Water quality, discharge and catchment attributes for large-sample studies in Germany - QUADICA

Pia Ebeling[1], Rohini Kumar[2], Stefanie R. Lutz[1,3], Tam Nguyen[1], Fanny Sarrazin[2], Michael Weber[2], Olaf Büttner[4], Sabine Attinger[2], Andreas Musolff[1]

[1]Department of Hydrogeology, Helmholtz Centre for Environmental Research-UFZ, Leipzig, 04318, Germany
[2]Department of Computational Hydrosystems, Helmholtz Centre for Environmental Research-UFZ, Leipzig, 04318, Germany
[3]Copernicus Institute of Sustainable Development, Utrecht, 3584 CB, the Netherlands
[4]Department Aquatic Ecosystems Analysis and Management, Helmholtz Centre for Environmental Research-UFZ, Magdeburg, 39114, Germany

Correspondence to: Pia Ebeling (pia.ebeling@ufz.de)

Abstract. Environmental data are the key to define and address water quality and quantity challenges at catchment scale. Here, we present the first large-sample water quality data set for 1386 German catchments covering a large range of hydroclimatic, topographic, geologic, land use and anthropogenic settings. QUADICA (water QUAlity, DIscharge and Catchment Attributes for large-sample studies in Germany) combines water quality with water quantity data, meteorological and nutrient forcing data, and catchment attributes. The data set comprises time series of riverine macronutrient concentrations (species of nitrogen, phosphorus and organic carbon) and diffuse nitrogen forcing data at catchment scale (nitrogen surplus, atmospheric deposition and fixation). Time series are generally aggregated to an annual basis; however, for 140 stations with long-term water quality and quantity data (more than 20 years), we additionally present monthly median discharge and nutrient concentrations, flow-normalized concentrations and corresponding mean fluxes as outputs from weighted regressions on time, discharge, and season (WRTDS). The catchment attributes include catchment nutrient inputs from point and diffuse sources and characteristics from topography, climate, land cover, lithology and soils. This comprehensive, freely available data collection with a large spatial and temporal coverage can facilitate large-sample data-driven water quality assessments at catchment scale as well as mechanistic modeling studies.

1 Introduction

Understanding hydrological and biogeochemical processes at various spatiotemporal scales is a major goal in catchment hydrology and is particularly relevant for robust predictions of water quantity and quality and adequate catchment management. Analyzing observations of spatial and temporal dynamics of water quantity and quality at the catchment scale can give insights into relevant processes using a *"pattern to process"* approach (Sivapalan, 2006). Especially large-sample studies covering a wide range of catchments can advance our knowledge on patterns across scales, catchment similarity, and dominant processes, beyond a single catchment or local behavior (Addor et al., 2020; Kingston et al., 2020). Such studies allow for generalizable theories and applications by *"balancing depth with breadth"* and facilitate classifications, regionalization and a better understanding of uncertainty in model predictions (Gupta et al., 2014). Thus, environmental data are the key for process understanding and hypothesis testing (Li et al., 2021). The collection and availability of water quantity and quality data are steadily increasing with technological advances (Rode et al., 2016), but particularly harmonized and quality controlled large-sample data sets of water quality and quantity along with catchment attributes are needed. These enable identifying and characterizing water quality and quantity response patterns and relationships with potential controls, facilitate hypothesis testing and thus advance our understanding of the complex coupled hydrological and biogeochemical systems across larger samples and domains (Li et al., 2021).

In recent years, the application of large-sample studies has been advancing fast for (surface) water quantity studies investigating dominant processes and drivers of water flow characteristics. Gupta et al. (2014) provided an overview of such studies, with the first ones published in the 1990s. They were followed by a recent surge in studies documenting and analyzing large-sample hydrologic data sets such as Newman et al. (2015), Kuentz et al. (2017), Do et al. (2017), Gnann et al. (2020), Tarasova et al. (2020), Merz et al. (2020). These studies identified catchment typologies, archetypal behavior and underlying controls e.g. regarding discharge variability across Europe (Kuentz et al., 2017), catchments with similar runoff event types (Tarasova et al., 2020) or how catchment discharge attenuates and shifts climate seasonality (Gnann et al., 2020).

In contrast, large-sample studies for water quality are less common. Nevertheless, some recent large-sample water quality studies provided a basis for enhancing our understanding of catchment functioning

in terms of mobilization, transport, and environmental fate of solutes and particulates and the generality of these functions. For example, Monteith et al. (2007) linked widespread positive trends in DOC concentrations observed in Europe and North America to decreasing atmospheric sulphur and chloride depositions. Godsey et al. (2009) and Godsey et al. (2019) provided wide evidence that weathering derived solutes are mostly exported chemostatically with low concentration variance. Basu et al. (2010) derived the hypothesis of chemostatic nutrient export resulting from homogenized sources due to the legacy of high inputs. More recently, Zarnetske et al. (2018) and Ebeling et al. (2021a) both provided evidence of widespread transport-limited DOC export from small to large catchments. However, several questions of general patterns, catchment similarities and typologies and their underlying controls remain open, for example, concerning the extent and recovery of nutrient legacy for both nitrogen (N) and phosphorous (P) (Chen et al., 2018), the extent of macronutrient interactions in differing landscape and anthropogenic settings and throughout the river network (Wollheim et al., 2018), and the impact of climate change on water quality trajectories in various catchments (Kaushal et al., 2018).

At the moment, large-sample studies are still hampered by limited availability (number of stations, number of samples, and covered regions) and accessibility of spatially and temporally harmonized large-sample data collections (e.g., Addor et al., 2020), despite recent efforts to make consistent large-sample data sets of catchment hydrology for both water quantity and water quality in streams publicly available (e.g., Virro et al., 2021). Prominent examples for large-sample hydrological data sets including catchment attributes are the Catchment Attributes and MEteorology for Large-sample Studies (CAMELS) data sets, available for the USA (Addor et al., 2017), Chile (Alvarez-Garreton et al., 2018), Brazil (Chagas et al., 2020), Great Britain (Coxon et al., 2020), and Australia (Fowler et al., 2021). More recently, the multi-national LArge-SaMple DAta for Hydrology and Environmental Sciences (LamaH, Klingler et al., 2021) have provided hydrometeorological time series at an hourly resolution together with catchment attributes. For stream water quality, currently available large-sample data sets focus on water quality time series only but lack additional data. Recently, two global databases of surface water quality were published, which combine data of several existing databases in homogenized and quality checked form: the Surface Water Chemistry database (SWatCh; Rotteveel and Sterling, 2021) with a focus on variables relevant for acidification and the Global River Water Quality Archive (GRQA; Virro et al., 2021) with a focus on

macronutrients. Both include the global databases Global Freshwater Quality Database (GEMStat; UNEP, 2018), GLObal RIver CHemistry database (GLORICH; Hartmann et al., 2014) and the European WaterBase (EEA, 2020), although spatiotemporal coverage of the data varies strongly. These are important recent advances towards open science in water quality research. However, to the authors' knowledge, there is currently no combined, ready-to-use data set of metrics of water quality, quantity, catchment attributes and forcing data (such as meteorological and nutrient inputs), which would allow investigating water quality dynamics and their controls. Moreover, large-sample and cross-regional studies are especially challenging in countries like Germany, where data responsibility is scattered between federal states, and where data are often not freely available nor homogenized between water quantity and quality stations. Nevertheless, there have been a few Germany-wide water quality studies on groundwater (Knoll et al., 2020) and surface water (Ebeling et al., 2021a) recently.

The key objective here is to provide a spatially and temporally consistent, comprehensive data set of joint water quality and quantity data, catchment attributes and nutrient inputs for German catchments, which is ready-to-use and freely available supporting an open science philosophy and FAIR data principles. In this "Water QUAlity, DIscharge and Catchment Attributes for large-sample studies in Germany" (QUADICA) data set, we have complemented available data sets of catchment attributes with new data on water quality and water quantity. These data include delineated catchment boundaries, catchment responses in terms of macronutrient concentrations (species of N, P and organic carbon (OC)) and discharge (Q), forcing data in terms of meteorological and diffuse nitrogen inputs and average catchment attributes. We distinguish stations with a high data availability, which allows further estimation of daily concentrations and fluxes using a regression approach, and stations with lower availability, for which aggregated observed concentrations are reported. For water quality (Section 3.1) and water quantity (Section 3.2), we provide

(1) time series of annual medians of observed macronutrient concentrations (dissolved and total forms of N, P and OC) and of observed discharge,

(2) time series of monthly and annual medians of estimated daily macronutrient concentrations and flow-normalized concentrations as well as mean nutrient fluxes and medians of observed discharge for stations with high data availability,

(3) monthly medians and 25th and 75th percentiles of observed concentrations and discharge over the whole time series.

Additionally, we provide time series of driving forces (Section 3.3 and 3.4) and catchment attributes (Section 4):

(4) time series of observed monthly meteorological forcing variables as catchment averages (Section 3.3),

(5) time series of estimated annual net diffuse nitrogen inputs to the catchments (Section 3.4),

(6) average catchment characteristics, i.e. topography, land cover, nutrient sources, lithology and soils, and hydroclimate (Section 4).

We envision that the QUADICA data set will directly enable large-sample assessments of mean concentrations and fluxes, and concentration and flux variability in terms of long-term trends, seasonality and relationships to discharge as well as their relationships to catchment attributes. We believe that the data set will allow better understanding of catchment functioning and water management beyond regional scales and stimulate provisioning and analysis of further water quality data at national to continental scales.

2 Catchment selection and delineation

The station selection and catchment delineation have been presented in a previous study (Ebeling et al., 2021a) and data repository (Ebeling, 2021) and are now included in the new QUADICA data set. All data sets use the same unique identifier (OBJECTID) for the stations and corresponding catchments. The station selection is based on riverine water quality data assembled from the German federal state environmental authorities, who are responsible for the routine monitoring of water quality in Germany (Musolff et al., 2020; Musolff, 2020) and take grab samples at approximately monthly intervals.

The following preprocessing steps were applied for each station and compound separately: we removed duplicates, negative and zero values and applied an outlier test for each time series (removing values above mean concentration + 4*standard deviation in logarithmic space, i.e. confidence level > 99.99 % for log normally distributed data). Finally, 1386 stations met the criteria concerning water quality data and catchment delineation as described in the following (Fig. 1):

1. Water quality data cover at least three years, include a minimum of 70 samples from 2000 to 2015 after preprocessing, and cover all seasons, i.e. seasonal coverage of at least 10 % of the samples in each quarter considering all possible combinations of three consecutive months (criteria one to three as described in Ebeling et al., 2021a). These criteria should ensure that a representative amount of data is available. Stations fulfilling these water quality data criteria for NO_3-N or PO_4-P were preselected (i.e., 1692 stations). Other variables (e.g. TP, TN, DOC) were not used in this initial step of station selection.

2. In a second step, we delineated the catchment area from topography for these preselected stations and verified them as described here. The topographic catchment boundaries were delineated based on a 100 m flow accumulation grid derived from a digital elevation model (DEM; resampled from 25 m to 100 m using the average; EEA, 2013) using spatial analysis tools and D8 flow direction type. The river network from the Rivers and Catchments of Europe - Catchment Characterisation Model (De Jager and Vogt, 2007) was used to burn by 10 m into the DEM before deriving the flow accumulation. The stations were snapped or manually moved towards the representative flow accumulation stream to define the catchment outlets (pour points). The resulting topography-based catchment polygons were quality-controlled manually by a comparison to the real river network. In case of major deviations, a few manual adaptations of the burned river segments were done if they substantially improved the overlap without hindering neighboring catchment delineations. In case of insufficient spatial overlap that could not be improved, stations were discarded from the selection. This resulted in a final set of 1386 catchments. The used DEM, flow direction and flow accumulation raster as well as the modified station locations and the river network are also provided in the data repository for further use.

The varying density of stations across Germany (Fig. 1a) has two main reasons: firstly, the provision of raw data varied in number of stations, number of samples per compound and station, and time series

length among the federal states; secondly, the topographic delineation of catchment boundaries was more successful where the topography is more pronounced, giving less delineable catchments in northern Germany. The delineated catchment boundaries are provided with the data set and enable the user to develop further geoinformation routines, e.g. to extract characteristics from other geographic data sets.

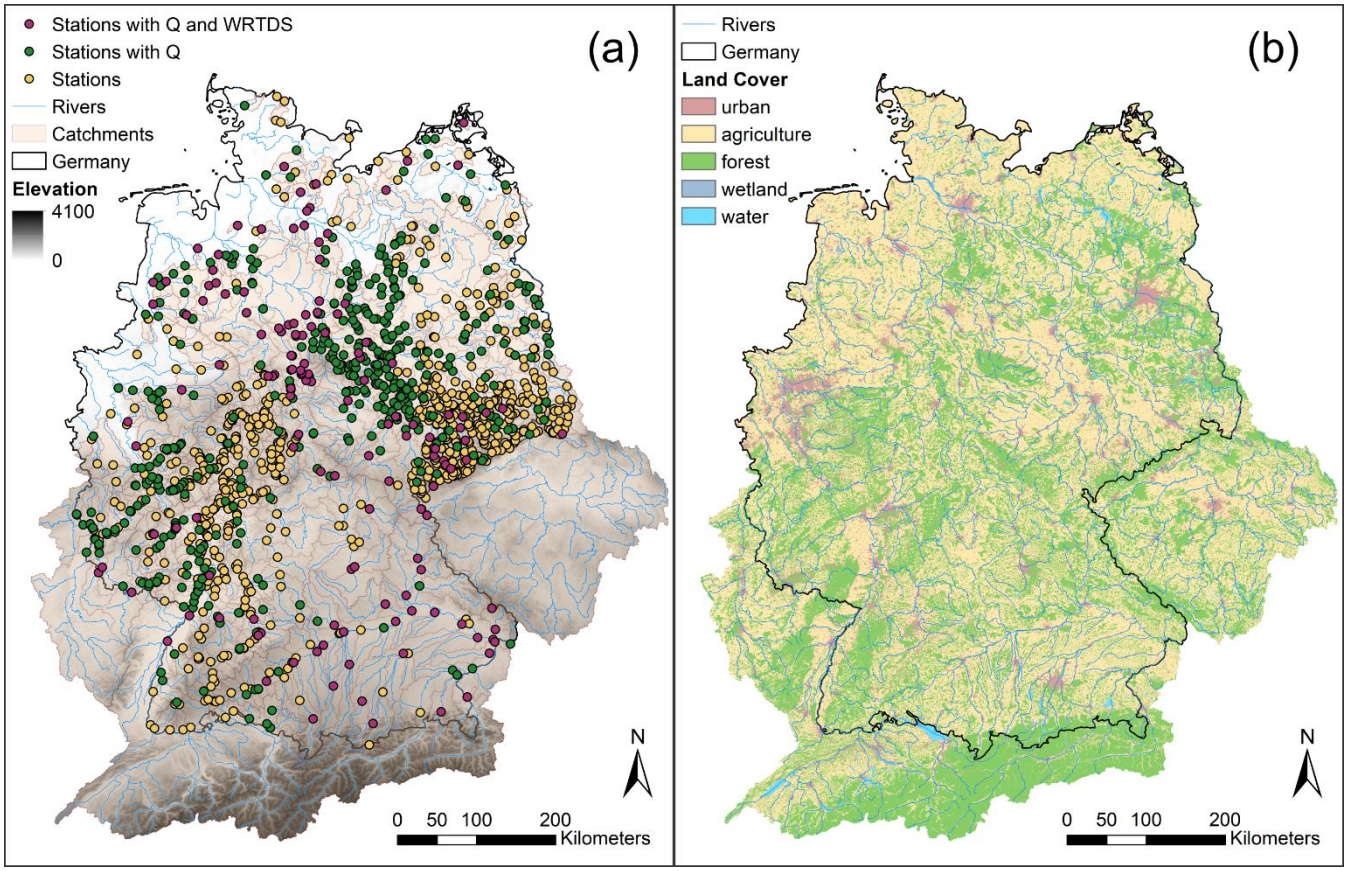

Figure 1: Map of (a) water quality stations, catchments and elevation (EEA, 2013) and (b) map of land cover (EEA, 2016a). Colors in (a) distinguish between stations with (green) and without discharge (Q) data (yellow) and long-term C-Q stations (dark purple) with high data availability (also WRTDS stations; for details see Section 3.1). WRTDS - Weighted Regression on Time, Discharge and Season.

3 Time series

For the 1386 delineated catchments, riverine concentration time series of nitrate (NO_3-N), mineral nitrogen (N_{min}), total nitrogen (TN), phosphate (PO_4-P), total phosphorus (TP), dissolved organic carbon (DOC), and total organic carbon (TOC) are provided (Table 1). They are supplemented by time series of

discharge (where available) and forcing variables (meteorological drivers and diffuse N input). Due to limited data availability, not all variables can be provided for all stations.

Table 1: Provided time series data, their basis (observed or estimated), aggregation type, temporal resolution and source of original data, which was used to calculate the aggregated data provided here. WRTDS -Weighted Regression on Time, Discharge and Season.

Variable	Section	Data basis	Temporal (Spatial) Aggregation	Temporal resolution	Source
Concentration of NO_3-N, N_{min}, TN, PO_4-P, TP, DOC, TOC	3.1	observed	median	annual	Musolff et al. (2020); Musolff (2020)
		daily estimated using WRTDS	median	monthly	Musolff et al. (2020); Musolff (2020)
		observed	long-term median	monthly	Musolff et al. (2020); Musolff (2020)
Discharge	3.2	observed	median	annual	Musolff et al. (2020); Musolff (2020)
		observed	median	monthly	Musolff et al. (2020); Musolff (2020)
		observed	long-term median	monthly	Musolff et al. (2020); Musolff (2020)
Precipitation	3.3	observed gridded	sum (average)	monthly	Cornes et al. (2018); E-Obs (2018)
Potential evapotranspiration	3.3	estimated	sum (average)	monthly	Cornes et al. (2018); E-Obs (2018)
Mean air temperature	3.3	observed gridded	average (average)	monthly	Cornes et al. (2018); E-Obs (2018)
Diffuse N input as total	3.4	estimated	(average)	annual	see Section 3.4
Diffuse N input from agricultural areas	3.4	estimated	(average)	annual	see Section 3.4

185

3.1 Water quality time series

3.1.1 Annual median concentrations

Annual medians of concentration data are presented for time series of the 1386 stations fulfilling the water quality criteria, analogously to the catchment selection criteria described in Section 2. . To calculate summary statistics, we substituted concentration values below the detection limit (left-censored data) with half the detection limit.

The resulting data density distributions over time and the number of years covered by each variable show the highest data availability for TOC, PO_4-P and NO_3-N in more recent years (Fig. 2). An overview of the time-series statistics for each variable is given in Table 2, while time series are shown in Appendix A Fig. A1. For NO_3-N concentrations, the number of stations with available data is 1339 and the median number of samples per station is 157. The earliest time series starts in 1954, while the median start across stations is in 1994. The median time series length is 19 years and the maximum time series length is 61 years. For PO_4-P concentrations, the number of stations with available data is 1330 and the median number of samples per station is 152. The earliest time series starts in 1965, while the median start across stations is in 1993. The median time series length is 20 years and the maximum time series length is 48 years. For TOC concentrations, the number of stations with available data is 1296 and the median number of samples per station is 139. The earliest time series starts in 1979, while the median start across stations is in 1999. The median time series length is 15 years and the maximum time series length is 36 years. For all water quality variables, the median of the first year of the time series is in the 1990s and the median number of samples per station and year is 12, indicating that grab samples were on average taken on a monthly basis. Note that the number of samples underlying the median values can differ between the different nutrient species so that the fraction of TN present as NO_3-N or TP present as PO_4-P may show inconsistencies for single stations (e.g. values above 1).

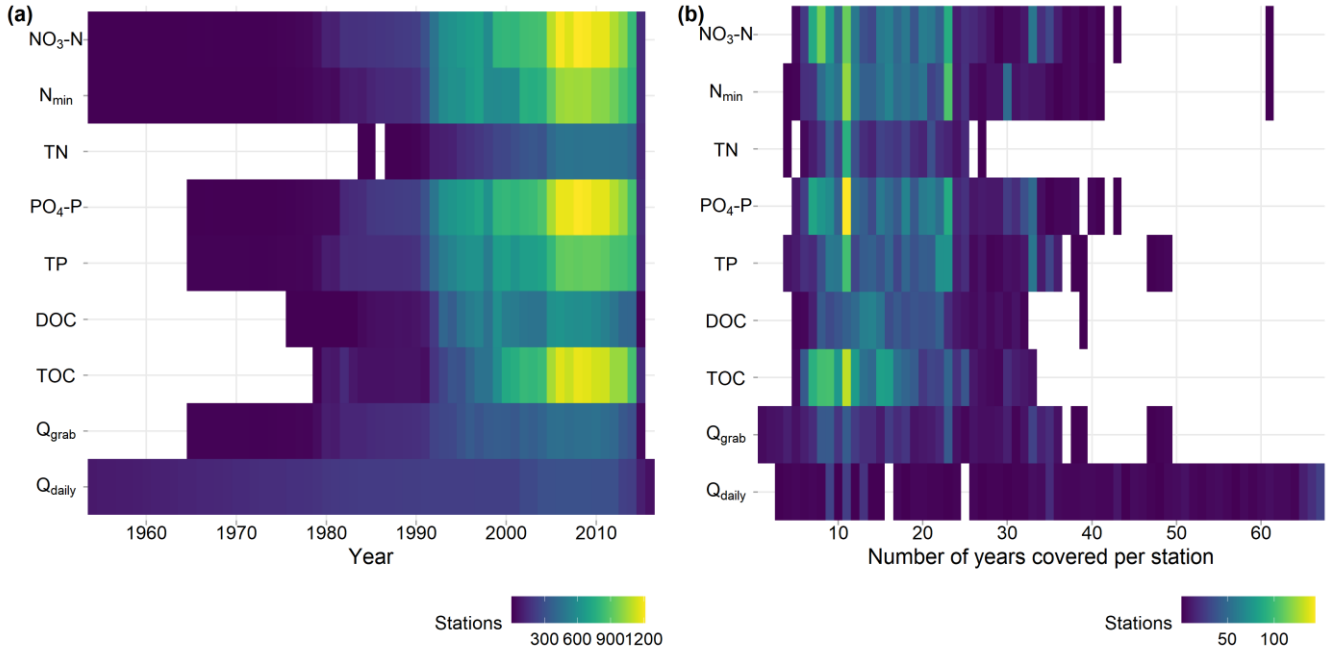

Figure 2: Heat map of (a) the number of stations with available annual medians over time and per variable and (b) the number of years covered by each station. Q_{grab} refers to the median Q from grab sample dates and Q_{daily} refers to median discharge from daily discharge (see Section 3.2.1 for details). For visualization purposes in (a) station counts from 1954 are shown, omitting one concentration and a few Q_{daily} records before 1954, and in (b) counts up to 67 years are shown, omitting three longer Q_{daily} records.

Table 2: Number of stations with available data for the water quality compounds and discharge during grab sampling dates, earliest and median start year of time series, maximum and median time series length and covered years (i.e. years with available data), median number of samples per stations and per station and year, and number of outliers removed. * omitting one sample from 1900.

Variable	**NO_3-N**	**N_{min}**	**TN**	**PO_4-P**	**TP**	**DOC**	**TOC**	**Q_{grab}**	**Q_{daily}**
Unit	**mg l^{-1}**	**mg l^{-1}**	**mg l^{-1}**	**mg l^{-1}**	**mg l^{-1}**	**mg l^{-1}**	**mg l^{-1}**	**m^3 s^{-1}**	**m^3 s^{-1}**
Number of stations	1339	1149	514	1330	1046	744	1296	581	324
Earliest start	1954*	1954	1984	1965	1965*	1976	1979	1965 *	1893
Median start year	1994	1993	1999	1994	1993	1993	1999	1993	1975
Median time series length per station (years covered)	19 (16)	21 (16)	15 (13)	20 (16)	21 (17)	20 (15)	15 (13)	19 (16)	38 (38)
Maximum time series length per station (years covered)	61* (61)	61 (61)	31 (27)	48 (43)	49* (49)	39 (39)	36 (33)	49* (49)	123 (123)

Total number of samples (incl. outliers)	309,965	235,015	92,876	297,591	258,059	139,440	239,282	156,388	>4 Mio
Median number of samples per station	157	153	149	152	165	164	139	170	13388
Median number of samples per station and year	12	12	12	12	12	12	12	13	365
Number of outliers	59	52	45	68	326	257	795	-	-
Maximum fraction of outliers per station [%]	1.7	1.7	1.7	1.7	2.5	3.1	3.6	-	-

3.1.2 Monthly median concentrations and mean fluxes for stations with high data availability

For the subset of stations with high data availability, a Weighted Regression on Time, Discharge and Season (WRTDS; Hirsch et al., 2010) was applied using the R package *EGRET* (version 3.0.2; Hirsch and De Cicco, 2015). We refer to these stations as 'WRTDS stations' for short. WRTDS represents long-term trends, seasonal components and discharge-related variability of the water quality variables (Hirsch et al., 2010). The criteria for a WRTDS application were checked for each station and compound separately using the preprocessed data as described in Section 2. The criteria were a time series of at least 20 years length, at least 150 samples of water quality, no data gaps larger than 20 % of the total time series length and a complete time series of daily discharge (see also Section 3.2.2). The number of WRTDS stations varies between 44 for TN and 126 for PO_4-P (Table 3), while the fraction of stations with high data availability varies between 4.9 % for TOC and 11.7 % for TP.

For WRTDS stations, we provide monthly and annual median estimated water quality and observed quantity data in addition to the annual observed data (see above). More specifically, we provide monthly and annual median concentration and flow-normalized concentration and mean flux estimates from the WRTDS model output and median observed discharge (see Section 3.2.2) if data are available for at least 80 % of the respective time frame. The median R^2 between WRTDS-modelled and observed

concentrations varies between 0.44 for DOC and TOC and 0.75 for TN (Table 3), while overall 69.3 % of the catchment and compound combinations have a median R^2 of at least 0.5. The median bias varies between -1.4 % for PO_4-P (negative values indicate overestimation) and 0.2 % for NO_3-N (positive values indicate underestimation). Overall, 51 % of the catchments have a bias below 1 % and 95 % below 5 % respectively. An overview of the availability of WRTDS stations and model performances is given in Table 3 and shown in Fig. A2, while their locations are shown in Fig. 1a and performances provided in the data repository.

Table 3: Number of stations with high data availability (WRTDS stations) for each compound and median coefficient of determination of WRTDS models.

Variable	**total**	**NO_3-N**	**N_{min}**	**TN**	**PO_4-P**	**TP**	**DOC**	**TOC**
Unit		**mg l^{-1}**	**mg l^{-1}**	**mg l^{-1}**	**mg l^{-1}**	**mg l^{-1}**	**mg l^{-1}**	**mg l^{-1}**
Number of WRTDS stations	140	125	97	44	126	122	61	64
Median R^2	0.61	0.63	0.71	0.75	0.69	0.53	0.44	0.44
Median bias [%]	-0.3	0.2	0.1	0.1	-1.4	-0.9	-0.6	-0.6

3.1.3 Monthly median concentrations over the time series

Next to annual and monthly time series, we provide long-term monthly medians over the complete time series of each station, enabling assessments of average seasonal variability. We also include the 25th and 75th percentiles to reflect the long-term variability of a given month. The provided data frame in QUADICA indicates the number of samples available for the corresponding month across the years, based on which representativeness can be assessed and quality criteria can be defined.

3.2 Water quantity time series

For about 43 % of the water quality stations (n=590), information on discharge is available (Fig. 1a) and provided harmonized with the water quality data (i.e. annual and monthly resolution). The discharge information is a collection of data provided by the federal states together with the concentration data

either as daily discharge time series or for the times of grab sampling of water quality. Additionally, we integrated daily discharge data from 53 stations available from the Global Runoff Data Center (GRDC) to increase the number of stations with available discharge time series. We matched GRDC gauging stations to the existing water quality stations using the "proximity/point distance" tool in ArcGIS with a search radius of 500 m. For each match, we checked the consistency of river names and visually the locations. The corresponding GRDC station numbers are indicated in the metadata of the water quality and quantity data set (Musolff, 2020). For the original daily discharge data, the reader may refer to the regularly published and accessible data at the GRDC portal (https://portal.grdc.bafg.de).

3.2.1 Annual median discharge

Annual median discharge is aggregated from available observed discharge data. For 324 water quality stations, a co-located Q station with a continuous daily Q record is available. However, the time series may include data gaps and for nine of the co-located discharge stations, the time series of discharge and concentration data do not overlap at all. For additional 266 stations, Q data was only available at the time that the grab samples were taken. This resulted in a set of 581 stations for which Q data are available on the sampling dates of concentration data. We extracted annual median discharge both from continuous daily data (Q_{daily}) and from dates when the water quality sample was taken (Q_{grab} with a median of 13 values per year, Table 2) for the water quality stations. The density distribution of stations with available annual discharge over time is shown in Fig. 2a. Similar to the concentration data, the data availability is higher in more recent years, with the maximum of 449 stations in 2010. The number of years covered is, however, higher compared to water quality data for several stations (Fig. 2b). For stations with available daily discharge data, both annual median values of the daily data and the data from grab sample days were compared (Fig. A3). Our results suggest that annual median values from grab sample dates can be considered to be robust estimates of annual median discharge as they have a negligible bias (bias=-0.5 %) and low scatter around the 1:1 line (R^2>0.99). The time series are shown in Appendix A Fig. A1. The data set additionally provides the number of samples used to calculate the medians as a measure of robustness.

3.2.2 Monthly median discharge

Monthly median discharge is provided for WRTDS stations. To fill gaps in the daily discharge time series of 45 stations required for WRTDS models (see Section 3.1.2), we used simulated discharge from the mesoscale hydrological model *mHM* (Kumar et al., 2013; Samaniego et al., 2010; Zink et al., 2017) if the regression coefficient (R^2) between observed and simulated discharge for the station was greater than 0.6. Subsequently, modelled discharge was bias-corrected with piecewise linear regressions and used for gap-filling (Ebeling et al., 2021b; Ehrhardt et al., 2021a). If modelled discharge was not available, small gaps (up to seven days) were interpolated with fixed-interval smoothing using the R package *baytrends* (Murphy et al., 2019). Note that the gap-filled discharge time series are used for the WRTDS models only. This includes the monthly and annual discharge data provided with the WRTDS data tables (as described in Section 3.1.2).

3.2.3 Monthly median discharge over the time series

Analogously to the water quality metrics (see Section 3.1.3), we provide long-term monthly median discharge and the 25^{th} and 75^{th} percentile over the whole time series if available for the station representing average discharge seasonality. The number of samples used for the calculation of medians is indicated as a measure of accuracy.

3.3 Meteorological time series

Meteorological time series are provided as spatial catchment averages on monthly resolution. We used the daily gridded product of climate variables (precipitation and maximum, minimum, and average air temperature) from the European Climate Assessment and Dataset project (E-OBS, v18.0e; Cornes et al., 2018). The advantage of a European data set is the coverage of transnational catchments as e.g. the Elbe or Rhine. The data sets are available at a spatial resolution of 0.1 degree over the period 1950-2018. The interpolation approach used to create the gridded fields uses a stochastic technique based on Gaussian Random Field, and involves several ground-based observation networks distributed across Europe (see Cornes et al., 2018 for more details). The daily fields of potential evapotranspiration are derived based on the method from Hargreaves and Samani (1985) at the same spatial resolution (0.1 degree) using the daily

(maximum, minimum and average) air temperature data sets. We then calculated the spatial averages of daily climate variables (precipitation, air temperature, potential evapotranspiration) for all water quality stations, considering the corresponding (upstream) catchment area. Monthly estimates of total precipitation and potential evapotranspiration, and average air temperatures were subsequently calculated for each study basin.

3.4 Time series of net N input from diffuse sources

For the period 1950-2015, we provide time series of catchment-scale N surplus, i.e. the net diffuse N input, which is the sum of N inputs minus the sum of N outputs from harvesting. At the catchment scale, the N surplus is the sum of N surplus on agricultural N_{agri} (kg y^{-1} ha^{-1}) and non-agricultural areas $N_{nonagri}$ (kg y^{-1} ha^{-1}) normalized to the catchment area. For transboundary catchments with area outside of Germany, N surplus is normalized to the German part only. On non-agricultural areas, the N surplus is composed of atmospheric N deposition and biological N fixation. On agricultural areas, the N surplus includes additional N inputs, i.e. mineral fertilizer and manure applications, and N outputs from harvesting.

For agricultural land, the N surplus data stem from two data sets: one at state level provided for the period 1950-1998 (Behrendt et al., 2003; which builds on Bach and Frede, 1998; and Behrendt et al., 2000), and one at county level provided for the period 1995-2015 (Häußermann et al., 2019). To create a consistent long-term data set (1950-2015), we harmonized the county and state level data sets based on the overlapping years (1995-1998) and downscaled the state level data to county level for the period 1950-1994. Specifically, we bias-corrected the state level data of Behrendt et al. (2003) using proportions as they commonly underestimated the values provided by Häußermann et al. (2019) for the period 1995-1998. To downscale the bias-corrected state level N surplus (1950-1994) to county level, we used a linear regression between the county and state totals for the period 1995-2015 (data from Häußermann et al., 2019). As data for city-states (Berlin, Bremen and Hamburg) are not provided in the state level dataset, we used the average value from 1995-1998 for the period 1950-1994 under the assumption that the error is acceptable considering the small agricultural areas. The N surplus data comprises values for five of the eleven agricultural land classes in Corine Land Cover (CLC; EEA, 2016a) (non-irrigated arable land,

vineyards, fruit trees and berry plantations, pastures, complex cultivation patterns). The data includes N inputs from applications of fertilizers in mineral and organic forms, from seeds and planting material (county level data only), from N deposition and from biological N fixation, and N outputs from harvested crops. To upscale agricultural N surplus from county level to catchment level, we used the fraction of agricultural area provided by CLC and a scaling factor. Since CLC overestimates agricultural areas compared to the census data at county level (Bach et al., 2006), we scaled the agricultural areas from CLC in each county with the mean ratio between the agricultural area from census data (Häußermann et al., 2019) and the CLC maps (years 2000, 2006 and 2012; median ratio of 1.24 across counties).

For non-agricultural land (CLC classes forest, water bodies, wetlands, grassland) and the remaining agricultural land CLC classes not covered by the N surplus data described above (e.g., permanently irrigated land), we used the atmospheric N deposition data from the Meteorological Synthesizing Centre-West (MSC-W) of the European Monitoring and Evaluation Programme (EMEP; Simpson et al., 2012). The EMEP database uses a chemical transport model to generate a consistent gridded field of Europe-wide wet and dry, and oxidized and reduced atmospheric N depositions (Simpson et al., 2012). The model assimilates varying levels of observational information on different atmospheric chemicals (e.g., Bartnicky and Benedictow, 2017; Bartnicky and Fagerli, 2006). The data was available for the period 1980-1995 with five-year steps, which we linearly interpolated to obtain an annual time series, and with annual steps for the period 1995-2015. For the data before 1980, we assumed constant values from 1980 due to missing information. Deposition on urban sealed surfaces was neglected, since we assume this component is collected by the sewer system and transported to the wastewater treatment plants. We thus assume it is not a diffuse N source but part of the point sources (Section 4.3). In contrast, deposition on urban grassland like public parks was considered. To account for the overestimated area of the five agricultural CLC classes of the agricultural N surplus data (see above), we added the corresponding missing fraction proportionally to the remaining land cover classes. We estimated terrestrial biological N fixation by plants for non-agricultural, vegetated areas using land-use specific rates provided by Cleveland et al. (1999) and Van Meter et al. (2017).

The catchment scale N surplus time series were calculated by intersecting the two N surplus components (N_{agri} and $N_{nonagri}$) with the respective land use and catchment area. As the N surplus data was only

available within Germany; data from transboundary catchments (e.g., main stretch of the Elbe or Rhine Rivers) need to be used cautiously, with higher uncertainty for catchments with a higher fraction of catchment area outside Germany (Section 4.3). Figure 3 shows the resulting N input time series of all catchments. The majority of N input stems from agriculture with a median of 64 % of the total catchment N surplus stemming from N_{agri} across all catchments (averages between 1950 and 2015). The agricultural N surplus (N_{agri}) as well as its fraction per catchment was highest during the 1980s with the median across catchments amounting to 52 kg ha^{-1} y^{-1} and 76 % (average between 1980 and 1989), respectively. The highest mean agricultural N surplus and its fraction per year across all catchments were reached in 1988 with 60.7 kg ha^{-1} y^{-1} and 74 %, while the values were already above 50 kg ha^{-1} y^{-1} and 70 % from 1976 to 1989. For the total N surplus, mean annual values across catchments were above 70 kg ha^{-1} y^{-1} in the same period (1976-1989), while values above 50 kg ha^{-1} y^{-1} were already reached since 1969 and the maximum of 76.7 kg ha^{-1} y^{-1} occurred in 1980.

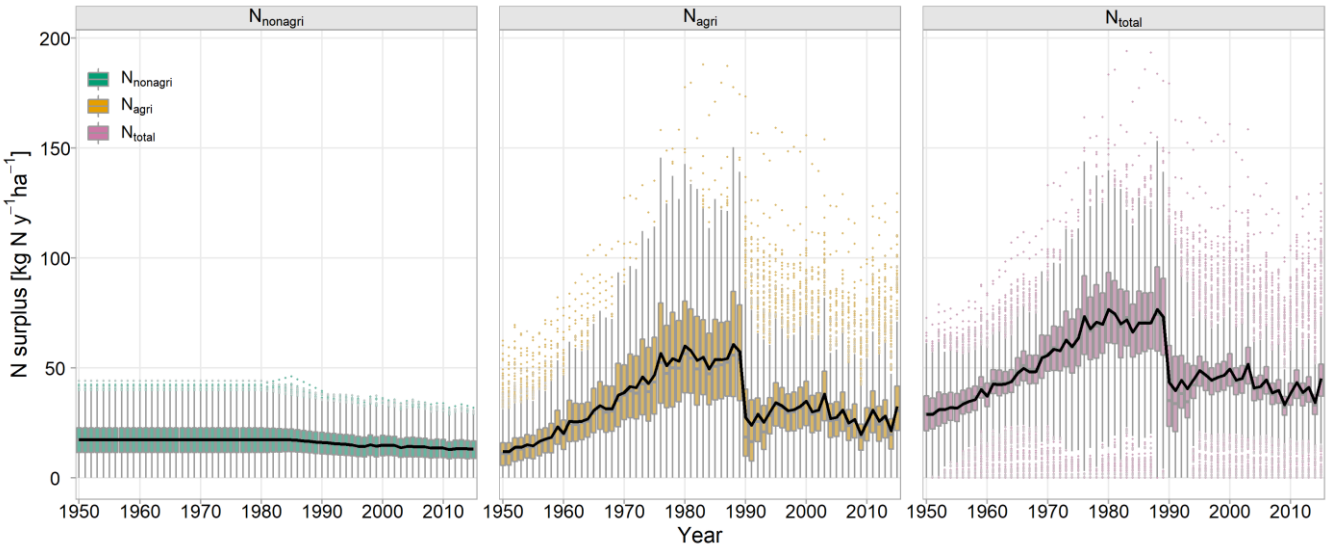

Figure 3: Time series of annual N surplus of all catchments for the different N surplus components: N surplus on non-agricultural areas (left), N surplus on agricultural areas (middle), and total N surplus from both non-agricultural and agricultural areas (right). Boxplots represent the distribution of annual N surplus as averages of the German catchment area across all catchments showing summary statistics (median, quartiles, and quartiles +/- 1.5 times the interquartile range) and individual points outside these ranges. The black lines represent mean annual values for each N surplus component across the catchments.

4 Catchment attributes

The provided catchment attributes characterize the catchments in terms of topography, land cover, nutrient sources, lithology and soils, and hydroclimate. The attributes were chosen with a focus on macronutrient sources and transport in line with the data set. Figure 4 shows the spatial distribution of a set of selected characteristics. All attributes, their variable names, original data sources and methods are listed in Appendix B Table B1 and the data repository (Ebeling et al., 2022a). This repository of catchment attributes is a composite of attributes from two existing repositories (Ebeling and Dupas, 2021; Ebeling, 2021).

4.1 Location and Topography

Catchment size was calculated from the delineated catchment boundaries described in Section 2. Catchment size ranges from 0.9 to 123,012 km² with a median of 171.2 km², a 25th percentile of 53.6 and a 75th percentile of 634.4 km². Additionally, the fraction of the catchment area lying within German borders was calculated (f_AreaGer). Mean and median catchment elevation and topographic slope were extracted from the DEM with 100 m resolution (see also Section 2; EEA, 2013). The 100 m grid of topographic wetness index (TWI) was calculated from the DEM by relating the upstream area (from flow accumulation) to the local slope at each grid cell following Beven and Kirkby (1979). For each catchment, we extracted mean, median and 90th percentile TWI values. The 90th percentile has been shown to be a proxy for the abundance of riparian wetlands in a catchment (Musolff et al., 2018). Drainage density, defined as the length of surface waters per area, closely relates to topography. Drainage density was calculated and provided in two ways: as the catchment average of the gridded drainage density (cell size 0.012 degree) provided in the Hydrologischer Atlas Deutschland (BMU, 2000) and as the river length from EU-hydro river network (EEA, 2016b) within the catchment divided by its area. For the latter, the level of detail was too coarse to yield plausible values for all catchments, which is why values are missing for 27 of the smaller catchments. However, the EU-hydro river network allowed us to derive further stream attributes such as the Strahler order.

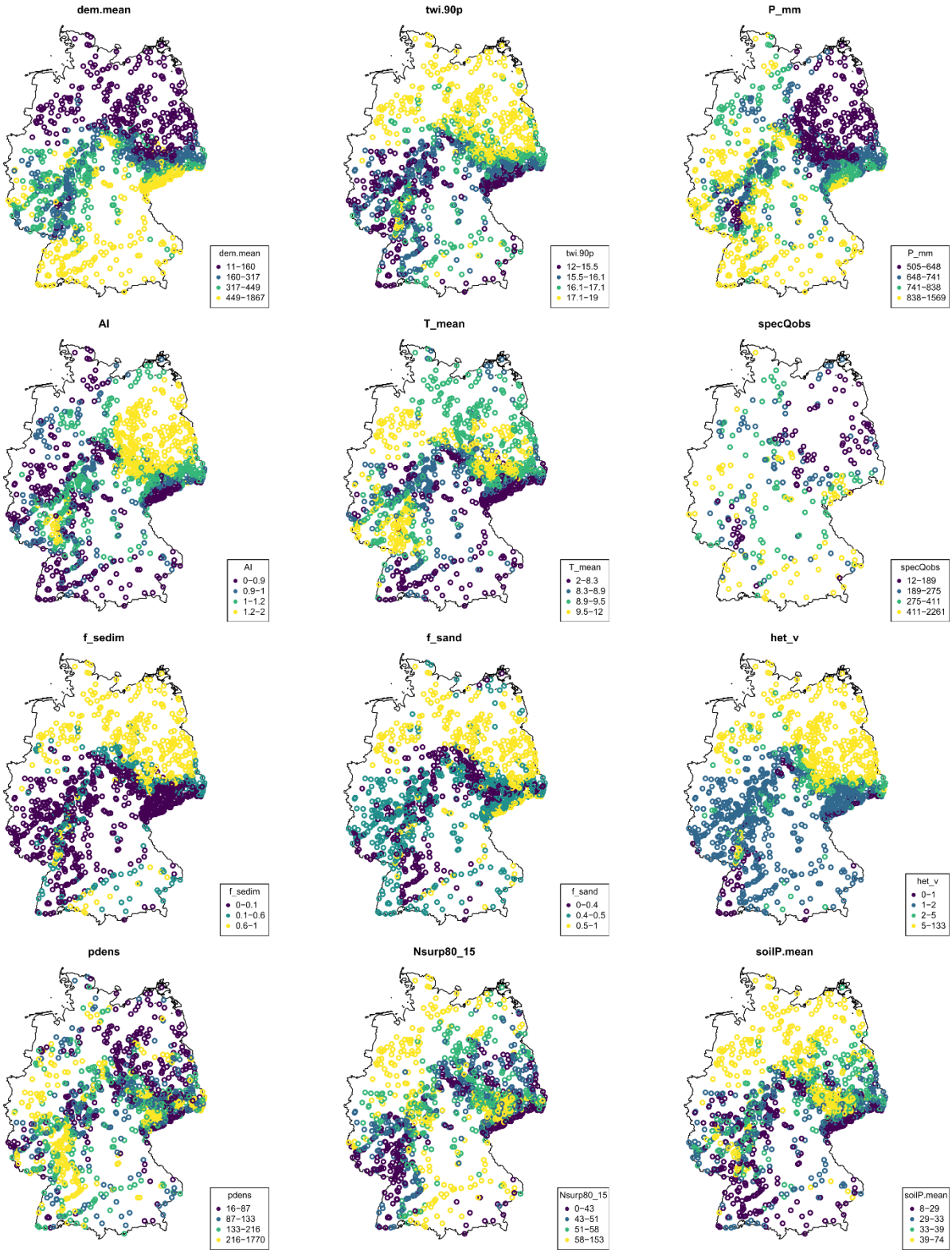
dem.mean
dem.mean
11–160
160–317
317–449
449–1867
twi.90p
twi.90p
12–15.5
15.5–16.1
16.1–17.1
17.1–19
P_mm
P_mm
505–648
648–741
741–838
838–1569
AI
AI
0–0.9
0.9–1
1–1.2
1.2–2
T_mean
T_mean
2–8.3
8.3–8.9
8.9–9.5
9.5–12
specQobs
specQobs
12–189
189–275
275–411
411–2261
f_sedim
f_sedim
0–0.1
0.1–0.6
0.6–1
f_sand
f_sand
0–0.4
0.4–0.5
0.5–1
het_v
het_v
0–1
1–2
2–5
5–133
pdens
pdens
16–87
87–133
133–216
216–1770
Nsurp80_15
Nsurp80_15
0–43
43–51
51–58
58–153
soilP.mean
soilP.mean
8–29
29–33
33–39
39–74

Figure 4: Maps of selected catchment characteristics. Each dot represents one station and the color represents the attribute of the corresponding catchment. Color bars are according to the quartiles of the data distribution of each attribute. Shown attributes: dem.mean – average elevation [m], twi.90p – 90th percentile of the topographic wetness index [-], P_mm – mean annual precipitation [mm y^{-1}], AI – aridity index [-], T_mean – mean air temperature [°C], specQobs – specific annual discharge [mm y^{-1}], f_sedim – fraction of sedimentary aquifer [-], f_sand – fraction of sandy soils [-], het_v – vertical concentration heterogeneity [-], pdens – population density [inhabitants km^{-2}], Nsurp80_15 – mean N surplus from 1980-2015 [kg N ha^{-1} y^{-1}], soilP.mean – phosphorus content in topsoil [mg kg^{-1}]. For more details on the characteristics refer to the text in Section 4 and Table B1.

4.2 Land cover and population density

The fractions of land cover classes were calculated from the level 1 classification of the CLC data set for 2012 (artificial, agricultural, forested land, wetland and surface water cover) (EEA, 2016a). For a finer distinction within these overall classes, fractions of land cover classes were additionally calculated from level 2 data. Note that there can be an overestimation of agricultural areas from these CLC land cover classes when compared to census data as described by Bach et al. (2006) and considered for N surplus time series (Section 3.4). Nevertheless, we expect that the relative distribution of agricultural fractions among the catchments is well captured. The mean catchment population density was calculated from the global data set Gridded Population of the World (CIESIN, 2017) for 2010.

4.3 Nutrient sources

The input from point sources is calculated as the sum of N and P load from wastewater treatment plants (WWTPs) with more than 2000 population equivalents (PE) from the database of the European Environment Agency (EEA, 2017) and data collected from 13 German federal states covering smaller WWTPs (PE < 2000) within Germany (Büttner, 2020). One PE is defined as the organic biodegradable load having a five–day biochemical oxygen demand (BOD5) of 60 g of oxygen per day (EC, 1991a). As a second data source, we calculated catchment averages of the European domestic waste emissions database (Vigiak et al., 2020; Vigiak et al., 2019) for N, P and BOD5 inputs from point sources. The average N, P und BOD5 input per person was estimated using the point source input divided by the number of inhabitants according to the population density. The advantage of these European data is the consistency for an extended, transnational data set, for example, as it is available for German and French catchments (Ebeling and Dupas, 2021).

The net N input from diffuse sources was determined as temporal averages of diffuse N surplus time series (Section 3.4) for different periods, representing the main sampling period with historic inputs

(1980-2015) and the current period (2000-2015). We also calculated averages for the periods before (1971-1990) and after (1990-2015) the EU Nitrogen Directive (EC, 1991b) and the difference between them as a characteristic of net input change. Note that the used N surplus data only cover Germany, but catchments can be transnational. The uncertainty with larger areas outside of Germany increases, for which f_AreaGer can be used as a measure. To estimate source apportionment between point and diffuse N sources, we calculated the fraction of catchment point source N loads (N_WW_frac) from total catchment N input as the sum of catchment point source N loads (N_T_YKM2) and N surplus (here using Nsurp80_15 for the period 1980-2015) on average based on the European wastewater database.

$$\mathrm{N\_WW\_frac} = \mathrm{N\_T\_YKM2}/(\mathrm{N\_T\_YKM2} + \mathrm{Nsurp80\_15})$$

We defined horizontal and vertical source heterogeneity in catchments to quantify the spatial distribution of diffuse nutrient sources with a focus on NO_3-N (Ebeling et al., 2021a). The horizontal source heterogeneity describes the distribution of agricultural land use in a catchment in relation to the stream network. We used the horizontal flow distance of the 100 m DEM (EEA, 2013; Section 2) to the EU-hydro river network (EEA, 2016b) and a highly resolved land use map of 2015 provided by Pflugmacher et al. (2018). We divided the grid into classes of flow distance to stream with 400 m steps. Subsequently, we fitted a linear regression to the share of agricultural source areas in each of the distance classes and the mean distance of the range of each distance class (i.e., 200 m for the class 0-400 m) weighted by the abundance of that specific class. The slope of the resulting linear model het_h characterizes if agricultural source areas tend to be located close to the stream network (het_h<0), equally distributed (het_h=0) or located far away from the stream network (het_h>0). For more details refer to Ebeling et al. (2021a). The vertical source heterogeneity het_v is the ratio of shallow and deep NO_3-N concentration. Shallow NO_3-N concentrations are estimated on a 1 km grid by Knoll et al. (2020) using a ten years average of N surplus and average groundwater recharge. This can be seen as a potential leachate concentration as denitrification in the soil's root zone and horizontal transport are not accounted for. The deep NO_3-N concentrations are estimated on the same grid using a random forest model that is trained on observed concentrations in groundwater (Knoll et al., 2020). The ratio of both was averaged across the catchment to yield het_v reported here. A ratio of one describes a catchment that has a vertical homogeneity in NO_3-N concentrations while a ratio above one describes stronger vertical concentration gradients.

4.4 Lithology and Hydrogeology

To characterize the lithological and the hydrogeological settings of the catchments we used the international hydrogeological map of Europe 1:1,500,00 (BGR & UNESCO, 2014). For the lithological settings, we derived the fraction of area covered by calcareous rocks, calcareous rocks and sediments, magmatic rocks, metamorphic rocks, siliciclastic rocks, siliciclastic rocks and sediments, and sediments (based on lithology data level four). Additionally, we determined the fractions of the more aggregated lithological classes (from lithology level five), i.e. consolidated, partly consolidated, and unconsolidated rocks. We furthermore quantified the areal fraction of aquifer type in the catchment differentiating porous aquifers, fissured hard rock aquifers (including karst), and locally aquiferous or non-aquiferous rocks. Finally, we extracted the catchment median estimate of depth to bedrock from the global map from Shangguan et al. (2017).

4.5 Soil properties

We calculated the fraction of the catchment covered with hydromorphic soils (stagnosols, semi-terrestrial, semi-subhydric, subhydric and peat soils) from the German soil map (1:250,000; BGR, 2018). As this data source only covers Germany, data might not be reliable for transboundary catchments (see also Section 4.3). We also calculated the average fraction of sand, silt and clay averaged across the soil horizons of the top 1 m based on the Harmonized World Soil Database (HWSD; v1.2) available at a 30 arc-second raster database (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012). We first estimated vertically weighted soil textural properties from the original HWSD data provided for two soil layers (upper 30 cm and 30-100 cm). Next, we calculated the areal averages of respective properties considering the boundary (polygon) of each study catchment.

We estimated the porosity of soil profiles (thetaS) based on the pedo-transfer function of Zacharias and Wessolek (2007) and the root-zone plant available water content (WaterRoots) which reflects the difference in water content between the field capacity and permanent wilting point. The field capacity is calculated based on a flux-based estimation approach proposed by Twarakavi et al. (2009) corresponding to a minimum drainage flux of 1 mm d^{-1}. The estimate of the permanent wilting point is derived using the Van Genuchten (1980) model of the matric potential at -1500 kPa and the corresponding model

parameters calculated from pedotransfer functions of Zacharias and Wessolek (2007). Similar to soil textural properties, for each of these soil hydraulic parameters (porosity, field capacity, and permanent wilting point), we calculated areal averages of the vertically weighted estimates for the upper 1 m soil profile for each study catchment. More details on this method using pedo-transfer functions and subsequent aggregations can be found in Livneh et al. (2015). Furthermore, we estimated average catchment soil chemistry of the topsoil (first 20 cm) for year 2009 from the European soil chemistry map, which is based on the LUCAS database (Ballabio et al., 2019). For this, we calculated mean C/N ratio, nitrogen content and phosphorus content from the maps for each catchment.

4.6 Hydroclimatic characteristics

Long-term average hydroclimatic characteristics were derived from the meteorological (Section 3.3) and discharge time series. All climatic characteristics were calculated for a period of 30 years from 1986 to 2015 based on the E-OBS data set from the European Climate Assessment & Dataset (ECA&D) project (v18.0e; Cornes et al., 2018). First, we provide mean annual precipitation, mean annual potential evapotranspiration, mean annual air temperature and the aridity index as the ratio between potential evapotranspiration and precipitation. The variability of precipitation is further characterized by the mean precipitation frequency and depth (Botter et al., 2013) and by two seasonality indices, i.e. the ratio between summer (June-August) and winter (December-February) precipitation (P_SIsw) and the average difference between average daily precipitation within each month and within a year (P_SI).

The hydrologic properties were characterized from stations with observed daily discharge data (Section 3.2) for different time periods according to the available data and study purposes of the original data sets. For current properties, daily discharge data from November 1999 (hydrological year 2000) were used for calculations (309 stations). Additionally, the hydrologic characteristics calculated from daily discharge data starting in 1986 are provided (319 stations), which are possibly more relevant for studies with a long-term perspective. If there was data only before 1986, we used the available time period (four stations). The actual starting and ending dates of the time series finally used for calculations are provided to inform on the exact time periods (StartQobs and EndQobs; Q_StartDate and Q_EndDate respectively, refer to Table B1). Provided average characteristics include mean, median, median summer (May-

October), median winter (November-April), and specific discharge. For the variability of discharge, we provide the coefficient of variation, the base flow index (according to WMO, 2008) and the flashiness index based on flow percentiles (ratio of 5^{th} to 95^{th} percentile) as well as discharge seasonality in terms of the ratio between summer and winter median discharge and the runoff coefficient (discharging fraction of precipitation).

5 Limitations

The presented data set has several limitations. More than half of the stations do not a have co-located gauging station and the ones that do are not homogeneously distributed across Germany. Existing concentration time series would benefit from available discharge data, as this allows the characterization of concentration-discharge relationships as well as the estimation of daily concentration, flow-normalized concentration and flux data for stations with high data availability using the WRTDS method. Generally, modelled discharge from hydrological models such as mHM (Section 3.2.2) or estimated discharge using other (mechanistic or statistical modeling) techniques could serve to extend the data set of joint water quality and water quantity and overcome missing station matches or data gaps. Other limitations are linked to data policies by federal state authorities, which sometimes do not permit publication of raw quality and quantity data. However, we aimed to make a virtue of necessity by providing aggregated data and further ready-to-use metrics of water quality and quantity (e.g., annual median concentrations and monthly median concentrations over the whole time series). Attributes derived from exclusively national data sets, such as N surplus, underlie higher uncertainties in transboundary catchments, as data outside Germany are either not available or not consistent. Additionally, there is uncertainty in the attributes, stemming from the inherent uncertainties in the data sets and the catchment boundaries. However, the provided description and references of the methods and the underlying data sources should enable users to evaluate the reliability of each descriptor in the data set and exclude stations from the analyses if necessary. This also leaves room for further improvements and extensions when new data and knowledge become available. Besides a higher number of water quality stations, longer time series and more co-located discharge data, it would be especially interesting to add time series of nutrient inputs from point

sources and from diffuse P sources, as well as information on tile drainage locations to the catchment attributes. For a better linkage of chemical water quality with ecological research questions, biological water quality variables such as chlorophyll-a concentrations would be highly valuable as well.

6 Data availability

The QUADICA data set presented here is freely available in two online repositories. The water quality and water quantity data described in this manuscript, as well as the time series of meteorological and diffuse nitrogen input can be accessed under https://doi.org/10.4211/hs.0ec5f43e43c349ff818a8d57699c0fe1 (Ebeling et al., 2022b). The catchment characteristics and boundaries are published under https://doi.org/10.4211/hs.88254bd930d1466c85992a7dea6947a4 (Ebeling et al., 2022a). Due to license agreements, the raw concentration and raw discharge data provided by the German federal states are not made public but are deposited in an institutional repository (Musolff et al., 2020), however, metadata of the data and stations are provided under https://doi.org/10.4211/hs.a42addcbd59a466a9aa56472dfef8721 (Musolff, 2020).

7 Conclusions

In this study, we provide a comprehensive homogenized data set with a large spatial and temporal coverage of both water quality and quantity observations along with catchment attributes. Specifically, the data set includes time series of water quality, co-located discharge, hydroclimatic data and diffuse nitrogen inputs, as well as catchment boundaries and more than 100 catchment attributes for 1386 German catchments. The presented QUADICA (water QUAlity, DIscharge and Catchment Attributes for large-sample studies in Germany) data set offers the opportunity to identify spatial and temporal patterns in water quality jointly with water quantity. This allows to formulate and test hypotheses on underlying processes by linking observed responses to the driving forces and catchment attributes. QUADICA also opens up opportunities to calibrate and validate water and solute transport models at the scale of single and multiple catchments as well as at national scale. Consequently, the data set has the potential to

advance our understanding about water quality processes across scales. More specifically, the data can be used to examine various spatio-temporal water quality patterns such as average concentrations, trends, and average seasonality. For stations with high data availability, analyses can be extended to trajectories of seasonality, flow-normalized concentrations and mass fluxes. The patterns can be investigated for the three different macronutrients nitrogen, phosphorus and organic carbon, their species as well as for nutrient ratios. In addition, interactions between the nutrients and their spatio-temporal patterns can be assessed. In the context of comparative large-sample hydrology (e.g., Gupta et al., 2014), the spatio-temporal water quality patterns can be linked to catchment attributes to identify underlying processes. This can, for example, support quantifying the impact of human disturbances on nutrient cycles and their interactions with natural controls. Some studies investigated spatio-temporal patterns and underlying controls in large-sample approaches using parts of the provided data set recently. For example, Ebeling et al. (2021a) assessed average nutrient concentrations and export dynamics, Ebeling et al. (2021b) evaluated long-term trajectories of nitrate seasonality, Ehrhardt et al. (2021b) quantified nitrogen legacies using nitrogen input and export time series, and Yang et al. (2021) modeled the impact of phosphorus inputs on stream network algae growth. These assessments and derived hypotheses can be further explored and extended with the provided data to increase our knowledge on catchment functioning.

Furthermore, the provided data can be merged with other water quality and quantity data sets e.g. to enable assessments across transnational, large scales and with higher variability in catchment attributes. Here, we hope to stimulate other researchers or environmental authorities to provide similar data sets of joint water quality and quantity data to make the wealth of spatiotemporal water quality data available, including long-term data that have been collected in research projects and during regular monitoring activities such as the EU Water Framework Directive (EC, 2000). Therefore, we call for joint efforts to further increase opportunities for catchment scale water quality assessments and modeling activities on regional, transnational and even continental scales.

Appendix A

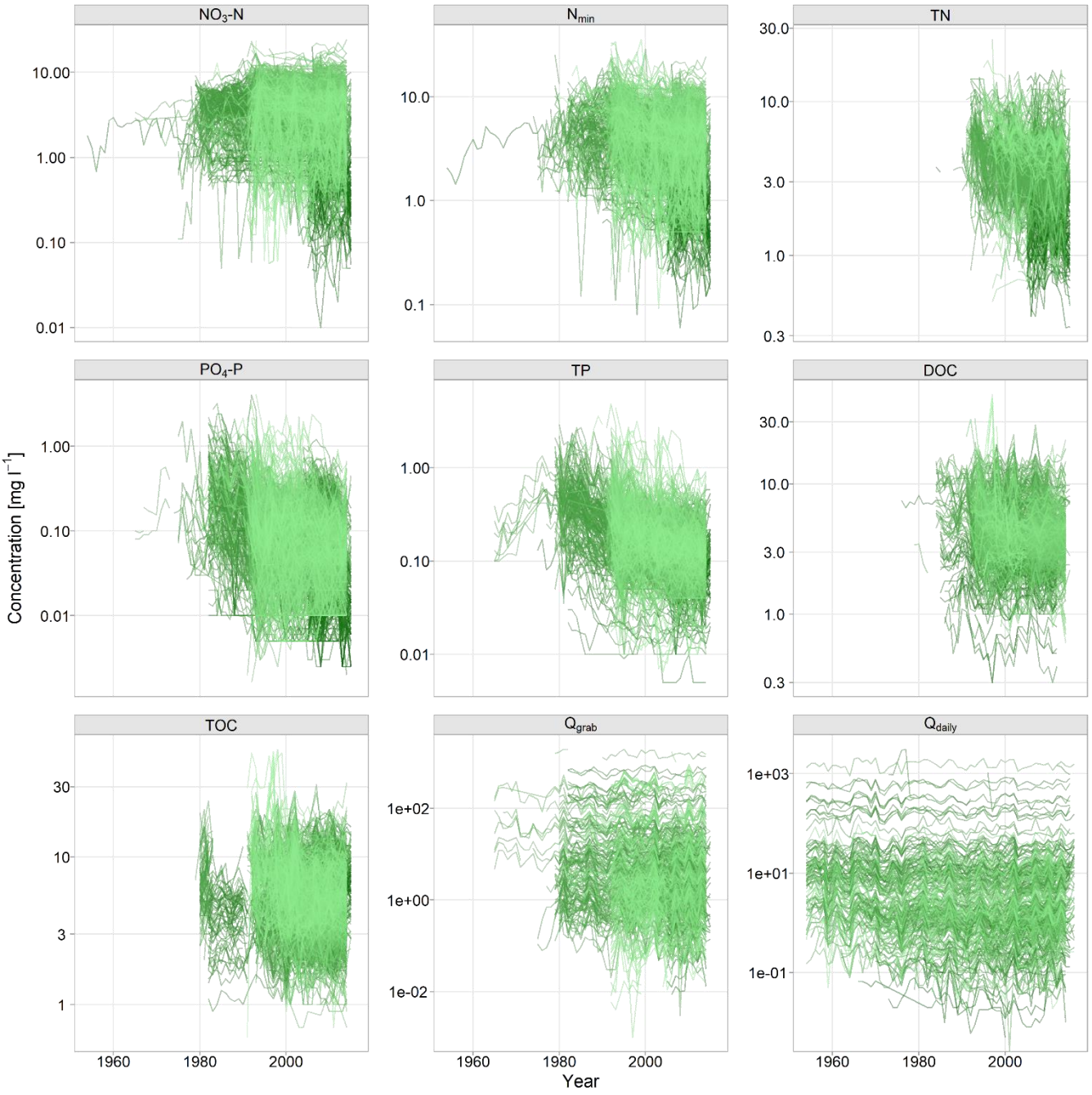

Fig. A1: Time series of annual median concentrations and discharge observed at the 1386 water quality stations during grab sampling as in Table 1, Fig. 1 and described in Section 3.1. Note: For visualization purposes, values before 1954 and values > 40 mg l^{-1} for N species (i.e. five NO_3-N, seven N_{min} and zero TN values) are not shown.

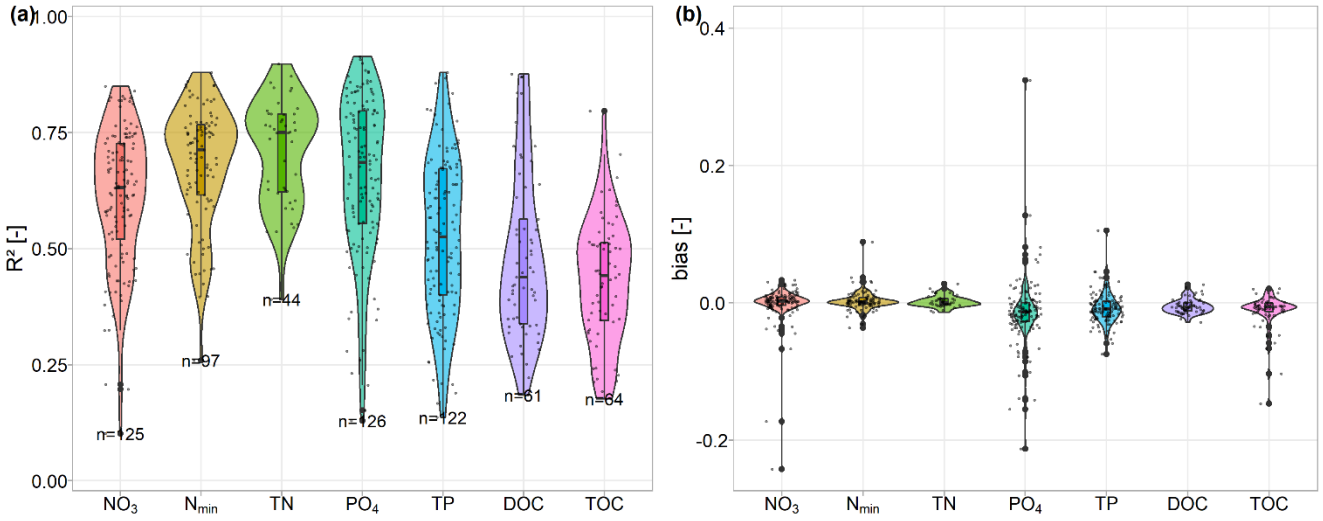

Fig. A2: Distribution of performances of WRTDS-models by compound, coefficient of determination R^2 (a) and bias (b). Boxes highlight the median and quartiles of each distribution; points display performance values of single catchments. Note that one bias value > 0.4 is not shown for TOC for better visibility.

605

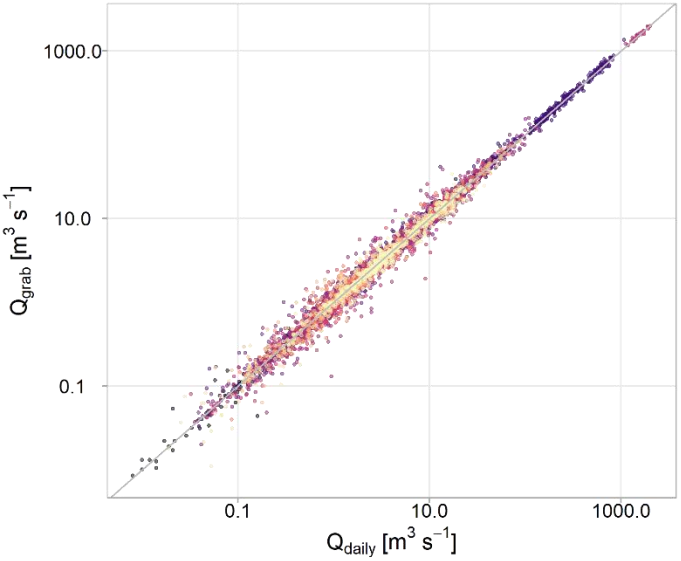

Fig. A3: Comparison of annual medians from continuous daily discharge (Qdaily) and discharge at the dates grab samples were taken (Qgrab). Colors represent different catchments.

Appendix B

Table B1: Catchment attributes, associated methods and original data sources used for calculating the attributes (Ebeling et al., 2022a). This collection of catchment attributes is merged and adapted from existing repositories (Ebeling, 2021; Ebeling and Dupas, 2021) and the related publications (Ebeling et al., 2021b; Ebeling et al., 2021a; Ehrhardt et al., 2021b). For more details see Section 4.

Category	Variable	Unit	Description and method	Data source
General	OBJECTID	-	Unique identifier	
	Station	-	Station name	
	Area_km2	km²	Catchment area	
	f_AreaGer	-	Fraction of catchment area within Germany	
Topography	dem.mean	mamsl	Mean elevation of catchment, from DEM rescaled from 25 to 100 m resolution using average	EEA (2013)
	dem.median	mamsl	Median elevation of catchment, from DEM rescaled from 25 to 100 m resolution using average	EEA (2013)
	slo.mean	°	Mean topographic slope of catchment, from DEM	EEA (2013)
	slo.median	°	Median topographic slope of catchment, from DEM	EEA (2013)
	twi.mean	-	Mean topographic wetness index (TWI, Beven & Kirkby, 1979)	EEA (2013)
	twi.med	-	Median topographic wetness index (TWI, Beven & Kirkby, 1979)	EEA (2013)
	twi.90p	-	90^{th} percentile of the TWI as a proxy for riparian wetlands (following Musolff et al., 2018)	EEA (2013)
	ddhad	km^{-1}	Average drainage density of the catchment. Gridded drainage density is provided as the length of surface waters (rivers and lakes) per area from a 75km² circular area around each cell centerd.	BMU (2000)
	DrainDens	km^{-1}	Average drainage density of the catchment, calculated from EU-Hydro River Network and intersection with Catchment polygons (contains several implausible values (often too small values due to coarser resolution of river network))	EEA (2016b)
Land cover	f_artif	-	Fraction of artificial land cover	EEA (2016a)

	f_agric	-	Fraction of agricultural land cover	EEA (2016a)
	f_forest	-	Fraction of forested land cover	EEA (2016a)
	f_wetl	-	Fraction of wetland cover	EEA (2016a)
	f_water	-	Fraction of surface water cover	EEA (2016a)
	f_urban	-	Fraction of Class 11 Level 2 CORINE Land Cover	EEA (2016a)
	f_industry	-	Fraction of Class 12 Level 2 CORINE Land Cover	EEA (2016a)
	f_mine	-	Fraction of Class 13 Level 2 CORINE Land Cover	EEA (2016a)
	f_urban_veg	-	Fraction of Class 14 Level 2 CORINE Land Cover	EEA (2016a)
	f_arable	-	Fraction of Class 21 Level 2 CORINE Land Cover	EEA (2016a)
	f_agri_perm	-	Fraction of Class 22 Level 2 CORINE Land Cover	EEA (2016a)
	f_pastures	-	Fraction of Class 23 Level 2 CORINE Land Cover	EEA (2016a)
	f_agri_hetero	-	Fraction of Class 24 Level 2 CORINE Land Cover	EEA (2016a)
	f_fores	-	Fraction of Class 31 Level 2 CORINE Land Cover	EEA (2016a)
	f_scrub	-	Fraction of Class 32 Level 2 CORINE Land Cover	EEA (2016a)
	f_open	-	Fraction of Class 33 Level 2 CORINE Land Cover	EEA (2016a)
	pdens	inhabitants km^{-2}	Mean population density	*CIESIN (2017)*
Nutrient sources	Nsurp00_15	kg N ha^{-1} y^{-1}	Mean nitrogen (N) surplus per catchment during sampling period (2000-2015) including the N surplus on agricultural land and atmospheric N deposition and biological N fixation on non-agricultural areas. Details on the N surplus data is given in Section 3.4.	Bach et al. (2006); Bach and Frede (1998); Bartnicky and Benedictow (2017); Bartnicky and Fagerli (2006); Behrendt et al. (1999); Cleveland et al. (1999); Häußermann et al. (2019); Van Meter et al. (2017)
	Nsurp91_15	kg N ha^{-1} y^{-1}	Mean N surplus per catchment from 1991 to 2015 (after Nitrogen Directive was introduced)	See Nsurp00_15

Nsurp80_15	kg N ha^{-1} y^{-1}	Mean N surplus per catchment from 1980 to 2015 (main sampling period)	See Nsurp00_15
Nsurp71_90	kg N ha^{-1} y^{-1}	Mean N surplus per catchment from 1971 to 1990 (historic (legacy) inputs)	See Nsurp00_15
dNsurp71_91	kg N ha^{-1} y^{-1}	Change in mean N surplus between the periods 1971-1990 and 1991-2015, i.e. dNsurp71_91=Nsurp71_90 - Nsurp91_15	See Nsurp00_15
N_WW	kg N ha^{-1} y^{-1}	Sum of N input from point sources including waste water treatment plants (WWTP) > 2000 person equivalents from the database of the European Environment Agency covering areas beyond Germany and data collected from 13 Federal German States covering smaller WWTP within Germany	EEA (2017); Yang et al. (2019)
P_WW	kg P ha^{-1} y^{-1}	Sum of P input from WWTP analogous to N_WW	EEA (2017); Yang et al. (2019)
N_T_YKM2	t N km^{-2} y^{-1}	Mean N input from point sources summing all N emission values provided in the EU domestic waste emissions data base	Vigiak et al. (2019); Vigiak et al. (2020)
P_T_YKM2	t P km^{-2} y^{-1}	Mean P input from point sources summing all P emission values provided in the EU domestic waste emissions data base	Vigiak et al. (2019); Vigiak et al. (2020)
BOD_T_YKM2	t O km^{-2} y^{-1}	Mean five-days biochemical oxygen demand (BOD) input from point sources summing all BOD emission values provided in the EU domestic waste emissions data base	Vigiak et al. (2019); Vigiak et al. (2020)
N_T_YEW	t N inh^{-1} y^{-1}	Calculated N input per person (from EU domestic waste emissions data base) N_T_YEW =N_T_YKM2 / nEW * Area_km2	Vigiak et al. (2019); Vigiak et al. (2020)
P_T_YEW	t P inh^{-1} y^{-1}	Calculated P input per person (from EU domestic waste emissions data base) P_T_YEW =P_T_YKM2 / nEW * Area_km2	Vigiak et al. (2019); Vigiak et al. (2020)
nEW	-	Calculated number of inhabitants, nEW=pdens * Area_km2	*CIESIN (2017)*
n_UWWTP	-	Number of point sources from European data base (UWWTP data base)	EEA (2017)
N_WW_frac	-	Fraction of point source loads from total N input loads N_WW_frac = N_T_YKM2 / (N_T_YKM2 + Nsurp80_15)	
f_sarea	-	Fraction of source area in the catchment. Source areas were defined as seasonal, perennial cropland and grassland land cover classes using	Source areas based on Pflugmacher et al. (2018)

			a highly resolved land use map (Pflugmacher et al., 2018)	
	het_h	m^{-1}	Slope of relative frequency of source areas in classes of flow distances to stream as a proxy for horizontal source heterogeneity. For details refer to Ebeling, Kumar, et al. (2021)	Source areas based on Pflugmacher et al. (2018)
	R2_het_h	-	Coefficient of determination of horizontal source heterogeneity het_h	
	sdist_mean	m	Mean lateral flow distance of source areas to stream. For details refer to Ebeling, Kumar, et al. (2021)	Source areas based on Pflugmacher et al. (2018)
	het_v	-	Mean ratio between potential seepage and groundwater NO_3-N concentrations as proxy for vertical concentration heterogeneity. For details refer to Ebeling, Kumar, et al. (2021)	Knoll et al. (2020)
Lithology and soils	f_calc	-	Fraction of calcareous rocks (Lithology level 4)	BGR & UNESCO (eds.) (2014)
	f_calc_sed	-	Fraction of calcareous rocks and sediments (Lithology level 4, coarse and fine sediments aggregated)	BGR & UNESCO (eds.) (2014)
	f_magma	-	Fraction of magmatic rocks (Lithology level 4)	BGR & UNESCO (eds.) (2014)
	f_metam	-	Fraction of metamorphic rocks (Lithology level 4)	BGR & UNESCO (eds.) (2014)
	f_sedim	-	Fraction of sedimentary aquifer (Lithology level 4, coarse and fine sediments aggregated)	BGR & UNESCO (eds.) (2014)
	f_silic	-	Fraction of siliciclastic rocks (Lithology level 4)	BGR & UNESCO (eds.) (2014)
	f_sili_sed	-	Fraction of siliciclastic rocks and sediments (Lithology level 4, coarse and fine sediments aggregated)	BGR & UNESCO (eds.) (2014)
	f_consol	-	Fraction of consolidated rocks (Lithology Level 5)	BGR & UNESCO (eds.) (2014)
	f_part_consol	-	Fraction of partly consolidated rocks (Lithology Level 5)	BGR & UNESCO (eds.) (2014)
	f_unconsol	-	Fraction of unconsolidated rocks (Lithology Level 5)	BGR & UNESCO (eds.) (2014)
	f_porous	-	Fraction of porous aquifer (code 1 and 2 of aquifer type)	BGR & UNESCO (eds.) (2014)
	f_porous1	-	Fraction of porous aquifer (code 1 of aquifer type)	BGR & UNESCO (eds.) (2014)

f_porous2	-	Fraction of porous aquifer (code 2 of aquifer type)	BGR & UNESCO (eds.) (2014)
f_fissured	-	Fraction of fissured aquifer (code 3 and 4 of aquifer type)	BGR & UNESCO (eds.) (2014)
f_fiss1	-	Fraction of fissured aquifer (code 3 of aquifer type)	BGR & UNESCO (eds.) (2014)
f_fiss2	-	Fraction of fissured aquifer (code 4 of aquifer type)	BGR & UNESCO (eds.) (2014)
f_hard	-	Fraction of locally aquiferous and non-aquiferous aquifer (code 5 and 6 of aquifer type)	BGR & UNESCO (eds.) (2014)
f_hard1	-	Fraction of locally aquiferous rocks (code 5 of aquifer type)	BGR & UNESCO (eds.) (2014)
f_hard2	-	Fraction of non-aquiferous rocks (code 6 of aquifer type)	BGR & UNESCO (eds.) (2014)
f_inwater		Fraction of inland water (code 200 of aquifer type)	BGR & UNESCO (eds.) (2014)
f_ice		Fraction of snow or ice field (code 300 of aquifer type)	BGR & UNESCO (eds.) (2014)
dtb.median	cm	Median depth to bedrock in the catchment	Shangguan et al. (2017)
f_gwsoils	-	Fraction of water-impacted soils in the catchment (from soil map 1:250,000), including stagnosols, semi-terrestrial, semi-subhydric, subhydric and moor soils	BGR (2018)
f_sand f_silt f_clay	-	Mean fraction of sand in soil horizons of the top 100 cm Mean fraction of silt in soil horizons of the top 100 cm Mean fraction of clay in soil horizons of the top 100 cm	FAO/IIASA/ISRIC/ISSCAS/JRC (2012)
f_clay_agri		Mean fraction of clay in soil horizons of the top 100 cm on agricultural land use (Class 2 Level 1 CORINE; see f_clay and f_agric)	FAO/IIASA/ISRIC/ISSCAS/JRC (2012), EEA (2016a)
WaterRoots	mm	Mean available water content in the root zone from pedo-transfer functions	Livneh et al. (2015); Samaniego et al. (2010); Zink et al. (2017)
thetaS	-	Mean porosity in catchment from pedo-transfer functions	Livneh et al. (2015); Samaniego et al. (2010); Zink et al. (2017)
soilN.mean	g kg^{-1}	Mean top soil N in catchment	Ballabio et al. (2019)
soilP.mean	mg kg^{-1}	Mean top soil P in catchment	Ballabio et al. (2019)
soilCN.mean	-	Mean top soil C/N ratio in catchment	Ballabio et al. (2019)

Hydrology	StartQobs	YYYY-MM-DD	Starting date of Q time series used for calculating hydrological indices (1999-11-01 or start of time series)	
	EndQobs	YYYY-MM-DD	End date of Q time series used for calculating hydrological indices	
	meanQobs	$m^3 s^{-1}$	Mean discharge (period from StartQobs to EndQobs)	Musolff (2020); Musolff et al. (2020)
	medQobs	$m^3 s^{-1}$	Median discharge (period from StartQobs to EndQobs)	Musolff (2020); Musolff et al. (2020)
	specQobs	$mm\ y^{-1}$	Mean annual specific discharge (period from StartQobs to EndQobs)	Musolff (2020); Musolff et al. (2020)
	CVQobs	-	Coefficient of variation of time series of daily Q (period from StartQobs to EndQobs)	Musolff (2020); Musolff et al. (2020)
	medSuQobs	$m^3 s^{-1}$	Median summer discharge (months May-October) (period from StartQobs to EndQobs)	Musolff (2020); Musolff et al. (2020)
	medWiQobs	$m^3 s^{-1}$	Median winter discharge (months November-April) (period from StartQobs to EndQobs)	Musolff (2020); Musolff et al. (2020)
	seasRQobs	-	Seasonality index of Q, as ratio between median summer and median winter Q (period from StartQobs to EndQobs)	Musolff (2020); Musolff et al. (2020)
	BFIQobs	-	Base flow index calculated according to WMO [2008] with *lfstat* package (version 0.9.4) in R (period from StartQobs to EndQobs)	Musolff (2020); Musolff et al. (2020)
	flashQobs	-	Flashiness index of Q as the ratio between 5 % percentile and 95 % percentile of Q time series (period from StartQobs to EndQobs)	Musolff (2020); Musolff et al. (2020)
	RCQobs	-	Runoff coefficient (fraction of mean annual precipitation discharging as specific discharge, specQobs/P_mm) (period from StartQobs to EndQobs)	Musolff (2020); Musolff et al. (2020)
	Q_StartDate	YYYY-MM-DD	Starting date of Q time series used for calculating hydrological indices (from 1986, if possible and at least 3 years of data, in few cases only earlier data was available)	
	Q_EndDate	YYYY-MM-DD	End date of Q time series used for calculating hydrological indices (as available)	
	Q_mean	$m^3 s^{-1}$	Mean discharge (data for the period Q_StartDate-Q_EndDate)	Musolff (2020); Musolff et al. (2020)
	Q_median	$m^3 s^{-1}$	Median discharge (data for the period Q_StartDate-Q_EndDate)	Musolff (2020); Musolff et al. (2020)

	Q_spec	mm y^{-1}	Mean annual specific discharge (data for the period Q_StartDate-Q_EndDate)	Musolff (2020); Musolff et al. (2020)
	Q_CVQ	-	Coefficient of variation of time series of daily Q (data for the period Q_StartDate-Q_EndDate)	Musolff (2020); Musolff et al. (2020)
	Q_medSum	m^3 s^{-1}	Median summer discharge (months May-October) (data for the period Q_StartDate-Q_EndDate)	Musolff (2020); Musolff et al. (2020)
	Q_medWin	m^3 s^{-1}	Median winter discharge (months November-April) (data for the period Q_StartDate-Q_EndDate)	Musolff (2020); Musolff et al. (2020)
	Q_Sum2Win	-	Seasonality index of Q, as ratio between median summer and median winter Q (data for the period Q_StartDate-Q_EndDate)	Musolff (2020); Musolff et al. (2020)
	BFI	-	Base flow index calculated according to WMO [2008] with *lfstat* package (version 0.9.4) in R (data for the period Q_StartDate-Q_EndDate)	Musolff (2020); Musolff et al. (2020)
	flashi	-	Flashiness index of Q as the ratio between 5 % percentile and 95 % percentile of Q time series (data for the period Q_StartDate-Q_EndDate)	Musolff (2020); Musolff et al. (2020)
Climate	P_mm	mm y^{-1}	Mean annual precipitation (period 1986-2015)	Cornes et al. (2018)
	P_SIsw	-	Seasonality of precipitation as the ratio between mean summer (Jun-Aug) and winter (Dec-Feb) precipitation (period 1986-2015)	Cornes et al. (2018)
	P_SI	-	Seasonality index of precipitation as the mean difference between monthly averages of daily precipitation and year average of daily precipitation (period 1986-2015)	Cornes et al. (2018)
	P_lambda	d^{-1}	Mean precipitation frequency λ as used by Botter et al. (2013) with rain days for precipitation above 1 mm (period 1986-2015)	Cornes et al. (2018)
	P_alpha	mm d^{-1}	Mean precipitation depth as used by Botter et al. (2013) with rain days for precipitation above 1 mm (period 1986-2015)	
	PET_mm	mm y^{-1}	Mean annual potential evapotranspiration (period 1986-2015)	Cornes et al. (2018)
	AI	-	Aridity index as AI=PET_mm/P_mm (period 1986-2015)	Cornes et al. (2018)
	T_mean	°C	Mean annual air temperature (period 1986-2015)	Cornes et al. (2018)

615

Author contributions. PE carried out the study, processed and curated the data and created the figures and tables. PE, AM and RK conceptualized and designed the study following initial ideas and acquired funding from AM and SA. Several authors contributed to the data collection and processing: RK provided the gridded meteorological time series, simulated discharge data and atmospheric deposition data, MW provided time series of N surplus data for the catchments, OB collected the point source data for Germany. PE produced the original draft of the manuscript with contributions of AM and RK. All authors contributed to the reviewing and editing of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We gratefully thank all data collectors, processors and providers including the federal state environmental agencies and all contributors to this data set. We especially thank Thomas Grau, Teresa Nitz, Joni Dehaspe, Stefanie Breese, Sophie Ehrhardt for their contributions for data processing. The authors thank two anonymous reviewers for their valuable comments. We gratefully acknowledge Martin Bach and Uwe Häußermann, Justus-Liebig-University of Giessen, for the provision of the two data sets on the agricultural N surplus data for Germany. We acknowledge the E-OBS data set from the EU-FP6 project UERRA (http://www.uerra.eu) and the Copernicus Climate Change Service, and the data providers in the ECA&D project (https://www.ecad.eu). The authors additionally acknowledge several organizations for the data products used here, including the BfG, BGR, SGD, EEA, FAO, IIASA, ISRIC, ISSCAS, and JRC. The authors thank for the funding by the Deutsche Forschungsgemeinschaft - DFG (Dominant controls on catchment water quality dynamics – a Germany-wide analysis using data-driven models, DFG 392886738).

References

Addor, N., Newman, A. J., Mizukami, N., and Clark, M. P.: The CAMELS data set: catchment attributes and meteorology for large-sample studies, Hydrology and Earth System Sciences, 21, 5293-5313, https://doi.org/10.5194/hess-21-5293-2017, 2017.

Addor, N., Do, H. X., Alvarez-Garreton, C., Coxon, G., Fowler, K., and Mendoza, P. A.: Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges, Hydrological Sciences Journal, 65, 712-725, https://doi.org/10.1080/02626667.2019.1683182, 2020.
Alvarez-Garreton, C., Mendoza, P. A., Boisier, J. P., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J., and Ayala, A.: The CAMELS-CL dataset: catchment attributes and meteorology for large sample studies – Chile dataset, Hydrology and Earth System Sciences, 22, 5817-5846, https://doi.org/10.5194/hess-22-5817-2018, 2018.
Bach, M. and Frede, H.-G.: Agricultural nitrogen, phosphorus and potassium balances in Germany —Methodology and trends 1970 to 1995, Zeitschrift für Pflanzenernährung und Bodenkunde, 161, 385-393, https://doi.org/10.1002/jpln.1998.3581610406, 1998.
Bach, M., Breuer, L., Frede, H. G., Huisman, J. A., Otte, A., and Waldhardt, R.: Accuracy and congruency of three different digital land-use maps, Landscape and Urban Planning, 78, 289-299, https://doi.org/10.1016/j.landurbplan.2005.09.004, 2006.
Ballabio, C., Lugato, E., Fernández-Ugalde, O., Orgiazzi, A., Jones, A., Borrelli, P., Montanarella, L., and Panagos, P.: Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression, Geoderma, 355, 113912, https://doi.org/10.1016/j.geoderma.2019.113912, 2019.
Bartnicky, J. and Benedictow, A.: Atmospheric Deposition of Nitrogen to OSPAR Convention waters in the period 1995–2014, EMEP/MSC-W Technical Report, 1/2007, Meteorological Synthesizing Centre-West (MSC-W), Norwegian Meteorological Institute, Oslo, 2017.
Bartnicky, J. and Fagerli, H.: Atmospheric Nitrogen in the OSPAR Convention Area in the Period 1990–2004. Summary Report for the OSPAR Convention., EMEP/MSC-W Technical Report, 4/2006, Meteorological Synthesizing Centre-West (MSC-W) of EMEP, Oslo, 2006.
Basu, N. B., Destouni, G., Jawitz, J. W., Thompson, S. E., Loukinova, N. V., Darracq, A., Zanardo, S., Yaeger, M., Sivapalan, M., Rinaldo, A., and Rao, P. S. C.: Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity, Geophysical Research Letters, 37, https://doi.org/10.1029/2010gl045168, 2010.
Behrendt, H., Huber, P., Opitz, D., Schmoll, O., Scholz, G., and Uebe, R.: Nutrient emissions into river basins of Germany, UBA-Texte, 23/00, 2000.
Behrendt, H., Bach, M., Kunkel, R., Opitz, D., Pagenkopf, W.-G., Scholz, G., and Wendland, F.: Nutrient Emissions into River Basins of Germany on the Basis of a Harmonized Procedure UBA-Texte, 82/03, 2003.
Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant, Hydrological Sciences Bulletin, 24, 43-69, https://doi.org/10.1080/02626667909491834, 1979.
BGR: Bodenübersichtskarte der Bundesrepublik Deutschland 1:250.000 (BUEK250). Soil map of Germany 1:250,000, Federal Institute for Geosciences and Natural Resources (BGR) [dataset], 2018.
BGR & UNESCO (eds.): International Hydrogeological Map of Europe 1 : 1,500,000 (IHME1500). Digital map data v1.1. [dataset], http://www.bgr.bund.de/ihme1500/, 2014.
BMU, Bundesministerium Für Umwelt, N. u. R. (Ed.): Hydrologischer Atlas von Deutschland, Datenquelle: Hydrologischer Atlas von Deutschland/BfG, 2000, Bonn/Berlin2000.
Botter, G., Basso, S., Rodriguez-Iturbe, I., and Rinaldo, A.: Resilience of river flow regimes, Proc Natl Acad Sci U S A, 110, 12925-12930, https://doi.org/10.1073/pnas.1311920110, 2013.
Büttner, O.: DE-WWTP - data collection of wastewater treatment plants of Germany (status 2015, metadata), HydroShare [dataset], https://doi.org/10.4211/hs.712c1df62aca4ef29688242eeab7940c, 2020.
Center for International Earth Science Information Network - CIESIN - Columbia University: Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 10, NASA Socioeconomic Data and Applications Center (SEDAC) [dataset], https://doi.org/10.7927/H4DZ068D, 2017.
Chagas, V. B. P., Chaffe, P. L. B., Addor, N., Fan, F. M., Fleischmann, A. S., Paiva, R. C. D., and Siqueira, V. A.: CAMELS-BR: hydrometeorological time series and landscape attributes for 897 catchments in Brazil, Earth Syst. Sci. Data, 12, 2075-2096, https://doi.org/10.5194/essd-12-2075-2020, 2020.
Chen, D., Shen, H., Hu, M., Wang, J., Zhang, Y., and Dahlgren, R. A.: Chapter Five - Legacy Nutrient Dynamics at the Watershed Scale: Principles, Modeling, and Implications, in: Advances in Agronomy, edited by: Sparks, D. L., Academic Press, 237-313, https://doi.org/10.1016/bs.agron.2018.01.005, 2018.
Cleveland, C. C., Townsend, A. R., Schimel, D. S., Fisher, H., Howarth, R. W., Hedin, L. O., Perakis, S. S., Latty, E. F., Von Fischer, J. C., Elseroad, A., and Wasson, M. F.: Global patterns of terrestrial biological nitrogen (N2) fixation in natural ecosystems, Global Biogeochemical Cycles, 13, 623-645, https://doi.org/10.1029/1999GB900014, 1999.
Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M., and Jones, P. D.: An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets, Journal of Geophysical Research: Atmospheres, 123, 9391-9409, https://doi.org/10.1029/2017jd028200, 2018.
Coxon, G., Addor, N., Bloomfield, J. P., Freer, J., Fry, M., Hannaford, J., Howden, N. J. K., Lane, R., Lewis, M., Robinson, E. L., Wagener, T., and Woods, R.: CAMELS-GB: hydrometeorological time series and landscape attributes for 671 catchments in Great Britain, Earth Syst. Sci. Data, 12, 2459-2483, https://doi.org/10.5194/essd-12-2459-2020, 2020.
De Jager, A. and Vogt, J.: Rivers and Catchments of Europe - Catchment Characterisation Model (CCM) (2.1), European Commission, Joint Research Centre (JRC) [dataset], 2007.

Do, H. X., Westra, S., and Leonard, M.: A global-scale investigation of trends in annual maximum streamflow, Journal of Hydrology, 552, 28-43, https://doi.org/10.1016/j.jhydrol.2017.06.015, 2017.
E-OBS: (v18.0) [dataset], 2018.
Ebeling, P.: CCDB - catchment characteristics data base Germany, HydroShare [dataset], https://doi.org/10.4211/hs.0fc1b5b1be4a475aacfd9545e72e6839, 2021.
Ebeling, P. and Dupas, R.: CCDB - catchment characteristics data base France and Germany, HydroShare [dataset], https://doi.org/10.4211/hs.c7d4df3ba74647f0aa83ae92be2e294b, 2021.
Ebeling, P., Kumar, R., and Musolff, A.: CCDB - catchment characteristics data base Germany, HydroShare [dataset], https://doi.org/10.4211/hs.88254bd930d1466c85992a7dea6947a4, 2022a.
Ebeling, P., Kumar, R., Weber, M., and Musolff, A.: QUADICA - water quality, discharge and catchment attributes for large-sample studies in Germany, HydroShare [dataset], https://doi.org/10.4211/hs.0ec5f43e43c349ff818a8d57699c0fe1, 2022b.
Ebeling, P., Kumar, R., Weber, M., Knoll, L., Fleckenstein, J. H., and Musolff, A.: Archetypes and Controls of Riverine Nutrient Export Across German Catchments, Water Resources Research, 57, e2020WR028134, https://doi.org/10.1029/2020WR028134, 2021a.
Ebeling, P., Dupas, R., Abbott, B., Kumar, R., Ehrhardt, S., Fleckenstein, J. H., and Musolff, A.: Long-Term Nitrate Trajectories Vary by Season in Western European Catchments, Global Biogeochemical Cycles, 35, e2021GB007050, https://doi.org/10.1029/2021GB007050, 2021b.
EC: Council Directive 91/271/EEC of 21 May 1991 concerning urban waste water treatment, Official Journal of the European Communities, 1991a.
EC: Council Directive 91/676/EEC of 12 December 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources, Official Journal of the European Communities, 1991b.
EC: Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy, Official Journal of the European Communities, L 327, 1 - 73, 2000.
EEA: DEM over Europe from the GMES RDA project (EU-DEM, resolution 25m) - version 1, Oct. 2013, European Environmental Agency [dataset], 2013.
EEA: CORINE Land Cover 2012 v18.5, European Environmental Agency [dataset], 2016a.
EEA: EU-Hydro River Network, European Environmental Agency. Copernicus Land Monitoring Service [dataset], 2016b.
EEA: Waterbase - UWWTD: Urban Waste Water Treatment Directive – reported data (6) [dataset], 2017.
EEA: Waterbase - Water Quality ICM [dataset], 2020.
Ehrhardt, S., Ebeling, P., and Dupas, R.: Exported french water quality and quantity data, HydroShare [dataset], http://www.hydroshare.org/resource/d8c43e1e8a5a4872bc0b75a45f350f7a 2021a.
Ehrhardt, S., Ebeling, P., Dupas, R., Kumar, R., Fleckenstein, J. H., and Musolff, A.: Nitrate Transport and Retention in Western European Catchments Are Shaped by Hydroclimate and Subsurface Properties, Water Resources Research, 57, e2020WR029469, https://doi.org/10.1029/2020WR029469, 2021b.
FAO/IIASA/ISRIC/ISSCAS/JRC: Harmonized World Soil Database (version 1.2), FAO, Rome, Italy and IIASA, Laxenburg, Austria [dataset], 2012.
Fowler, K. J. A., Acharya, S. C., Addor, N., Chou, C., and Peel, M. C.: CAMELS-AUS: hydrometeorological time series and landscape attributes for 222 catchments in Australia, Earth Syst. Sci. Data, 13, 3847-3867, https://doi.org/10.5194/essd-13-3847-2021, 2021.
Gnann, S. J., Howden, N. J. K., and Woods, R. A.: Hydrological signatures describing the translation of climate seasonality into streamflow seasonality, Hydrol. Earth Syst. Sci., 24, 561-580, https://doi.org/10.5194/hess-24-561-2020, 2020.
Godsey, S. E., Hartmann, J., and Kirchner, J. W.: Catchment chemostasis revisited: Water quality responds differently to variations in weather and climate, Hydrological Processes, 33, 3056-3069, https://doi.org/10.1002/hyp.13554, 2019.
Godsey, S. E., Kirchner, J. W., and Clow, D. W.: Concentration-discharge relationships reflect chemostatic characteristics of US catchments, Hydrological Processes, 23, 1844-1864, https://doi.org/10.1002/hyp.7315, 2009.
Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., and Andréassian, V.: Large-sample hydrology: a need to balance depth with breadth, Hydrol. Earth Syst. Sci., 18, 463-477, https://doi.org/10.5194/hess-18-463-2014, 2014.
Hargreaves, G. H. and Samani, Z. A.: Reference Crop Evapotranspiration from Temperature, Applied Engineering in Agriculture, 1, 96-99, https://doi.org/10.13031/2013.26773, 1985.
Hartmann, J., Lauerwald, R., and Moosdorf, N.: A Brief Overview of the GLObal RIver Chemistry Database, GLORICH, Procedia Earth and Planetary Science, 10, 23-27, https://doi.org/10.1016/j.proeps.2014.08.005, 2014.
Häußermann, U., Bach, M., Klement, L., and Breuer, L.: Stickstoff-Flächenbilanzen für Deutschland mit Regionalgliederung Bundesländer und Kreise – Jahre 1995 bis 2017. Methodik, Ergebnisse und Minderungsmaßnahmen, Texte, 131/2019, 2019.
Hirsch, R. M. and De Cicco, L. A.: User Guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R Packages for Hydrologic Data, U.S. Geological Survey Techniques and Methods book 4, chap. A10, 93, https://dx.doi.org/10.3133/tm4A10, 2015.
Hirsch, R. M., Moyer, D. L., and Archfield, S. A.: Weighted Regressions on Time, Discharge, and Season (WRTDS), with an Application to Chesapeake Bay River Inputs, JAWRA Journal of the American Water Resources Association, 46, 857-880, https://doi.org/10.1111/j.1752-1688.2010.00482.x, 2010.

Kaushal, S. S., Gold, A. J., Bernal, S., and Tank, J. L.: Diverse water quality responses to extreme climate events: an introduction, Biogeochemistry, 141, 273-279, https://doi.org/10.1007/s10533-018-0527-x, 2018.

Kingston, D. G., Massei, N., Dieppois, B., Hannah, D. M., Hartmann, A., Lavers, D. A., and Vidal, J. P.: Moving beyond the catchment scale: Value and opportunities in large-scale hydrology to understand our changing world, Hydrological Processes, 34, 2292-2298, https://doi.org/10.1002/hyp.13729, 2020.

Klingler, C., Schulz, K., and Herrnegger, M.: LamaH-CE: LArge-SaMple DAta for Hydrology and Environmental Sciences for Central Europe, Earth Syst. Sci. Data, 13, 4529-4565, https://doi.org/10.5194/essd-13-4529-2021, 2021.

Knoll, L., Breuer, L., and Bach, M.: Nation-wide estimation of groundwater redox conditions and nitrate concentrations through machine learning, Environmental Research Letters, 15, 064004, https://doi.org/10.1088/1748-9326/ab7d5c, 2020.

Kuentz, A., Arheimer, B., Hundecha, Y., and Wagener, T.: Understanding hydrologic variability across Europe through catchment classification, Hydrology and Earth System Sciences, 21, 2863-2879, https://doi.org/10.5194/hess-21-2863-2017, 2017.

Kumar, R., Samaniego, L., and Attinger, S.: Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations, Water Resources Research, 49, 360-379, https://doi.org/10.1029/2012wr012195, 2013.

Li, L., Sullivan, P. L., Benettin, P., Cirpka, O. A., Bishop, K., Brantley, S. L., Knapp, J. L. A., van Meerveld, I., Rinaldo, A., Seibert, J., Wen, H., and Kirchner, J. W.: Toward catchment hydro-biogeochemical theories, WIREs Water, 8, e1495, https://doi.org/10.1002/wat2.1495, 2021.

Livneh, B., Kumar, R., and Samaniego, L.: Influence of soil textural properties on hydrologic fluxes in the Mississippi river basin, Hydrological Processes, 29, 4638-4655, https://doi.org/10.1002/hyp.10601, 2015.

Merz, R., Tarasova, L., and Basso, S.: The flood cooking book: ingredients and regional flavors of floods across Germany, Environmental Research Letters, 15, 114024, https://doi.org/10.1088/1748-9326/abb9dd, 2020.

Monteith, D. T., Stoddard, J. L., Evans, C. D., de Wit, H. A., Forsius, M., Høgåsen, T., Wilander, A., Skjelkvåle, B. L., Jeffries, D. S., Vuorenmaa, J., Keller, B., Kopácek, J., and Vesely, J.: Dissolved organic carbon trends resulting from changes in atmospheric deposition chemistry, Nature, 450, 537-540, https://doi.org/10.1038/nature06316, 2007.

Murphy, R., Perry, E., Keisman, J., Harcum, J., and Leppo, E. W.: baytrends: Long Term Water Quality Trend Analysis. R package version 1.1.0, https://CRAN.R-project.org/package=baytrends, 2019.

Musolff, A.: WQQDB - water quality and quantity data base Germany: metadata, HydroShare [dataset], https://doi.org/10.4211/hs.a42addcbd59a466a9aa56472dfef8721, 2020.

Musolff, A., Fleckenstein, J. H., Opitz, M., Büttner, O., Kumar, R., and Tittel, J.: Spatio-temporal controls of dissolved organic carbon stream water concentrations, Journal of Hydrology, 566, 205-215, https://doi.org/10.1016/j.jhydrol.2018.09.011, 2018.

Musolff, A., Grau, T., Weber, M., Ebeling, P., Samaniego-Eguiguren, L., and Kumar, R.: WQQDB: water quality and quantity data base Germany [dataset], 2020.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T., and Duan, Q.: Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance, Hydrol. Earth Syst. Sci., 19, 209-223, https://doi.org/10.5194/hess-19-209-2015, 2015.

Pflugmacher, D., Rabe, A., Peters, M., and Hostert, P.: Pan-European land cover map of 2015 based on Landsat and LUCAS data, PANGAEA [dataset], https://doi.org/10.1594/PANGAEA.896282, 2018.

Rode, M., Wade, A. J., Cohen, M. J., Hensley, R. T., Bowes, M. J., Kirchner, J. W., Arhonditsis, G. B., Jordan, P., Kronvang, B., Halliday, S. J., Skeffington, R. A., Rozemeijer, J. C., Aubert, A. H., Rinke, K., and Jomaa, S.: Sensors in the Stream: The High-Frequency Wave of the Present, Environmental Science & Technology, 50, 10297-10307, 10.1021/acs.est.6b02155, 2016.

Rotteveel, L. and Sterling, S. M.: The Surface Water Chemistry (SWatCh) database: A standardized global database of water chemistry to facilitate large-sample hydrological research, Earth Syst. Sci. Data Discuss., 2021, 1-17, https://doi.org/10.5194/essd-2021-43, 2021.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resources Research, 46, https://doi.org/10.1029/2008WR007327, 2010.

Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyíri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J. P., Valdebenito, Á., and Wind, P.: The EMEP MSC-W chemical transport model - technical description, Atmos. Chem. Phys., 12, 7825-7865, https://doi.org/10.5194/acp-12-7825-2012, 2012.

Sivapalan, M.: Pattern, Process and Function: Elements of a Unified Theory of Hydrology at the Catchment Scale, in: Encyclopedia of Hydrological Sciences, edited by: Anderson, M. G., and McDonnell, J. J., https://doi.org/10.1002/0470848944.hsa012, 2006.

Tarasova, L., Basso, S., Wendi, D., Viglione, A., Kumar, R., and Merz, R.: A Process-Based Framework to Characterize and Classify Runoff Events: The Event Typology of Germany, Water Resources Research, 56, e2019WR026951, https://doi.org/10.1029/2019WR026951, 2020.

Twarakavi, N. K. C., Sakai, M., and Šimůnek, J.: An objective analysis of the dynamic nature of field capacity, Water Resources Research, 45, https://doi.org/10.1029/2009WR007944, 2009.

UNEP: GEMStat database of the Global Environment Monitoring System for Freshwater (GEMS/Water) Programme., United Nations Environment Programme [dataset], 2018.

van Genuchten, M. T.: A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils, Soil Science Society of America Journal, 44, 892-898, https://doi.org/10.2136/sssaj1980.03615995004400050002x, 1980.

Vigiak, O., Grizzetti, B., Zanni, M., Aloe, A., Dorati, C., Bouraoui, F., and Pistocchi, A.: Domestic waste emissions to European freshwaters in the 2010s (v. 1.0), European Commission, Joint Research Centre (JRC) [Dataset] [dataset], 2019.

Vigiak, O., Grizzetti, B., Zanni, M., Aloe, A., Dorati, C., Bouraoui, F., and Pistocchi, A.: Domestic waste emissions to European waters in the 2010s, Scientific Data, 7, 33, https://doi.org/10.1038/s41597-020-0367-0, 2020.

Virro, H., Amatulli, G., Kmoch, A., Shen, L., and Uuemaa, E.: GRQA: Global River Water Quality Archive, Earth Syst. Sci. Data Discuss., 2021, 1-30, https://doi.org/10.5194/essd-2021-51, 2021.

WMO: Manual on Low-flow Estimation and Prediction, World Meteorological Organization, 2008.

Wollheim, W. M., Bernal, S., Burns, D. A., Czuba, J. A., Driscoll, C. T., Hansen, A. T., Hensley, R. T., Hosen, J. D., Inamdar, S., Kaushal, S. S., Koenig, L. E., Lu, Y. H., Marzadri, A., Raymond, P. A., Scott, D., Stewart, R. J., Vidon, P. G., and Wohl, E.: River network saturation concept: factors influencing the balance of biogeochemical supply and demand of river networks, Biogeochemistry, 141, 503-521, https://doi.org/10.1007/s10533-018-0488-0, 2018.

Yang, S., Bertuzzo, E., Büttner, O., Borchardt, D., and Rao, P. S. C.: Emergent spatial patterns of competing benthic and pelagic algae in a river network: A parsimonious basin-scale modeling analysis, Water Research, 193, 116887, https://doi.org/10.1016/j.watres.2021.116887, 2021.

Zacharias, S. and Wessolek, G.: Excluding Organic Matter Content from Pedotransfer Predictors of Soil Water Retention, Soil Science Society of America Journal, 71, 43-50, https://doi.org/10.2136/sssaj2006.0098, 2007.

Zarnetske, J. P., Bouda, M., Abbott, B. W., Saiers, J., and Raymond, P. A.: Generality of Hydrologic Transport Limitation of Watershed Organic Carbon Flux Across Ecoregions of the United States, Geophysical Research Letters, 45, 11,702-711,711, https://doi.org/10.1029/2018gl080005, 2018.

Zink, M., Kumar, R., Cuntz, M., and Samaniego, L.: A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty, Hydrology and Earth System Sciences, 21, 1769-1790, https://doi.org/10.5194/hess-21-1769-2017, 2017.