe@jrc.it

Cesar Alfaro
Departament of Computer Science and Statistics, Rey Juan Carlos University
Calle Tulipán S/N, 28933, Móstoles, Madrid
Spain
ORCiD: 0000-0002-9146-4067
cesar.alfaro@urjc.es

Javier Gomez
Departament of Computer Science and Statistics, Rey Juan Carlos University
Calle Tulipán S/N, 28933, Móstoles, Madrid
Spain
ORCiD: 0000-0001-6434-7263
javier.gomez@urjc.es