Towards understanding of long-term nitrogen transport and retention dynamics across German catchments

Tam V. Nguyen[1], Fanny J. Sarrazin[2,*], Pia Ebeling[1,*], Andreas Musolff[1], Jan H. Fleckenstein[1,3], Rohini Kumar[2]

[1]Department of Hydrogeology, Helmholtz Centre for Environmental Research - UFZ, 04318 Leipzig, Germany.

[2]Department of Computational Hydrosystems, Helmholtz Centre for Environmental Research – UFZ, Leipzig, Germany.

[3]Bayreuth Centre of Ecology and Environmental Research, University of Bayreuth, Bayreuth, Germany.

Corresponding author: Tam V. Nguyen (tam.nguyen@ufz.de)

[*]These authors contributed to the manuscript equally.

Key Points:

- We provided insights into the long-term (1950-2014) nitrogen (N) transport and retention across various German catchments
- Large-sample assessment shows that most of N surplus was removed by denitrification and accumulated in the soil zone
- Four catchment clusters with distinct nitrogen transport and retention dynamics could be distinguished and explained by catchment attributes

Abstract

Elevated nitrate concentrations in German water bodies are a widespread problem, potentially resulting from a long history of excess nitrogen (N) inputs. Here, we investigated long-term (1950-2014) N dynamics across 89 German catchments using a process-based model. Results showed that the mean fractions of N surplus (excess) exported to the river, removed by denitrification, accumulated in the soil zone, and accumulated in groundwater across all catchments are 27%, 58%, 14%, and 1%, respectively. Dissolved inorganic N in groundwater could affect the stream N levels over decades as indicated by long groundwater transit times. A cluster identified four catchment groups with distinct archetypal long-term N transport and retention dynamics, which can be partly linked to the catchments' topographic and geological conditions. This hints at underlying mechanisms that explain spatial differences in the fate of diffuse N inputs to catchments and opens the possibility for better-targeted management.

Plain language summary

High nitrate concentrations in German water bodies are a widespread problem, potentially linked to a long history of excess nitrogen (N) inputs on agricultural fields. In this study, we analyzed the long-term N transport and accumulation in various catchments across Germany from 1950 to 2014 using a process-based model. We further clustered these catchments into different types according to their long-term N patterns and linked these groups with their catchment characteristics. Our results show that only a small part of the net N input was exported to rivers while most of the net N input was lost to the atmosphere (denitrified). The majority of the remaining N surplus was stored in the soil zone. The age of N in discharge was found to be years to decades, suggesting that past N inputs will still have an impact on the future stream water quality status. A cluster identified four catchment groups, which can be partly explained by the catchment's topographic and geological conditions. This hints at underlying mechanisms that explain spatial differences in the fate of diffuse N inputs to catchments and opens the possibility for better-targeted management.

1 Introduction

Human activities, especially agricultural management practices, have drastically changed the Earth's landscape and disturbed the global nitrogen (N) cycle (Foley, 2017; Vitousek et al., 1997). N surplus (excess of N inputs to the soil that were not taken up by crops) from global croplands increased more than fivefold from 16 Tg N yr^{-1} in 1961 to 86 Tg N yr^{-1} in 2010 (Zhang et al., 2021) and it is likely to continue increasing until at least 2050 (Bouwman et al., 2013). In many areas, excess use of N fertilizers for crop production was identified as one of the main causes of surface water and groundwater deterioration, resulting in negative impacts on human health and aquatic ecosystems (Evans et al., 2019). Regulations at the national and international levels, e.g., the Clean Water Act in the United States (EPA, 1972), the Nitrates Directive in Europe (CEC, 1991), and the Action Plan for the Zero Increase of Fertilizer Use in China (Ju et al., 2016)**,** have been introduced to reduce excess N inputs to agricultural lands and to protect water quality. However, the implementation of such mitigation regulations does not always lead to immediate or clear responses of surface water and groundwater quality (Brown & Froemke, 2012; EEA, 2021; Smith et al., 1987). This requires a sound understanding of long-term N transport and retention.

The lag times from changes in N management practices and changes in groundwater or surface water quality vary from years to decades (Chen et al., 2014, 2018; Meals et al., 2010). The reason for these lag times was discussed to be the accumulation of N in the soil (mainly soil organic nitrogen - SON) as biogeochemical legacy and in the subsurface (unsaturated and groundwater, mainly dissolved inorganic nitrogen - DIN) as hydrological legacy (e.g., Basu et al., 2022; Chen et al., 2018; Van Meter et al., 2016, 2017). SON and groundwater DIN accumulations are controlled by the soil mineralization rate in the soil and groundwater transit times, respectively. Several studies suggest that most of the N surplus in the catchment is stored as SON while groundwater DIN is comparatively small (Ascott et al., 2017; Chen et al., 2018; Galloway et al., 2003; Liu et al., 2021; Van Meter et al., 2016). Nevertheless, groundwater DIN storage could affect stream water quality status over decades due to long transit times (Chen et al., 2018). There have been several studies explored the biogeochemical and hydrological lag times, for example, in the Mississippi River basin (Van Meter et al., 2016, 2017), the Susquehanna River basin (Van Meter et al., 2017), the Weser River basin (Sarrazin et al., 2022), and in other basins (Chen et al., 2018). The aforementioned studies, however, were conducted in individual or only a few catchments. Understanding and predicting long-term N transport and retention across a variety of landscape characteristics, hydroclimatic drivers, and anthropogenic impacts rather requires studies with a large sample of catchments.

In recent years, some studies have linked long-term N transport and retention with catchment attributes using a large sample of catchments to discuss underlying processes controlling the build-up of N legacies. However, only a few studies have explicitly separated the soil (biogeochemical legacy) and groundwater (hydrological legacy) N dynamics. For example, McDowell et al. (2021) found that lag times between soil N leaching and riverine N export in 34 catchments in New Zealand varied from 1 to 12 years with higher lag times in catchments with higher altitudes, less steep slope, higher stream order, and higher evapotranspiration. In 14 nested catchments located in the Grand River Watershed, Liu et al. (2021) reported that about 82-96% of the catchment N was stored in the soil and the remaining was stored in groundwater. The mean transit times in groundwater in these catchments ranged from 5 to 34 years with longer transit times found in catchments with higher tile drainage density.

Some recent studies have directly linked N surplus to riverine N export without an explicit separation between the soil zone and groundwater (e.g., Dupas et al., 2020; Ehrhardt et al., 2021). In these studies, 'missing N' is often used to refer to the amount of N that can be either stored in the catchment or be permanently removed via denitrification. For example, lag times between N surplus and the peak riverine N export (mode of the N transport time distribution) in 16 catchments located in Western France were found to vary from 2 to 14 years, depending on catchment lithology (Dupas et al., 2020). In these catchments, about 45-88% of N surplus was missing N. At a larger scale spanning over 238 catchments in Western Europe, the mode of N transport times were reported to be around 5 years, on average with a higher mode of N transport times in catchments with higher potential evapotranspiration and lower precipitation seasonality (Ehrhardt et al., 2021). They also found that catchments with thicker unconsolidated aquifers have a larger amount of missing N while a higher fraction of consolidated and porous aquifers show a smaller amount of missing N. While these studies provided empirical (data-based) evidence on the fate of missing N, there is generally a lack of understanding of the different components of the missing N (e.g., soil N storage, groundwater N storage, soil and groundwater N denitrification) and their relation to catchment characteristics, especially in German landscape. This knowledge gap is important for a more mechanistic understanding of long-term N characteristics in catchments and allows better-targeted management strategies for abating N pollution.

The aims of this study are (1) to provide quantitative estimations of different components of the 'missing N' across German catchments and (2) to discuss the linkages between long-term N transport and retention, and catchment characteristics. To this end, we investigated long-term N transport and retention in different terrestrial components (soil and groundwater) across 89 catchments in Germany with diverse settings. We used a parsimonious, process-based model that allows for an explicit characterization of biogeochemical and hydrological legacies. Moreover, we discussed how our findings could be used for management purposes and provide potential implications for other catchments.

2 Materials and Methods

2.1 Study area and data

The study uses data from 89 catchments (out of which 70 are non-nested catchments) located in Germany (Figure 1a). In total, the study area has a non-overlapping area of 120,596 km^2, which is about one-third of the German territory. The catchment area varies between 19 and 49,760 km^2 with a median area of 742 km^2, covering both German lowlands and mountainous areas. Agriculture is the dominant land use in most of the study catchments, accounting for (median value) 56% of the catchment area. Consolidated rock was found to be the dominant aquifer material in more than half of the selected catchments. The distribution of precipitation, air temperature, topographic gradient (slope), aquifer depth (Figure 1b), and other catchment characteristics (Figure S4) indicate that the selected catchments have diverse settings.

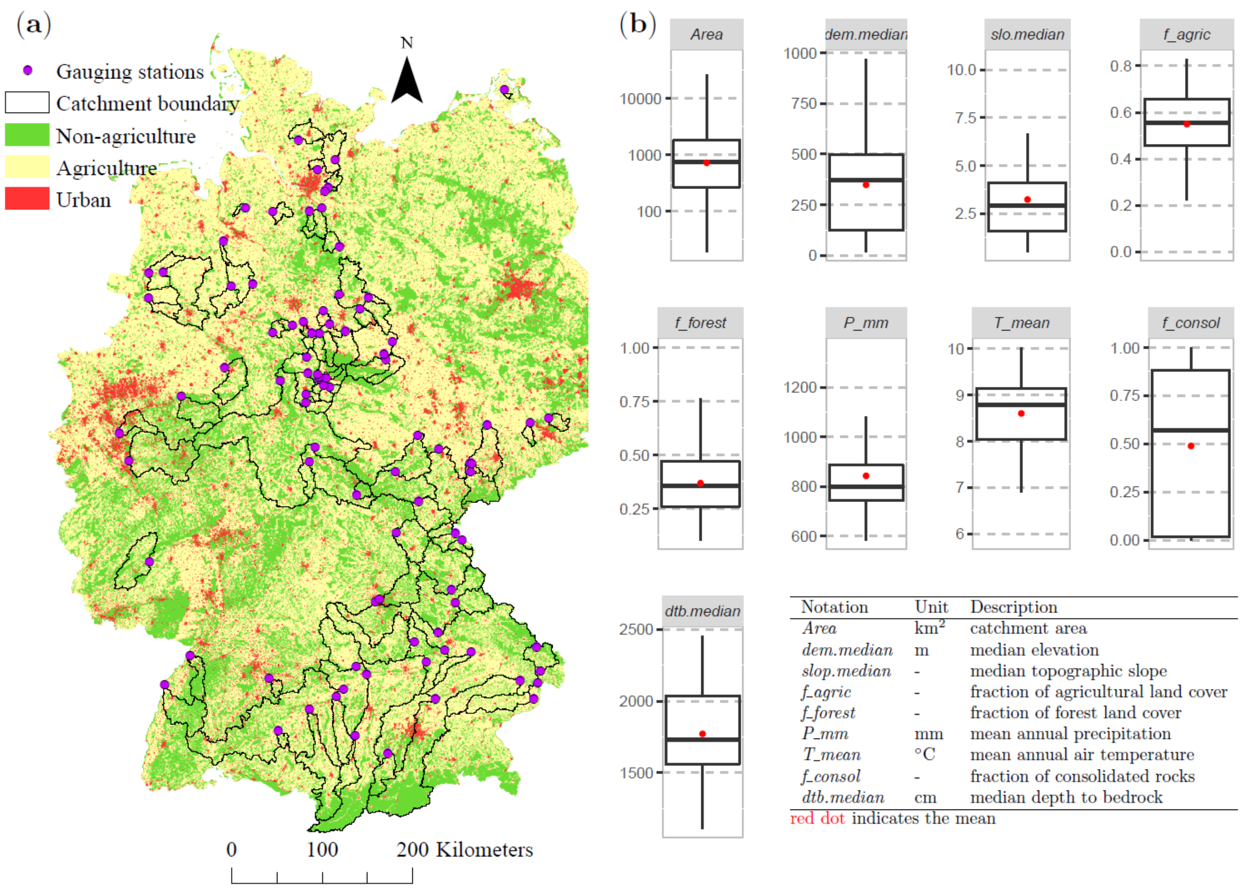

Figure 1. Overview of the selected catchments, represented by (a) the land use map (EEA, 2019) and location of the study catchments, and (b) the boxplots of some catchment attributes (Ebeling et al., 2022). For better visualization, the boxplots only show the values within 1.5 times interquartile range below and above the 25th and 75 percentiles.

The catchment-scale annual N-surplus from 1950 to 2014 was calculated from the fractional contribution from agricultural and non-agricultural land uses (forest, buildup, and other vegetated and non-vegetated lands), based on their relative areas. Land uses were constructed by combining the Corine Land Cover dataset (EEA, 2019), the History Database of the Global Environment dataset (HYDE dataset, Goldewijk et al., 2017), and statistical agricultural area data of Germany (Statistisches Bundesamt, 2021) similar to Sarrazin et al. (2022). The N surplus for agricultural areas is available at the county level for the period 1995-2014 (Häußermann et al., 2020) and at the state level for the period 1950-1998 (Behrendt et al., 2003). The two datasets were harmonized to create consistent time series of N surplus for the period 1950-2014 following Ehrhardt et al. (2021) and Ebeling et al. (2022). The N surplus for non-agricultural areas was estimated as the sum of atmospheric N deposition (Lamarque et al., 2012; Tilmes et al., 2016) and biological N fixation. Biological N fixation rates of 16 kg ha^{-1} yr^{-1} for forest and 2.7 kg ha^{-1} yr^{-1} for the other vegetated land were taken based on the mean rates reported in Cleveland et al., (1999) for temperate forest and natural grassland, respectively (Sarrazin et al., 2022). The catchment-scale annual N point sources for the period 1950-2014 were constructed using the methodology of Morée et al. (2013) and information on population counts (HYDE dataset), protein supply (FAO, 1951, 2021a, 2021b), and population connection to sewer and wastewater treatment plants (WWTPs; Eurostat, 2016, 2021; Seeger, 1999) (for further details see Sarrazin et al., 2022). The reconstructed N loading from WWTPs was constrained to follow the N loading reported by the authority for the period 2012-2016 (Büttner, 2020; Yang et al., 2019), following Sarrazin et al. (2022).

Daily instream nitrate concentrations were reconstructed from irregularly observed instream NO_3-N data using Weighted Regression on Time, Discharge and Season (WRTDS, Hirsch et al., 2010) and were aggregated (discharge-weighted mean) to yearly estimates. Simulated daily discharges from the mesoscale Hydrologic Model (mHM, Kumar et al., 2013; Samaniego et al., 2010) were bias-corrected using piece-wise linear regression and used for gap filling if observed discharges were not available for WRTDS (Ehrhardt et al., 2021). Further details on instream nitrate (NO_3-N) concentrations and discharge data at outlets of selected catchments can be obtained from Ebeling et al. (2022). For all of the selected gauging stations, the minimum time series length of instream NO_3-N concentrations was 20 years and the median number of observations was 426 [min = 154, max = 1294]. In general, the performance of the WRTDS is acceptable (Figure S3) with a median R^2 of 0.63 (interquartile range = [0.49, 0.73]).

2.2 Representation of N transport in the catchment

In this study, we used a parsimonious representation of soil N dynamics and a mechanistic representation of N transport in groundwater using the concept of StorAge Selection (SAS) function (Botter et al., 2011; Nguyen et al., 2021, 2022; Van der Velde et al., 2010). The model, called the StorAge Selection function for Nitrate (SAS-N, Figure S1), consists of two dominant N storages representing the soil zone and groundwater (e.g., Nguyen et al., 2021; Van Meter et al., 2017). The SAS-N model (1) can be considered as an improved version of the catchment-scale lumped transfer function approach (Ehrhardt et al., 2021) with an explicit representation of the soil and groundwater compartments, and (2) has a more realistic representation of groundwater transport with dynamics groundwater transit times compared to other models (e.g., Van Meter et al., 2017). The SAS-N model operates at a yearly time step and is driven by N surplus and effective precipitation (the difference between precipitation and actual evapotranspiration). N surplus can be accumulated in the soil zone as soil organic nitrogen (SON), denitrified, or leached to the groundwater as dissolved inorganic nitrogen (DIN, nitrate). Leached N to the groundwater can be further denitrified and exported to the stream using the SAS approach (Benettin et al., 2013; Nguyen et al., 2022). N point sources (e.g., from WWTPs) are added to the riverine N export and routed to the catchment outlet taking into account instream removal (Sarrazin et al., 2022). A detailed description of the SAS-N model is given in the supporting information (Text S1).

The SAS-N model contains six calibration parameters (Text S1 and Table S1). These parameters were identified by running the model for each catchment with 50,000 parameter sets generated by uniform Latin Hypercube Sampling within their pre-defined ranges (Table S1). The model performance was evaluated against instream nitrate concentrations at the corresponding catchment outlet with the root mean square error. The model was run from 1800 to 2014 with 1800-1949 taken as the warm-up period. Results from the 30 best simulations from each catchment were used for all of the following analyses (see Text S2 for more detail on the model performance).

2.3 Cluster analysis

The objectives of the cluster analysis were to find distinct archetypes of long-term N transport and retention and to characterize their relationships with catchment attributes. In water quality studies, the k-means clustering algorithm (Hartigan & Wong, 1979) has been used, e.g., to understand patterns and controls of catchment-scale nitrate storage (Ascott et al., 2017), groundwater geochemistry (Frapporti et al., 1993), and aquifer vulnerability (Javadi et al., 2017). As an unsupervised machine learning approach, k-means clustering does not require prior knowledge about the underlying patterns of the datasets. The modelled long-term (1950-2014) mean behavioral N fluxes and stores characterizing transport and retention processes, including the transit times, from the 30 best model simulations (behavioral simulations) for each catchment were used for the clustering (Text S3). Then, statistical properties of various catchment attributes (Figure 1b and S4) within each cluster were calculated to identify differences in the catchment attributes among clusters. The tuning parameter of the k-means is the number of clusters that we optimized using a combination of the silhouette (Rousseeuw, 1987), elbow (Kodinariya & Makwana, 2013), and gap statistic (Tibshirani et al., 2001) methods to have a robust estimation (Figure S5).

3 Results and discussion

3.1 Long-term N transport and retention

The simulated long-term (1950-2014) N fluxes and stores across all catchments (Figure 3a and Text S3) shows that only 27 (mean of mean behavioral simulations) ± (standard deviation of mean behavioral simulations) 13% of N surplus was exported to the stream, in other words, the 'missing N' accounts for 73 ± 13% of N surplus (equivalent to 35 ± 6 kg ha^{-1} year^{-1}). These estimated values are well within the range reported by Ehrhardt et al. (2021) for Western European catchments. Results from our study suggest that the majority of N surplus was removed by denitrification in the soil zone (30 ± 15%) and groundwater (27 ± 11%). This is in line with the findings from Sarrazin et al. (2022), who showed that more than half of the N surplus in the Weser catchment in Germany was removed via denitrification. Seitzinger et al. (2006) also found that denitrification in the soil was generally higher than in groundwater at a global scale. About 14 ± 11% of N surplus that entered the catchments between 1950 and 2014 was accumulated in the soil zone while only 1 ± 0.9% in the groundwater with an average groundwater N storage of nearly 33 kg ha^{-1} in 2014 across all catchments. A dominance of soil N accumulation over groundwater N accumulation in catchments has been confirmed in earlier studies across western France (Dupas et al., 2020), the Danube (Malagó et al., 2017), the Weser (Sarrazin et al., 2022), and the Mississippi (Van Meter et al., 2016) river basins. An independent estimation based on groundwater N-stocks and maximum increase in groundwater nitrate concentration also showed that only around 1% of N surplus was accumulated in the European groundwater zone (Howarth et al., 1996). Although groundwater N accumulation during the study period (1950-2014) was found to be low compared to soil N storage, groundwater N storage predominantly consists of dissolved inorganic N in the form of nitrate, which could affect stream water quality status over decades in catchments with very long transit times. For example, we found that the mean transit times of discharge (and dissolved N), the time elapsed since a water parcel enters the groundwater to the time it leaves the catchment via discharge, varied between 3.2 and 20.3 years with a median value of 7.1 years (Figure 2b). It should be noted that there is also a variability in the simulated long-term N fluxes and stores among behavioral simulations within a catchment. In general, higher simulated fluxes or storages have higher standard deviations, except the instream N export because it is the calibrated variable (Figure S6).

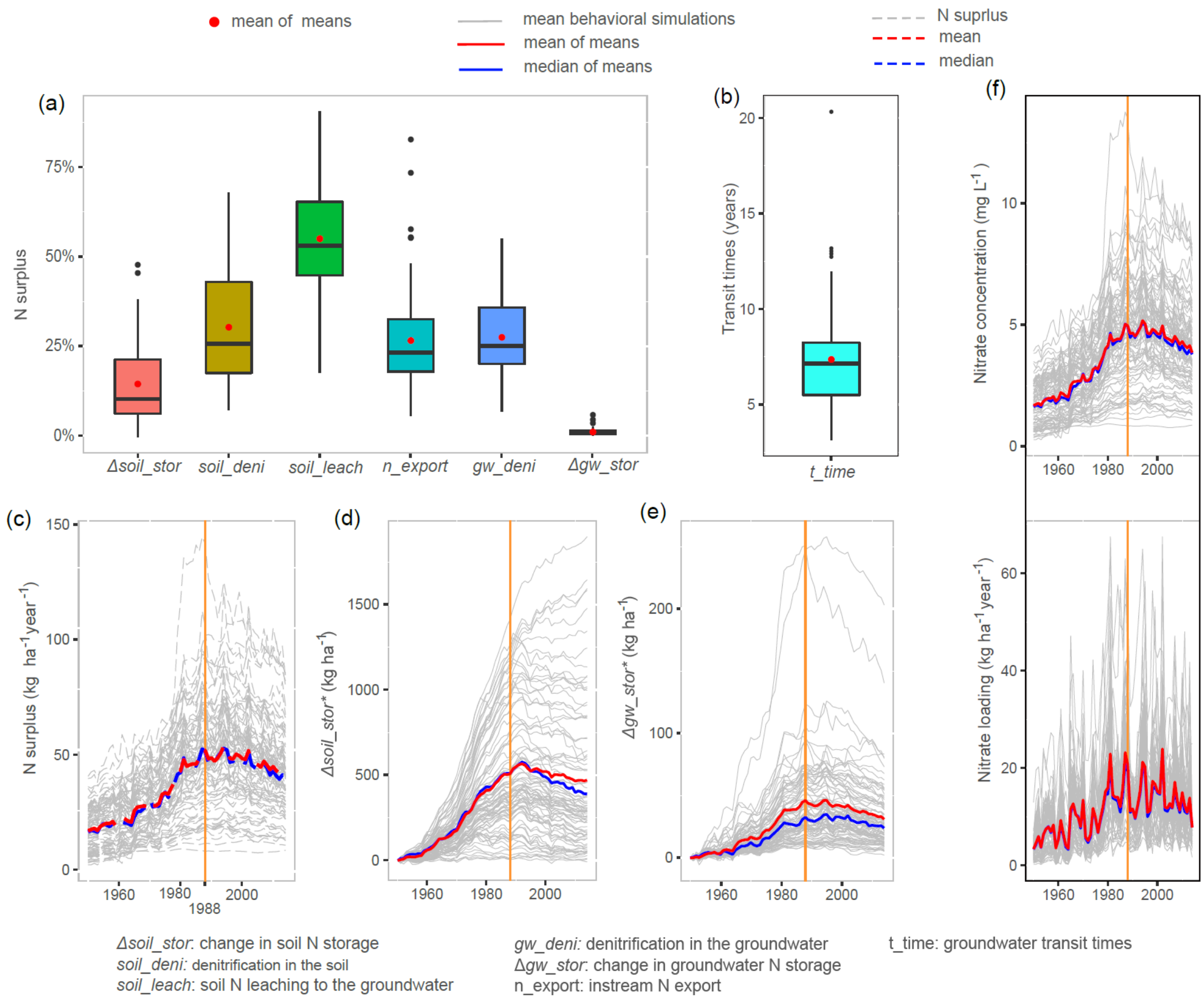

Figure 2. Long-term (1950-2014) N transport and retention in the study catchments (Text S3), represented by (a) boxplots of the long-term mean behavioral N fluxes and stores and (b) mean behavioral groundwater transit times, and the simulated time series of (c) N surplus, (d, e) changes in the mean behavioral soil (*Δsoil_stor**) and groundwater (*Δgw_stor**) N storages, respectively, since 1950, and (f) riverine N export in terms of the mean behavioral concentration and loading. Vertical lines in orange in panels (c-f) depict the year 1988.

The time series of N fluxes and stores among different catchments show a wide range of variations in levels but also similarities in patterns (Figure 2c-f). The N surplus, mean behavioral soil N and groundwater N accumulations from all catchments show a significant increasing trend (Mann-Kendall trend test (MK, Mann, 1945; Kendall, 1975) with *p-value* < 0.001) during the 1950-1988 period (Figure 2c-e). After 1988, N surplus declined significantly (MK, *p-value* < 0.05, mean slope = -0.93) in 74 catchments, out of which 13 and 3 catchments nevertheless showed an increasing trend in soil N and groundwater N accumulation (MK, *p-value* < 0.05, mean slope < -0.77), respectively. While the median N surplus across all 89 catchments in 2014 was reduced by 57% compared to that of 1988, the median of the mean behavioral soil, groundwater N accumulation, instream N concentrations and loadings decreased only by 15% and 16%, 23%, and 49%, respectively (Figure 2c-f, blue lines). The small reduction of groundwater N storage since 1988 found in this study is also in line with a slight decline in observed groundwater nitrate concentrations in recent decades across Germany (Van Grinsven et al., 2012).

3.2 Linking N characteristics to landscape attributes

Results from the k-means analysis indicate that the study catchments can be grouped into four clusters based on their underlying N export and retention dynamics (Figures 3a-b and S5). In general, catchments in the same cluster are located closer to each other (Figure 3a). This is expected as spatial similarity in neighborhood exists for many hydrological and water quality processes (Detenbeck et al., 1996; Western et al., 2004). The number of catchments within each cluster varies from 12 to 32 and each catchment cluster shows distinct long-term N dynamics (Figures 3a-b). The salient features of the four clusters can be summarized as: catchment cluster one has high soil N leaching (*soil_leach*) and riverine N export (*n_export*), and short groundwater transit times (*t_time)*, catchment cluster two is characterized by high soil (*Δsoil_stor*) and groundwater (*Δgw_stor*) N accumulation and long groundwater transit times (*t_time*), catchment cluster three shows high soil denitrification (*soil_deni*), and catchment cluster four has high groundwater denitrification (*gw_deni*).

Regarding the catchment attributes, catchments in **cluster one** are characterized by high altitude (*dem.median*), high precipitation (*P_mm*), high topographic slopes (*slo.median*), low topographic wetness index (*twi_mean*) (Figure 3c). We argue that these conditions lead to a dominance of fast shallow flow paths with short transit times in both soil and groundwater, resulting in low soil and groundwater N storage, high soil N leaching, low denitrification, and high riverine N export relatively to N surplus (Figure 3a). In addition, shallow aquifers and high fraction of consolidated rocks (*f_consol*) in catchment cluster one are also factors that may lead to low N storage and short transit times in groundwater (Figure 3b-c). The catchments in cluster one are also minimally disturbed mountainous forested catchments with low N surplus (Figure 3c). In contrast, catchments in **cluster two** can be interpreted as managed lowland catchments (low altitudes and slopes) with agriculture-dominated landscapes and high N surplus (Figure 3c). Lower precipitation and higher aridity (*AI*) in these catchments could cause lower soil moisture that restricts soil denitrification and flushing (leaching) of soil N, leading to higher soil N storage. Lower topographic slopes and deeper aquifers observed in these catchments facilitate deeper flow paths with longer transit times. Long transit times in combination with low aquifer denitrification rate could be an explanation for the relatively high fraction of groundwater N accumulation compared to the other clusters. High N accumulation in the catchments leads to low riverine N export.

Catchment **cluster three** is located in a comparable range of altitudes to catchment cluster one but with lower slopes and higher fractions of agriculture, lower precipitation, higher precipitation seasonality, and lower mean temperature. Soil denitrification in the catchment cluster three was found to be the highest among the four catchment clusters. The precipitation seasonality (*P_SIsw*) with higher summer precipitation causes higher soil moisture during the warm and biologically active season and could thus enhance soil denitrification. Addionally, high soil pH might cause high soil denitrification in the cluster three, as shown for southern

Germany (Müller et al., 2022). Groundwater denitrification in the catchment cluster three is relatively low compared to the others due to low soil N leaching. Catchment **cluster four** is located in the lowland areas as is catchment cluster two, but with slightly higher precipitation, causing higher soil N leaching and lower soil N storage (Figure 3a, c). The mean fraction of sedimentary aquifers (*f_sedim*) in catchment cluster four is the highest among the four catchment clusters with deep aquifer. This could indicate long transit times, high anoxic conditions and abundance of electron donors (Ebeling et al., 2021; Knoll et al., 2019), resulting in high groundwater denitrification in the catchment cluster four.

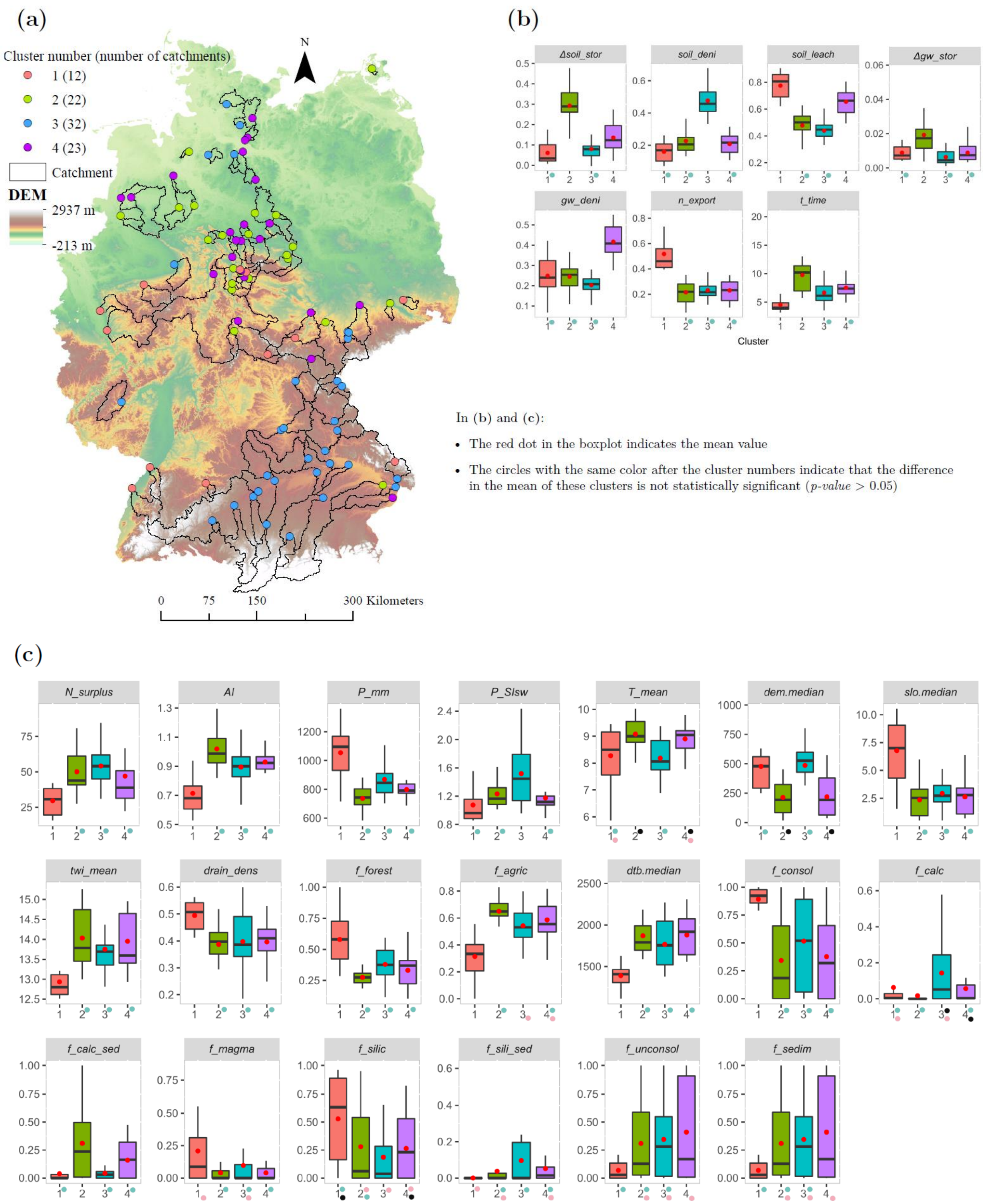

Figure 3. Clustering of catchment functioning based on long-term N characteristics (a) spatial distribution of catchment clusters, (b) boxplots of long-term mean N characteristics (Text S3) in four clusters, and (c) boxplots of catchment attributes (Table S3) of the corresponding clusters. The *aov* (Analysis of Variance Model) and *TukeyHSD* (Tukey Honest Significant Differences) R (R Core Team, 2021) functions were used for comparing the means among clusters.

Catchment attributes that are not statistically different at least in one cluster are not shown here. For better visualization, the boxplots only show the values within 1.5 times interquartile range below and above the 25^{th} and 75 percentiles.

4 Summary and implication

In this modeling study, we were able to shed new light on the fate of the 'missing N' across German catchments from 1950 to 2014 and their linkage with catchment characteristics. The results found in this study, however, are subjected to uncertainty due to, for example, data and parameter uncertainties that have not been fully explored (Sarrazin et al., 2022). Nevertheless, our findings are quantitatively in line with existing studies within the study area or elsewhere (Section 3.1) and the cluster analysis gave plausible results regarding existing process understanding (Section 3.2). Our results suggest that there is in general a large amount of accumulated N in the soil zone as biogeochemical legacy while the magnitude of groundwater N (in form of dissolved inorganic N) accumulation is low. Still, both biogeochemical and hydrological N legacies could have a significant impact on instream water quality for the next few decades as shown by the mean transit time of discharge could be up to 20.3 years. The k-means clustering identified four catchment clusters with different N transport and retention characteristics, which are further explained by some of the selected landscape attributes (e.g., climatic, topographic, and aquifer properties).

We propose that results from the cluster analysis can be used for a qualitative assessment of long-term N characteristics in other catchments within and beyond the physical boundaries of our study area. In particular, our results have shown that catchments located in close spatial proximity tend to behave more similarly than catchments located at more distant from each other. Therefore, long-term N characteristics in ungauged catchments can possibly be inferred from their neighboring catchments. On the other hand, knowing the catchment attributes could help to identify the catchment archetype of N transport (cluster) as demonstrated in this study (Section 3). The linkage between catchment characteristics and dominant N transport, storage and removal processes could inform the development of robust parameter regionalization techniques in future modelling studies (e.g., Kumar et al., 2013; Samaniego et al., 2010).

This study highlights the importance of considering N legacy effects in water quality modeling, management, evaluation programs, and having catchment-specific N management approaches as catchment responses to N surplus are highly heterogeneous. Neglecting N legacies in catchment water quality modeling could provide "the right results for the wrong reasons", leading to false conclusions for management practices. In catchments with a high accumulation of N in the soil zone, a long-term effort could be needed to achieve good chemical status for the groundwater bodies as N in the soil zone will continue to leach to the groundwater, potentially causing elevated groundwater N concentrations over a long period. This, together with long transit times in groundwater, could delay the effects of current management practice and improvement in surface water quality, which should be taken into account for evaluation programs. To have effective, locally adapted management and evaluation programs, it is necessary to answer questions about the expected timing and magnitude of improvements in surface water and groundwater quality after new mitigation measures have been introduced.

Acknowledgments

We thank Stefanie R. Lutz for her comments on the initial version of the manuscript. Pia Ebeling was funded by the Deutsche Forschungsgemeinschaft (DFG).

Open research

The model code and model results are available at https://doi.org/10.5281/zenodo.6788552. All catchment attributes can be obtained from https://www.hydroshare.org/resource/88254bd930d1466c85992a7dea6947a4/

References

Ascott, M. J., Gooddy, D. C., Wang, L., Stuart, M. E., Lewis, M. A., Ward, R. S., & Binley, A. M. (2017). Global patterns of nitrate storage in the vadose zone. *Nature Communications*, *8*, 1416. https://doi.org/10.1038/s41467-017-01321-w

Basu, N. B., Van Meter, K. J., Byrnes, D. K., Van Cappellen, P., Brouwer, R., Jacobsen, B. H., Jarsjö, J., Rudolph, D. L., Cunha, M. C., Nelson, N., Bhattacharya, R., Destouni, G., & Olsen, S. B. (2022). Managing nitrogen legacies to accelerate water quality improvement. *Nature Geoscience*, *15*, 97–105. https://doi.org/10.1038/s41561-021-00889-9

Behrendt, H., Bach, M., Kunkel, R., Opitz, D., Pagenkopf, W.-G., Scholz, G., & Wendland, F. (2003). Nutrient Emissions into River Basins of Germany on the Basis of a Harmonized Procedure. *Federal Environmental Agency*.

Benettin, P., Van Der Velde, Y., Van Der Zee, S. E. A. T. M., Rinaldo, A., & Botter, G. (2013). Chloride circulation in a lowland catchment and the formulation of transport by travel time distributions. *Water Resources Research*, *49*(8), 4619–4632. https://doi.org/10.1002/wrcr.20309

Botter, G., Bertuzzo, E., & Rinaldo, A. (2011). Catchment residence and travel time distributions: The master equation. *Geophysical Research Letters*, *38*(11), L11403. https://doi.org/10.1029/2011GL047666

Bouwman, L., Goldewijk, K. K., Van Der Hoek, K. W., Beusen, A. H. W., Van Vuuren, D. P., Willems, J., Rufino, M. C., & Stehfest, E. (2013). Exploring global changes in nitrogen and phosphorus cycles in agriculture induced by livestock production over the 1900-2050 period. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(52), 20882–20887. https://doi.org/10.1073/pnas.1012878108

Brown, T. C., & Froemke, P. (2012). Nationwide assessment of nonpoint source threats to water quality. *BioScience*, *62*(2), 136–146. https://doi.org/10.1525/bio.2012.62.2.7

Büttner, O. (2020). *DE-WWTP - data collection of wastewater treatment plants of Germany (status 2015, metadata). HydroShare*. https://doi.org/10.4211/hs.712c1df62aca4ef29688242eeab7940c (accessed 07 November 2021)

CEC. (1991). Council Directive 91/676/EEC of 12 December 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources. (No L 375/1). *Official Journal of the European Communities*.

Chen, D., Huang, H., Hu, M., & Dahlgren, R. A. (2014). Influence of lag effect, soil release, and climate change on watershed anthropogenic nitrogen inputs and riverine export dynamics. *Environmental Science and Technology*, *48*(10), 5683–5690. https://doi.org/10.1021/es500127t

Chen, D., Shen, H., Hu, M., Wang, J., Zhang, Y., & Dahlgren, R. A. (2018). Legacy Nutrient Dynamics at the Watershed Scale: Principles, Modeling, and Implications. *Advances in Agronomy*, *149*, 237–313. https://doi.org/10.1016/bs.agron.2018.01.005

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., Hlseroad, A., & Wasson, M. F. (1999). Global patterns of terrestrial biological nitrogen (N2) fixation in natural ecosystems. *Global Biogeochemical Cycles*, *13*(2). https://doi.org/10.1002/(ISSN)1944-9224

Detenbeck, N. E., Taylor, D. L., Lima, A., & Hagley, C. (1996). Temporal and spatial variability in water quality of wetlands in the Minneapolis/St. Paul, MN metropolitan area: Implications for monitoring strategies and designs. *Environmental Monitoring and Assessment*, *40*, 11–40. https://doi.org/10.1007/BF00395165

Digizeitschriften. (2021). *Statistisches Jahrbuch*. http://www.digizeitschriften.de/dms/toc/?PID=PPN635628112 (accessed 07 November 2021)

Dupas, R., Ehrhardt, S., Musolff, A., Fovet, O., & Durand, P. (2020). Long-term nitrogen retention and transit time distribution in agricultural catchments in western France. *Environmental Research Letters*, *15*(11), 115011. https://doi.org/10.1088/1748-9326/abbe47

Ebeling, P., Kumar, R., R. Lutz, S., Nguyen, T., Sarrazin, F., Weber, M., Büttner, O., Sabine, A., & Musolff, A. (2022). Water quality, discharge and catchment attributes for large-sample studies in Germany – QUADICA. *Earth Syst. Sci. Data Discuss. [Preprint]*. https://doi.org/10.5194/essd-2022-6

Ebeling, P., Kumar, R., Weber, M., Knoll, L., Fleckenstein, J. H., & Musolff, A. (2021). Archetypes and Controls of Riverine Nutrient Export Across German Catchments. *Water Resources Research*, *57*(4), e2020WR028134. https://doi.org/10.1029/2020wr028134

EEA. (2019). Corine Land Cover Version 20, Release date: 14-06-2019. *European Environment Agency (EEA) under the Framework of the Copernicus Land Monitoring Service [Data Set]*. https://land.copernicus.eu/pan-european/corine-land-cover (accessed 07 November 2021)

EEA. (2021). *REPORT FROM THE COMMISSION TO THE COUNCIL AND THE EUROPEAN PARLIAMENT on the implementation of Council Directive 91/676/EEC concerning the protection of waters against pollution caused by nitrates from agricultural sources based on Member State reports fo*. https://op.europa.eu/en/publication-detail/-/publication/2596c08f-2a8b-11ec-bd8e-01aa75ed71a1/language-en (accessed 03 June 2022)

Ehrhardt, S., Ebeling, P., Dupas, R., Kumar, R., Fleckenstein, J. H., & Musolff, A. (2021). Nitrate Transport and Retention in Western European Catchments Are Shaped by Hydroclimate and Subsurface Properties. *Water Resources Research*, *57*(10), e2020WR029469. https://doi.org/10.1029/2020WR029469

EPA, U. S. E. P. A. (1972). *Federal water pollution control amendments of 1972. Clean water act. 33 U.S.C. § 1251 et seq.* https://www3.epa.gov/npdes/pubs/cwatxt.txt

Eurostat. (2016). Population connected to wastewater collection and treatment systems by NUTS 2 regions [env_wwcon_r2] (last updated 11.04.2016). *European Commission*. Population connected to wastewater collection and treatment systems by NUTS 2 regions (accessed 07 November 2021)

Eurostat. (2021). Population connected to wastewater treatment plants [env_ww_con] (last update 02-03-2022). *European Commission*. http://appsso.eurostat.ec.europa.eu/nui/show.do?lang=en&dataset=env_ww_con (accessed 07 November 2021)

Evans, A. E., Mateo-Sagasta, J., Qadir, M., Boelee, E., & Ippolito, A. (2019). Agricultural water pollution: key knowledge gaps and research needs. *Current Opinion in Environmental Sustainability*, *36*, 20–27. https://doi.org/10.1016/j.cosust.2018.10.003

FAO. (1951). The state of food and agriculture 1951 - Review and overlook. *Food and Agriculture Organization of the United Nations (FAO), Rome, Italy*. https://www.fao.org/publications/sofa/en/ (accessed 07 November 2021)

FAO. (2021a). Food Balances (-2013, old methodology and population) (last update 12.12.2017). *Food and Agriculture Organization of the United Nations (FAO), Rome, Italy*. http://www.fao.org/faostat/en/#data/FBSH (accessed 07 November 2021)

FAO. (2021b). New Food Balances (last update 14.04.2021). *Food and Agriculture Organization of the United Nations (FAO), Rome, Italy*. http://www.fao.org/faostat/en/#data/FBS (accessed 07 November 2021)

Foley, J. (2017). Living by the lessons of the planet. *Science*, *356*(6335), 251–252. https://doi.org/10.1126/science.aal4863

Frapporti, G., Vriend, S. P., & van Gaans, P. F. M. (1993). Hydrogeochemistry of the shallow dutch groundwater: Interpretation of the National Groundwater Quality Monitoring Network. *Water Resources Research*, *29*(9), 2993–3004. https://doi.org/10.1029/93WR00970

Galloway, J. N., Aber, J. D., Erisman, J. W., Seitzinger, S. P., Howarth, R. W., Cowling, E. B., & Cosby, B. J. (2003). The nitrogen cascade. *BioScience*, *53*(4), 341–356. https://doi.org/10.1641/0006-3568(2003)053[0341:TNC]2.0.CO;2

Goldewijk, K. K., Beusen, A., Doelman, J., & Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene - HYDE 3.2. *Earth System Science Data*, *9*(2). https://doi.org/10.5194/essd-9-927-2017

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *Applied Statistics*, *28*(1), 100–108. https://doi.org/10.2307/2346830

Häußermann, U., Klement, L., Breuer, L., Ullrich, A., Wechsung, G., & Bach, M. (2020). Nitrogen soil surface budgets for districts in Germany 1995 to 2017. *Environmental Sciences Europe*, *32*(1). https://doi.org/10.1186/s12302-020-00382-x

Howarth, R. W., Billen, G., Swaney, D., Townsend, A., Jaworski, N., Lajtha, K., Downing, J. A., Elmgren, R., Caraco, N., Jordan, T., Berendse, F., Freney, J., Kudeyarov, V., Murdoch, P.,

& Zhu, Z. L. (1996). Regional nitrogen budgets and riverine N & P fluxes for the drainages to the North Atlantic Ocean: Natural and human influences. *Biogeochemistry*, *35*, 75–139. https://doi.org/10.1007/BF02179825

Ilampooranan, I., Van Meter, K. J., & Basu, N. B. (2019). A Race Against Time: Modeling Time Lags in Watershed Response. *Water Resources Research*, *55*(5), 3941–3959. https://doi.org/10.1029/2018WR023815

Javadi, S., Hashemy, S. M., Mohammadi, K., Howard, K. W. F., & Neshat, A. (2017). Classification of aquifer vulnerability using K-means cluster analysis. *Journal of Hydrology*, *549*, 27–37. https://doi.org/10.1016/j.jhydrol.2017.03.060

Ju, X., Gu, B., Wu, Y., & Galloway, J. N. (2016). Reducing China's fertilizer use by increasing farm size. *Global Environmental Change*, *41*, 26–32. https://doi.org/10.1016/j.gloenvcha.2016.08.005

Kendall, M. G. (1975). Rank Correlation Methods (4th edn.) Charles Griffin. *San Francisco, CA*.

Knoll, L., Breuer, L., & Bach, M. (2019). Large scale prediction of groundwater nitrate concentrations from spatial data using machine learning. *Science of the Total Environment*, *668*, 1317–1327. https://doi.org/10.1016/j.scitotenv.2019.03.045

Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal of Advance Research in Computer Science and Management Studies*, *1*(6), 90–95.

Kumar, R., Samaniego, L., & Attinger, S. (2013). Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. *Water Resources Research*, *49*(1), 360–379. https://doi.org/10.1029/2012WR012195

Lamarque, J. F., Emmons, L. K., Hess, P. G., Kinnison, D. E., Tilmes, S., Vitt, F., Heald, C. L., Holland, E. A., Lauritzen, P. H., Neu, J., Orlando, J. J., Rasch, P. J., & Tyndall, G. K. (2012). CAM-chem: Description and evaluation of interactive atmospheric chemistry in the Community Earth System Model. *Geoscientific Model Development*, *5*(2). https://doi.org/10.5194/gmd-5-369-2012

Liu, J., Van Meter, K. J., McLeod, M. M., & Basu, N. B. (2021). Checkered landscapes: Hydrologic and biogeochemical nitrogen legacies along the river continuum. *Environmental Research Letters*, *16*, 115006. https://doi.org/10.1088/1748-9326/ac243c

Malagó, A., Bouraoui, F., Vigiak, O., Grizzetti, B., & Pastori, M. (2017). Modelling water and nutrient fluxes in the Danube River Basin with SWAT. In *Science of the Total Environment* (Vols. 603–604, pp. 196–218). https://doi.org/10.1016/j.scitotenv.2017.05.242

Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, *13*(3), 245–259. https://doi.org/10.2307/1907187

McDowell, R. W., Simpson, Z. P., Ausseil, A. G., Etheridge, Z., & Law, R. (2021). The implications of lag times between nitrate leaching losses and riverine loads for water quality policy. *Scientific Reports*, *11*, 16450. https://doi.org/10.1038/s41598-021-95302-1

Meals, D. W., Dressing, S. A., & Davenport, T. E. (2010). Lag Time in Water Quality Response to Best Management Practices: A Review. *Journal of Environmental Quality*, *39*(1), 85–96. https://doi.org/10.2134/jeq2009.0108

Morée, A. L., Beusen, A. H. W., Bouwman, A. F., & Willems, W. J. (2013). Exploring global nitrogen and phosphorus flows in urban wastes during the twentieth century. *Global Biogeochemical Cycles*, *27*(3). https://doi.org/10.1002/gbc.20072

Müller, T. S., Dechow, R., & Flessa, H. (2022). Inventory and assessment of pH in cropland and grassland soils in Germany. *Journal of Plant Nutrition and Soil Science*, *185*(1). https://doi.org/10.1002/jpln.202100063

Nguyen, T. V., Kumar, R., Lutz, S. R., Musolff, A., Yang, J., & Fleckenstein, J. H. (2021). Modeling Nitrate Export From a Mesoscale Catchment Using StorAge Selection Functions. *Water Resources Research*, *57*(2), e2020WR028490. https://doi.org/10.1029/2020wr028490

Nguyen, T. V., Kumar, R., Musolff, A., Lutz, S. R., Sarrazin, F., Attinger, S., & Fleckenstein, J. H. (2022). Disparate Seasonal Nitrate Export From Nested Heterogeneous Subcatchments Revealed With StorAge Selection Functions. *Water Resources Research*, *22*(3), e2021WR030797. https://doi.org/10.1029/2021WR030797

R Core Team. (2021). R core team (2021). In *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www. R-project. org*.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Samaniego, L., Kumar, R., & Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, *46*(5), W05523. https://doi.org/10.1029/2008WR007327

Sarrazin, F., Kumar, R., Basu, N. B., Musolff, A., Weber, M., Van Meter, K., & Attinger, S. (2022). Characterizing catchment-scale nitrogen legacies and constraining their uncertainties. *Water Resources Research*, e2021WR031587. https://doi.org/10.1029/2021WR031587

Seeger, H. (1999). The history of German wastewater treatment. *European Water Management*, *2*(5).

Seitzinger, S., Harrison, J. A., Böhlke, J. K., Bouwman, A. F., Lowrance, R., Peterson, B., Tobias, C., & Van Drecht, G. (2006). Denitrification across landscapes and waterscapes: A synthesis. *Ecological Applications*, *16*(6), 2064–2090. https://doi.org/10.1890/1051-0761(2006)016[2064:DALAWA]2.0.CO;2

Smith, R. A., Alexander, R. B., & Wolman, M. G. (1987). Water-quality trends in the nation's rivers. *Science*, *235*(4796), 1607–1615. https://doi.org/10.1126/science.235.4796.1607

Statistisches Bundesamt. (2021). Online data portal Genesis, Tables 41120-01-02-4-B and 41141-01-01-4-B. *Gemeinsames Neues Statistisches Informations-System (GENESIS) [Data Set]*. *Wiesbaden, Germany*. https://www.regionalstatistik.de/genesis/online (accessed 07 November 2021)

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *63*(2), 411–423. https://doi.org/10.1111/1467-9868.00293

Tilmes, S., Lamarque, J. F., Emmons, L. K., Kinnison, D. E., Marsh, D., Garcia, R. R., Smith, A. K., Neely, R. R., Conley, A., Vitt, F., Val Martin, M., Tanimoto, H., Simpson, I., Blake, D. R., & Blake, N. (2016). Representation of the Community Earth System Model (CESM1) CAM4-chem within the Chemistry-Climate Model Initiative (CCMI). *Geoscientific Model Development*, *9*(5). https://doi.org/10.5194/gmd-9-1853-2016

Van der Velde, Y., De Rooij, G. H., Rozemeijer, J. C., Van Geer, F. C., & Broers, H. P. (2010). Nitrate response of a lowland catchment: On the relation between stream concentration and travel time distribution dynamics. *Water Resources Research*, *46*(11), W11534. https://doi.org/10.1029/2010WR009105

Van Grinsven, H. J. M., Ten Berge, H. F. M., Dalgaard, T., Fraters, B., Durand, P., Hart, A., Hofman, G., Jacobsen, B. H., Lalor, S. T. J., Lesschen, J. P., Osterburg, B., Richards, K. G., Techen, A. K., Vertès, F., Webb, J., & Willems, W. J. (2012). Management, regulation and environmental impacts of nitrogen fertilization in northwestern Europe under the Nitrates Directive; A benchmark study. *Biogeosciences*, *9*(12). https://doi.org/10.5194/bg-9-5143-2012

Van Meter, K. J., Basu, N. B., & Van Cappellen, P. (2017). Two centuries of nitrogen dynamics: Legacy sources and sinks in the Mississippi and Susquehanna River Basins. *Global Biogeochemical Cycles*, *31*(1), 2–23. https://doi.org/10.1002/2016GB005498

Van Meter, K. J., Basu, N. B., Veenstra, J. J., & Burras, C. L. (2016). The nitrogen legacy: Emerging evidence of nitrogen accumulation in anthropogenic landscapes. *Environmental Research Letters*, *11*, 035014. https://doi.org/10.1088/1748-9326/11/3/035014

Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., Schlesinger, W. H., & Tilman, D. G. (1997). Human alteration of the global nitrogen cycle: Sources and consequences. *Ecological Applications*, *7*(3), 737–750. https://doi.org/10.1890/1051-0761(1997)007[0737:HAOTGN]2.0.CO;2

Western, A. W., Zhou, S. L., Grayson, R. B., McMahon, T. A., Blöschl, G., & Wilson, D. J. (2004). Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes. *Journal of Hydrology*, *286*(1–4), 113–134. https://doi.org/10.1016/j.jhydrol.2003.09.014

Yang, S., Büttner, O., Kumar, R., Jäger, C., Jawitz, J. W., Rao, P. S. C., & Borchardt, D. (2019). Spatial patterns of water quality impairments from point source nutrient loads in Germany’s largest national River Basin (Weser River). *Science of the Total Environment*, *697*. https://doi.org/10.1016/j.scitotenv.2019.134145

Zhang, X., Zou, T., Lassaletta, L., Mueller, N. D., Tubiello, F. N., Lisk, M. D., Lu, C., Conant, R. T., Dorich, C. D., Gerber, J., Tian, H., Bruulsema, T., Maaz, T. M. C., Nishina, K., Bodirsky, B. L., Popp, A., Bouwman, L., Beusen, A., Chang, J., … Davidson, E. A. (2021). Quantification of global and national nitrogen budgets for crop production. *Nature Food*, *2*(7), 529–540. https://doi.org/10.1038/s43016-021-00318-5