

Assessing global regionalized impacts of eutrophication on freshwater fish biodiversity

Zhou, J.

Citation

Zhou, J. (2024, January 30). *Assessing global regionalized impacts of eutrophication on freshwater fish biodiversity*. Retrieved from https://hdl.handle.net/1887/3715136

Note: To cite this publication please use the final published version (if applicable).

Chapter 2

Regionalized nitrogen fate in freshwater systems on a global scale

This chapter has been published as Zhou, J., Scherer, L., van Bodegom, P.M., Beusen, A., Mogollón, J.M., 2022. Regionalized nitrogen fate in freshwater systems on a global scale. *J Ind Ecol* 26, 907–922.

Abstract

Excessive nitrogen (N) use in agriculture, industry, and household waste leads to widespread N release throughout the environment, causing eutrophication in both freshwater and coastal areas. To better understand N-induced eutrophication and other N-use-related environmental impacts at the local scale, improvements in the spatial resolution of LCIA measures are required. Here, we present a method to estimate gridded fate factors (FFs) at a half-degree resolution based on the IMAGE-Global Nutrient Model (GNM) to provide eutrophication indicators for global N-related manufacture, trade, and consumption in LCA. Across global freshwater systems, our cumulative FFs have a 5th percentile of 0.9 days and a 95th percentile of 184.0 days. Aggregated FFs for administrative units range from 0.3 days to 211.9 days. The hotspots of cumulative FFs are mainly distributed upstream of large reservoirs or lakes. On a global level, advection is the dominant process controlling the FF (69.7% of areas), followed by retention (29.0%), and water consumption (1.3%). N retention dominates in advection-favoring, high-discharge regions due to the high residence times, while water consumption tends to dominate water-scarce zones. The results demonstrate the importance of gridded information to assess eutrophication impacts, as it characterizes N emissions from anthropogenic sources at high spatial resolution in comparison to basin- or country-level assessments. Introducing soil-freshwater N fate complements existing P-related fates to improve global assessments of eutrophication.

2.1 Introduction

Human activities, including food production, detergent and fertilizer use, and waste such as sewage from households and industry, have exacerbated nutrient emissions to the environment, posing a pervasive threat to aquatic ecosystems. Nitrogen (N) is an important nutrient for life, but in excess can lead to eutrophication, hypoxia, and the deterioration of ecosystems (Jenny et al., 2016; Muller et al., 2012; Vonlanthen et al., 2012). Global N input to freshwater systems has grown from 34 to 64 Tg year-1 from 1900 to 2000 (Beusen et al., 2016), leading to an increase in eutrophic (and hypoxic) areas in freshwater and coastal systems (Jenny et al., 2016). Eutrophication can induce excessive reproduction of pernicious algae blooms (Chislock et al., 2013), whose decomposition consumes oxygen and can lead to hypoxia/anoxia in the water column. These conditions may be unable to sustain many aquatic organisms and thus jeopardize biodiversity (Schindler and Vallentyne, 2008), and may lead to a collapse of the aquatic ecosystem. This deterioration of aquatic ecosystems may last months or even years (e.g. algae blooms in Taihu Lake, (Duan et al., 2015)) and is highly likely to intensify because of increasing demand for food, fertilizer use, and industrial production with population growth (Jenny et al., 2016; Mogollón et al., 2018a; Tilman et al., 2001).

Life cycle assessment (LCA) provides a widely recognized framework to quantify environmental impacts, such as eutrophication, throughout the whole life cycle of a specific product (Hellweg and Milà i Canals, 2014; Payen et al., 2019). In the Life Cycle Impact Assessment (LCIA) phase of LCA, characterization factors (CF) relate the emissions or resource use from various life cycle stages to associated environmental impacts. For eutrophication, the fate factor (FF) describes the nutrient fate originating from various anthropogenic emissions and serves as the first step toward assessing the environmental impact of the nutrients from an LCIA standpoint.

Historically, LCA research has evaluated eutrophication indicators ignoring geospatial variation, a limitation that has been pointed out in previous research (Hauschild, 2006; Hauschild and Potting, 2005; Morelli et al., 2018). For instance, the Tool for the Reduction and Assessment of Chemical and other environmental Impacts (TRACI) (Bare, 2011, 2002; Bare et al., 2012) provides a midpoint eutrophication indicator by multiplying a nutrient factor and a transport factor. Nonetheless, TRACI does not model explicit N processes; instead, it derives the nutrient factor from the Redfield ratio to describe the relative influence of P versus N (Norris, 2002). Further, the transport factor, which is the same for N and P, ranges from 0 to 1 to represent the probability of the release arriving in an aquatic environment. TRACI assumes that all emitted nitrogen contributes to eutrophication, and ignores biogeochemical transformations of N before

reaching water bodies and during transport through water bodies (Payen and Ledgard, 2017). More recently, however, coupling LCA with geographic information systems (GIS) has allowed for the ability to identify locations undergoing (or that are susceptible to) N-induced impacts. Helmes et al. (2012) made the first big step towards regionalizing eutrophication impacts by developing a gridded FF model for phosphorus (P). They simulated P fate from its emissions and their model was later integrated into ReCiPe 2016 (Huijbregts et al., 2017), IMPACT World+ (Bulle et al., 2019), and LC-IMPACT (Verones et al., 2020). LCA models focusing specifically on N have mainly been developed for marine ecosystems (Payen et al., 2019), even though N has also been regarded as a nutrient sometimes contributing to freshwater eutrophication (Dodds and Smith, 2016; Lewis et al., 2011; Payen et al., 2019; Schindler, 2006; Vollenweider, 1968). The study of Cosme and Hauschild (2017) estimated CFs for N in 66 large marine ecosystems (LMEs) and their corresponding watershed based on the global Nutrient Export from WaterSheds (NEWS) 2 model (Mayorga et al., 2010). However, a grid-scale FF model for freshwater N is not available globally. Large watersheds are often quite heterogeneous. Thus, gridded models, consisting of much smaller spatial units, can better help evaluate local hotspots where nutrients may accumulate within watersheds. Furthermore, grid cells can be aggregated more accurately than watersheds to any scale, such as the country scale, which is the typical spatial unit of life cycle inventory data. Based on a review of existing spatially explicit fate models, NEWS 2, SWAT (Soil and Water Assessment Tool, (Kalcic et al., 2015)), and IMAGE-GNM (the Integrated Model to Assess the Global Environment–Global Nutrient Model, (Beusen et al., 2015)) have been the recommended options for the quantification of N fate factors for use in LCIA on a watershed scale (Morelli et al., 2018). Among them, NEWS 2 can simulate the nutrient fate for rivers and watersheds, but the resolution is limited to the watershed scale. SWAT can simulate organic nitrogen, organic phosphorus, nitrate, and dissolved inorganic phosphorus at the scale of user-defined hydrologic response units within a basin, but it is seldom applied to wide geographic coverage, let alone globally. Since it provides global nutrient loads and emissions at a half-degree resolution, Cosme et al. (2018) and Morelli et al. (2018) suggested IMAGE-GNM as the most comprehensive option for developing a grid-scale FF model.

In this study, we present a grided, spatially explicit FF model for N emissions to freshwater systems over the globe. We extract information about inland N fate from IMAGE-GNM at $0.5^{\circ} \times 0.5^{\circ}$ grid cells on a global scale for the year 2000, and also run the model for 1998 and 1999 to display the dynamics of N fate in subsequent years. IMAGE-GNM is a dynamic, distributed model with a yearly time step. It depicts nutrient reaction and delivery processes in soils and freshwater systems. The N retention, the N withdrawn via water consumption, and the N advection towards downstream cells – henceforth collectively termed " N removal processes" following Cosme et al. (2018) and Helmes et al. (2012) – as well as the drivers and accompanying uncertainties, are analyzed to better contextualize the meaning of the obtained FFs. By highlighting the importance of the link between the hydraulic drivers and FFs, this analysis allows identifying the possible impact of N in distinctive regions to improve the management of emission sources from production activities.

2.2 Methods

2.2.1 Model structure

N fate in soil and freshwater depends on its input, transport, and removal processes (Helmes et al., 2012; Mayorga et al., 2010; Seitzinger et al., 2005). At any location on land, N is imported from applications, depositions, erosion, and fixation, and further transported to the freshwater. For each grid cell, the N sources compose so-called soil N budgets (the difference between those inputs and N eliminations due to harvesting, grass cutting, and grazing) in IMAGE-GNM. Surplus N is transported via leaching into groundwater or surface runoff. During the soil to freshwater transport, N concentration declines due to absorption, uptake by plants, and denitrification. N is transported to open freshwater bodies via surface runoff, shallow groundwater transport, and deep groundwater transport.

Denitrification occurs in surface water and shallow groundwater that feeds into rivers along various flow paths, while N percolating from shallow to deep groundwater is assumed to not undergo denitrification (Beusen et al., 2015). Finally, the N contained in surface runoff and groundwater arrives at rivers and large water bodies (e.g. lakes and reservoirs). Note that in IMAGE-GNM all N processes (e.g. reaction and transport) taking place within multi-grid water bodies are assumed homogeneous and modeled at the single-cell outlet. IMAGE-GNM also models the treated sewage as a point source emitted directly to freshwater system. The spatial data used in IMAGE-GNM includes land cover, soil, lithology, and climate obtained from open-access databases.

IMAGE-GNM simulates the overall fate of N inputs. While IMAGE-GNM splits emissions into natural sources (e.g. biological fixation) and anthropogenic sources (e.g. synthetic fertilizers and manure), once these sources enter the compartment, they constitute total nitrogen (TN). The ratio of N decay from soil emissions to freshwater does not change within a cell, as it is determined by climate, soil texture, aeration, and soil organic carbon (C) content as opposed to the N soil content. The same applies to N emitted to freshwater, as the retention and residence time depend on the hydrological conditions of the water bodies. Therefore, the separation of natural and anthropogenic does not affect the calculation of cumulative FFs, but it could influence the emissionweighted FFs (i.e., regional average FFs) for diffuse emissions.

Here we estimated FFs for the year 2000, as it represents the most recent year available in IMAGE-GNM (Beusen et al., 2016). To show the temporal variation of FFs, we also examined the years 1998 and 1999, and displayed the relative standard deviation (RSD) of FFs for direct emissions to freshwater in the Supporting Information Figure S2.9. Our method can be replicated for other (more recent) years, once the data become available.

Figure 2.1 Model structure and spatial relation between the model grid cells. λ_{adv} , λ_{ret} , and λ_{con} indicate N removal rate constants for advection, retention, and water consumption, respectively.

N fate in this study is determined by retention and water consumption, as well as advection transporting N to downstream grid cells. We ran IMAGE-GNM and used the model inputs (e.g. emission data), intermediate variables (e.g. retention), and outputs (e.g. nutrient loading) to calculate rate constants ($\lambda_{adv}, \lambda_{ret}$, and λ_{con} , respectively), which we implemented as the advection, retention, and consumption removal processes [\(Figure 2.1\)](#page-8-0) in the fate factor model. These rate constants, which were calculated for each grid cell, are explained further in sections 2.2.1.1 to 2.2.1.3.

The cumulative fate factors ($FF_{e\rightarrow i}$, days, Eq. 2.1) follow the approach of Helmes et al. (2012) and LC-IMPACT (Azevedo et al., 2020). FF $_{e\rightarrow i}$ denotes the sum of individual fate factors from emission source e in cell *i* and all the downstream cells. The individual fate factors (FF_{e→i→j}, days) indicate FFs of the N emitted from source e in cell i to a specific receiving cell j. They are the product of the fraction of N transported from the emission to freshwater ($fr_{e\rightarrow i}$, dimensionless), the fraction of N delivered from the source cell i to receptor cell j by the freshwater system $(f_{i,j}$, dimensionless), and the persistence of N in the receiving compartment j (τ_j , year⁻¹):

$$
FF_{e\to i} = \sum_{j} FF_{e\to i\to j} = 365 \cdot fr_{e\to i} \cdot \sum_{j} f_{i,j} \cdot \tau_{j}
$$
 (2.1)

The resulting $FF_{e\rightarrow i}$ is spatially differentiated and it provides a basis for the environmental impact analysis caused by N emissions. A larger cumulative FF suggests that emissions in the source cell result in a higher possibility and duration of N remaining in the receiving water bodies.

Point sources are regarded as direct loads to the water bodies, and thus $fr_{e\rightarrow i, freehwater}$ equals 1. In contrast, nutrients in the soil are transformed/removed/retained during the transport from the soil to freshwater. For example, during the process of fertilizer application, N may be partially left in the soil compartment and absorbed by plants, and only the remainder of N can be delivered to the freshwater system. IMAGE–GNM distinguishes two emission routes from the soil to freshwater: losses from recent nutrient applications in the form of fertilizer, manure, or organic matter transported by runoff and subsurface delivery (Hart et al., 2004), as well as losses from long-term accumulation in soil compartments, which may be subject to erosion (McDowell and Sharpley, 2001; Tarkalson and Mikkelsen, 2004).

In IMAGE-GNM, the diffuse emissions excluding erosion only include agricultural applications (i.e. it excludes sludge). The transfer fraction of N losses from recent nutrient applications (i.e. nutrient budget which contains fertilizer, animal manure application, and biological N fixation subtracting crop harvesting, grass cutting, and ammonia volatilization) is considered as an export fraction of diffuse loads from the

soil through runoff, drainage, and leaching into groundwater $(fr_{e\rightarrow i,diffuse}, kg N_{water})$ / kg N_{emission}).

$$
fr_{e \to i, \text{diffuse}} = \frac{L_{e \to i, \text{diffuse}}}{E_{e \to i, \text{diffuse}}}
$$
(2.2)

where $E_{e \to i, diffuse}$ is the diffuse emission of source *e* from recent nutrient applications to the soil within grid cell *i* (kg year⁻¹), and $L_{e\rightarrow i,diffuse}$ is the load produced by source *e* in grid cell *i* (kg year⁻¹).

The transfer fraction of soil erosion ($fr_{i,erosion}$, kg N_{water} / (km²·year)) depends on the land use type.

$$
fr_{i, erosion, landscape} = \frac{L_{i,erosion, landscape}}{A_{i, landscape}}
$$
 (2.3)

where $L_{i,erosion, landscape}$ is the soil N eroded to freshwater systems within grid cell *i* (kg year⁻¹), A_i is the area of a given land use type in grid cell *i* (km²). IMAGE-GNM distinguishes arable land, grassland, and natural land. Note that in each grid cell, *f* , and τ_j is the same for all land types, but erosion FFs of different land-use types are distinguished via the enhanced transfer fraction of soil erosion ($fr_{i,erosion, landscape}^*$). This latter parameter reflects anthropogenic pressures due to a relative change from natural land to grassland and arable land $(c_{landuse})$.

$$
fr_{i,erosion, landscape}^{*} = fr_{i,erosion, landscape} - fr_{i,erosion,natural}
$$

$$
= \frac{fr_{erosion,landuse} - fr_{erosion,natural}}{fr_{crossion,natural}} \cdot fr_{i,erosion,natural}
$$

$$
= c_{landuse} \cdot fr_{i,erosion,natural}
$$
(2.4)

Calculating $fr_{i.erosion.$ landare for arable land and grassland as well as $fr_{i,erosion, natural}$ by using Eq. 2.3, we found that $c_{landuse}$ is a constant with values of 2.41 and 45.30 for grassland and arable land, respectively. Given the constant conversion between different land uses and thus equal spatial patterns, FFs of natural land erosion were shown as the baseline in section 2.3.1.

The persistence of N in the receiving water τ_i is defined as the reciprocal of the removal rates. It is related to advection rate $(\lambda_{adv,j},$ year⁻¹), the retention rates $(\lambda_{ret,j},)$ year⁻¹) and the removal rates by consumptive water use $(\lambda_{con,i},$ year⁻¹) in estuaries, river reaches, and lakes.

$$
\tau_j = \frac{1}{\lambda_{adv,j} + \lambda_{ret,j} + \lambda_{con,j}}
$$
(2.5)

The transport fraction of freshwater $f_{i,j}$ can be expressed as a ratio between advection rate and the combined removal rates (Eq. 2.6). In contract to the persistence τ_i , the transport fraction of N in the source cell *i* delivered to receptor cell *j* $(f_{i,j})$ is dimensionless, and is calculated as the ratio of N that reaches the receptor cell *j* to the N export from the source cell *i* (where n denotes the grid cell along the flow path between the source cell *i* to the upstream cell adjacent to the receptor cell *j*). Since the index cell *j* starts in source cell *i*, when $j = i$, $f_{i,i} = 1$. Due to the removal by retention and water consumption along the flow path, the further away from source cell *i*, the less impact is caused in the receptor cell *j*.

$$
f_{i,j} = \prod_{n=1}^{j-1} \frac{\lambda_{adv,n}}{\lambda_{adv,n} + \lambda_{ret,n} + \lambda_{con,n}}
$$
(2.6)

We elaborate on the components of Eq. 2.5 and Eq. 2.6 in sections $2.2.1.1 - 2.2.1.3$.

In our global analysis, we excluded FFs from arid and low-discharge cells. In many of these systems, surface water is highly likely to be unavailable or insufficient to meet the local water demand and could bias the resulting FFs. Here, arid cells were defined to be cells with an aridity index (AI, obtained from Trabucco and Zomer, (2019)) of less than 0.2 (Middleton and Thomas, 1997). However, we still included the dominant rivers (e.g. the Nile River) flowing within arid zones. For arid zones, cells, where the discharge is higher than the median of discharge in non-arid zones (325 mm year⁻¹), were kept, whereas, in non-arid zones, the cells where the discharge is lower than the median of discharge in arid zones (6 mm year⁻¹) were excluded.

2.2.1.1 Advection

The advection rate constant (λ_{adv} in [Figure 2.1\)](#page-8-0) was calculated following Helmes et al. (2012), following the principle that each grid cell undergoes advection from the local cell to a downstream cell.

IMAGE-GNM uses the hydrological model PCRaster Global Water Balance (PCR-GLOBWB (Wood et al., 2011)), where the river channel network is based on the DDM30 flow direction map of Döll and Lehner (2002), which links the upstream and downstream cells. The advection rate constant, $\lambda_{adv,i}$, is related to the water travel rate in river channels. This parameter equals the reciprocal of water residence time $t_{r,j}$ for water bodies, which is determined by the discharge Q_i (m³ year⁻¹) and volume V_i $(m³)$ of the water body.

$$
\lambda_{adv,i} = \frac{1}{t_{r,j}} = \frac{Q_i}{V_i} \tag{2.7}
$$

Water from the ground surface, soil, or aquifer is transported to the river network. Besides exchange between surface and subsurface water through infiltration and percolation, PCR-GLOBWB also simulates direct runoff, interflow, and base flow which are converted into discharge (Beusen et al., 2015). Reservoir regulation is also introduced in discharge modeling. However, the discharge does not reflect consumptive water use in IMAGE-GNM. Therefore, lakes and reservoirs are only included if the volume outstrips the water storage capacity within a cell (Beusen et al., 2015). Lake volumes and areas were taken from the Global Lakes and Wetlands Database version 1 (GLWD1) (Lehner and Döll, 2004), while reservoir data are from the Global Reservoir and Dam (GRanD) database (Lehner et al., 2011). In PCR-GLOBWB, the reservoirs were included in the model dynamically, according to their reported construction time.

2.2.1.2 Retention

N retention (λ_{ret} in Figure 2.1) consists of denitrification in water, sedimentation, and uptake by aquatic plants. As opposed to P, N undergoes little absorption in sediments. Thus, the advection rate of N better approximates the reciprocal of the water residence time. The analysis of hundreds of rivers and lakes by present studies (Behrendt and Opitz, 1999; Seitzinger et al., 2005; Venohr et al., 2005) that also indicated the link between N retention and hydrology. IMAGE-GNM employs the empirical retention equation of Wollheim et al. (2008), where the retention R_i (dimensionless) in cell *i* is a first-order degradation process, shown in Eq. 2.8.

$$
R_i = 1 - exp\left(-\frac{\nu_{f,i}}{H_{L,i}}\right) \tag{2.8}
$$

where $H_{\rm L}$ (m year⁻¹, Eq. 2.9) is the hydraulic load, and v_f (m year⁻¹, Eq. 2.10) is the net uptake velocity. The hydraulic load represents the hydrological characteristics of water bodies. It is determined by the depth (D_i, m) and residence time $(t_{r,i},$ year) of the water body within a cell.

$$
H_{L,i} = \frac{D_i}{t_{r,i}}\tag{2.9}
$$

The net uptake velocity v_f is affected by the biological and chemical features of the nutrient. In IMAGE-GNM, v_f for N takes a base value of 35 m year⁻¹ from Wollheim et al. (2008, 2006) and is modified by the annual temperature T (\degree C) and N concentration C_N :

$$
v_f = 35 \cdot f(T) \cdot f(C_N) \tag{2.10}
$$

 $f(C_N)$ represents the effect of concentration on denitrification resulting from electron donor limitation if excessive N is transported into the water (Mulholland et al., 2008).

Alexander et al. (2004) proposed that the retention rate $\lambda_{ret,i}$ (year⁻¹, Eq. 2.11) in cell *i* is related to the net uptake velocity v_f and the depth (D_i, m) of water bodies. Based

on the in-stream retention R_i given by IMAGE-GNM, $\lambda_{ret,i}$ can be derived from a function of a natural logarithm of $(1 - R_i)$ and advection rate $\lambda_{adv,i}$:

$$
\lambda_{ret,i} = \frac{v_{f,i}}{D} = -\ln(1 - R_i) \cdot \lambda_{adv,i}
$$
\n(2.11)

2.2.1.3 Water consumption

Humans withdraw water from rivers, lakes, and reservoirs for irrigation, industrial production, and households. Some of the water withdrawal returns to the freshwater system, while the rest of the water is consumed, along with a net removal of N from freshwater (λ_{con}), which is not considered as an N output by IMAGE-GNM. Therefore, we introduce N removal from consumption of both surface water and groundwater into the removal rates. The removal rate due to water consumption $(\lambda_{con,i}, \text{year}^1, \text{Eq. 2.12})$ corresponds to a product of all fractions of water consumption $(f_{con,i},$ dimensionless, Eq.2.13) and the advection removal rate $(\lambda_{adv,i}, \text{year}^{-1})$.

$$
\lambda_{con,i} = \sum f_{con,i} \cdot \lambda_{adv,i} = (f_{agr,i} + f_{dom,i} + f_{elc,i} + f_{man,i} + f_{lvs,i}) \cdot \lambda_{adv,i} \quad (2.12)
$$

where $f_{agr,i}$, $f_{dom,i}$, f_{elci} , $f_{man,i}$, and $f_{lvs,i}$ (dimensionless) are the fractions of water consumption for agriculture, domestic, thermoelectric, manufacturing, and livestock use, respectively. The fraction of water consumption $f_{con,i}$ is the ratio between the volumetric extraction rate of water consumption $U_{con,i}$ (m³ year⁻¹) and the available water in the form of river discharge Q_i within a cell.

$$
f_{con,i} = \frac{U_{con,i}}{Q_i} \tag{2.13}
$$

Global agricultural water consumption data were obtained from Pfister and Bayer (2014). The domestic, industrial, and livestock water consumption data are from Flörke and Eisner (2011).

2.2.2 Aggregation of FFs

With the emissions of N applications (i.e. diffuse and point sources) and land use (through soil erosion) in any location quantified, the cumulative FFs can be used to predict the N fate at a half-degree resolution. Unlike Helmes et al. (2012) who weighted FFs based on population, we aggregated FFs by weighting according to the respective inventories for each emission route. We use the emission-weighting data of direct emissions to freshwater and diffuse emissions to the soil, while we weight erosion FFs using the areas of three land use types (Supporting Information Figures S2.1 to S2.5). These weighting data were given at the same spatial resolution and for the same representative year as the cumulative FFs. The impact of direct N emissions to freshwater and diffuse emissions to the soil over a region was assessed via emissionweighted FFs (Eq. 2.14):

$$
\text{FF}_{region,e}^{average} = \frac{1}{\sum_{i} E_{e \to i\in r}} \cdot \sum_{i} \text{FF}_{e \to i\in r} \cdot E_{e \to i\in r} \tag{2.14}
$$

The regional (e.g. country) average fate factor ($FF_{region}^{average}$, days) is used to represent the aggregation of FFs over a region *r*. $E_{e\rightarrow i\epsilon r}$ is the emission from diffuse or point source *e* in grid cell *i* (kg year⁻¹), provided by IMAGE-GNM.

Regional FFs of erosion $FF_{region, erosion}^{average}$ (days·kg N_{water} / (km²·year)) aggregate nonzero FFs of erosion over a region and all land-use types through area-weighting:

$$
FF_{region, erosion}^{average} = \frac{1}{\sum_{i} A_{i \in r}} \sum_{i, landscape} FF_{i \in r, erosion, landscape} \cdot A_{i \in r, landscape}
$$
 (2.15)

where $A_{i\in r}$ (km²) is the total land area of grid cell *i*; $A_{i\in r,$ landuse (km²) is the area of this land-use type; $FF_{i\in r, erosion, landscape}$ (days·kg Nwater / (km²·year)) is the individual FF of soil erosion of the site-specific land use in grid cell *i*.

2.2.3 Net removal rate

Helmes et al. (2012) developed a method to calculate the net removal rate to assess dominant processes for phosphorous persistence in freshwater, and we apply this method to N. For the advection process, the net removal rate $(k_{adv},$ dimensionless) can be calculated directly by excluding retention and water consumption processes in the reciprocal of the FF of freshwater (Eq. 2.16). However, during the calculation of the FF, advection cannot be omitted and thus the net removal rates for retention and water use were estimated indirectly as the difference between the overall net removal rate $(k_{all},$ dimensionless) and the net removal rate excluding the corresponding process (Eq. 2.17 and Eq. 2.18). Finally, the dominant process is determined by the largest net removal rate occupied in each cell.

$$
k_{adv} = \frac{1}{\text{FF}_{adv}} \tag{2.16}
$$

$$
k_{ret} = k_{all} - k_{noret} = \frac{1}{\text{FF}_{all}} - \frac{1}{\text{FF}_{noret}} \tag{2.17}
$$

$$
k_{con} = k_{all} - k_{nocom} = \frac{1}{\text{FF}_{all}} - \frac{1}{\text{FF}_{nocom}} \tag{2.18}
$$

2.3 Results

2.3.1 Global spatially explicit fate factors

The cumulative FFs [\(Figure 2.2\)](#page-18-0) show a distinctive spatial differentiation pattern over the globe. For instance, the fate factor of freshwater ($FF_{i, freshwater}$) has hotspots mainly located in North America, Central Asia, Russia, and Turkey; with high values also occurring in the east of South America, South Africa, East Asia, and East Europe. Furthermore, FF_{i, freshwater} has a considerable variability, as its 5th and 95th percentiles are 0.9 and 184.0 days, respectively. The hotspots and high values of the $FF_{i, freshwater}$ are distributed in large reservoirs and lakes and their upstream sources. For instance, in North America, the hotspots of the $FF_{i, freshwater}$ are distributed in the upper reach of Colorado River, at the upstream of Lake Powell and Lake Mead, together with Missouri River, at the upstream of Lake Sakakawea, Manicouagan reservoir, and Lake Oahe. In Asia, the hotspots are situated Lake Qinghai, Lake Baikal,

and the Keban Baraji reservoir. In Europe, the hotspots of FFs appear in North Europe, Spain, and Turkey. Low values of $FF_{i, freshwater}$ are commonlty situated near the coast.

The cumulative FFs of direct emission to freshwater, diffuse emission, and erosion show similar patterns. The 5th and 95th percentiles are 0.04 and 27.3 days·kg N_{water}/kg Nemission for $FF_{i,diffuse}$, and 5.2 and 2496.6 days·kg N_{water} / (km²·year) for $\overline{\text{FF}}_{i,erosion, natural}$.

Figure 2.2 Cumulative FFs (fate factors) for N emission to freshwater for $0.5^{\circ} \times 0.5^{\circ}$ grid cells. (a) diffuse sources excluding erosion, (b) baseline erosion on natural land, and (c) direct emissions to freshwater, including point sources. For erosion, the

difference of FFs between anthropogenic and natural erosion can be derived by multiplying with $c_{landuse}$. $c_{landuse}$ for arable land and grassland are 45.30 and 2.41, respectively.

2.3.2 Regional averages of fate factors

Since life cycle inventories are usually reported at the national level (e.g. ecoinvent, (Wernet et al., 2016)), we also analyze the regional average FFs to match that spatial scale [\(Figure 2.3,](#page-19-0) data can be found in Supporting Information Table S2.1). Generally, geographic regions that contain no large lakes or reservoirs tend to have lower cumulative FFs, and thus emissions from those regions typically have less impact on the regional FFs. Conversely, regions that have a large portion of lakes and reservoirs, tend to exhibit higher regional FFs. For instance, there are 5 geographic regions with a regional average FF of direct emission to freshwater larger than 100 days, while 39 geographic regions have an average FF lower than 3 days (Figure 2.3). As for continents, the regional average FFs of direct emission to freshwater varies from 20.0 days in Africa to 41.2 days in North America as calculated from aggregated values in Figure 2.3. The emission-weighted global average FFs for direct emission to freshwater is 29.3 days.

 33 **Figure 2.3** FF_{region} . This figure presents cumulative FFs (fate factors) of direct

emission of N to freshwater over geographic regions (e.g. country scale). Other regional FFs are given in the Supporting Information Table S2.1.

2.3.3 Dominant removal process for N fate

The dominant removal process for N transported to freshwaters differs across the globe [\(Figure 2.4\)](#page-20-0). In the northern hemisphere, retention dominates the cumulative FFs in most areas of North America, Eastern Europe, and Central Asia, while in the southern hemisphere, retention dominates in the eastern side of the continents. In contrast, advection is the main contributor in coastal areas as well as South America, northern and eastern Asia. Water consumption dominates in some water-deficient areas, e.g. Northern India and the Beijing-Tianjin Metropolitan Region in Northern China. Globally, advection is the largest net removal process, dominating 69.7% of the global area; while retention is the main removal process for 29.0% of the global area; and water consumption is the prevailing process in 1.3% of the global area. The global map of the contribution of each removal process can be found in the Supporting Information Figures $S2.6 - S2.8$.

34___ **Figure 2.4** Dominant processes of net removal rate for cumulative FFs (fate factors) on a global scale. Because FFs of different emission routes only vary in the fraction of N transported from the emission to water (f_{e-i}) which is irrelevant for removal rates,

dominant processes are analyzed based on FFs of freshwater.

The overall statistical distribution of cumulative FFs and their main drivers (residence time, discharge, and aridity index, whose global maps can be found in Supporting Information Figures S2.10 - S2.12), are grouped according to the corresponding dominant process for each grid cell: advection, retention, or water consumption in [Figure 2.5.](#page-22-0) The grid cells dominated by advection are mainly clustered in the interval of low residence time (the 95% quantile is 12.9 days) and their average residence time is 5.3 days. For these cells, the average discharge is $792.5 \text{ m}^3/\text{s}$, which is much higher than the $263.6 \text{ m}^3/\text{s}$ and $16.9 \text{ m}^3/\text{s}$ for cells dominated by retention and water consumption respectively. Advection is also the main contributor to all grid cells with high discharge ($\geq 10^4$ m³/s). High discharge and low residence time are the typical hydrological features of large rivers, especially near the mouth. Therefore, most of the grid cells controlled by advection are distributed in the river basins of large rivers. Retention is the most significant process in grid cells where the residence time is high (the average residence time in retention-dominated cells is 92.5 days) and discharge is low, which means that retention controls the removal process in lakes, reservoirs, and near the source. Grid cells dominated by water consumption have a firm relationship with the aridity index (AI), and they are all distributed in the low-AI zone (AI < 1.24), showing that water scarcity plays an important role in these regions. In particular, the AI of 97% of these grid cells is lower than 1, which indicates that evaporation is higher than precipitation.

The cumulative FFs of cells dominated by retention are high, as its average is 99.5 days, while the averages of the cells controlled by advection and water consumption are 20.1 days and 16.2 days respectively. These findings agree with Helmes et al. (2012), as they found that most of the retention-dominated cells have a high residence time and the cells dominated by water use are distributed in the arid zone.

Figure 2.5 Letter-value plots of cumulative FFs (fate factors), residence time, discharge, and aridity index of N for different dominant processes. The black line in each letter-value plot denotes the median of the data, and the black dot indicates the average value. The widest box is the range of approximate 1/4 to 3/4 quantile, the lower box of the second widest is the approximate 1/8 to 1/4 quantile, and the next lower box is the approximate 1/16 to 1/8 quantile, recursively (Heike et al., 2017). Note that the width of the box does not denote the probability density of element and its value is arbitrary.

The FFs spatial variability is thus largely driven by the local hydrological conditions, especially residence time and the discharge. For instance, according to GLWD1 and the World Lake Database (Herschy, 2012), Zambia contains six major lakes (surface area

 \geq 1000 km²) and three large water bodies (surface area ranging from 100 to <1000 km²), which favor long residence times, and it is a hotspot of regional FFs in our study. In contrast, its neighboring country, Mozambique, has an extensive river network with high discharges with only one major lake and four large lakes, which favor removal through advection and is in the lowest FF class. Examples of discharge driving FF spatial heterogeneity include the difference between the Amazon River $(209,000 \text{ m}^3/\text{s})$ versus the São Francisco River (2,943 m³/s) and Paraná River (17,290 m³/s) in South America. The higher advection rate in the Amazon River results in comparatively lower FFs. Similarly, river sub-basins tend to have lower discharge and thus a higher FF than the main downstream branch. An example of this latter phenomenon is the Missouri River, which, as an upstream branch of the Mississippi River, has a fraction of its discharge $(2,478 \text{ m}^3/\text{s} \text{ vs } 16,790 \text{ m}^3/\text{s})$, and thus a higher FF.

While this analysis shows that the cumulative FFs are highly related to the hydrological condition, the regional FFs also depend on the amount and location of the emission sources (e.g. synthetic fertilizer use). Due to the high residence time and low discharge in cells with large lakes and reservoirs, emissions from the nearby upstream to these cells increase the risk of N enrichment and persistence. For instance, agricultural emissions upstream of lakes and reservoirs (e.g. Lake Qinghai and Lake Baikal) may result in over 300 days of N persistence in the region, while fertilizing the same amount downstream of these lakes, reservoirs, or large rivers (e.g. Amazon River and Nile River) may only let N reside for 10 days, owing to high removal rates, thus causing lower eutrophication impacts.

2.4 Discussion

Our research calculates N FFs of three different emission routes at a half-degree spatial resolution and reveals the influence of hydrological conditions on N persistence, which affects the vulnerability of freshwater bodies to eutrophication. The local hydrological conditions depend on geological features, climate, and the presence of dams. For instance, large rivers with high discharge and low residence time always appear in humid regions with steep terrain. These areas thus tend to have low FFs, whereas a dam increases the residence time, hence the higher FFs in river basins with a dam. Emissions from anthropogenic sources (e.g. via industrial and agricultural activity) in those regions with high FFs may cause severe eutrophic impacts on downstream areas. Through our analysis of FFs, spatial patterns have for the first time been quantified for inland N at the sub-degree grided scale and build the foundation to allow LCA practitioners to assess the regional eutrophic impact of N over the globe.

We also highlight that FF temporal variations even in subsequent years can be quite substantial in urban regions with a massive population, as attested by the RSD of FFs between 1998 and 2000 (Supporting Information Figure S2.9). This reveals that the nutrient fate in freshwater systems is a dynamic process and reinforces the necessity of using dynamic models to derive FFs to complement current steady-state LCIA models.

2.4.1 Comparison with other models

This research builds on previous studies, and it provides FFs of inland N emitted both from the soil and directly to freshwater. Previous research of Cosme et al. (2018) extracted hydrological parameters from the Global NEWS 2 model (Mayorga et al., 2010), in which the residence time was also used to estimate denitrification, and constructed a FF model for the global coast. Cosme et al. (2018) aimed at modeling the persistence of dissolved inorganic nitrogen (DIN) in the receiving coastal large marine ecosystems, and also provided information of inland N fate at the watershed scale as a complementary result to N discharge toward the ocean. The basin-area-weighted riverine FF for DIN of Cosme et al. (2018) is 96 days, while Payen et al. (2021), who also applied NEWS2, calculated the global average of freshwater FF for DIN as 257 days. Compared to their results, our global average FF of TN (29.3 days) is lower. Hotspots partly agree, but not always. For example, both Payen et al. (2021) and we identified hotspots in the Mississippi River and Ob River, while these are not hotspots

according to Cosme et al. (2018); Payen et al. (2021) also identified hotspots in the Ganges River and the Hudson Bay, while these are not hotspots according to our FFs or the ones of Cosme et al. (2018). The discrepancy between our FFs and their FFs results from the difference in nitrogen indicators (DIN vs. TN), as well as the different mechanisms of nutrient models, methods of calculating FFs, and spatial delineations. For instance, van Vliet et al. (2019) showed that global TN export of NEWS2 (45 Tg year-1) is higher than that of IMAGE-GNM (37 Tg year-1) due to the difference in hydrological input data, spatial resolution, and the estimation of retention. Besides, the ratios between DIN and TN also exert large variation in different rivers. For example, the ratios have been found to be 50% for the Yangtze River (Yan et al., 2001) and 86% for the Mississippi River (Goolsby et al., 1999).

The model by Helmes et al. (2012) put forward an inland FF model for P for $0.5^{\circ} \times 0.5^{\circ}$ grid cells. Due to different hydrodynamic and biochemical processes, there is a clear difference between N and P cycles (e.g. N has a variety of redox forms, and undergo denitrification and exchange with the atmosphere, while most P in nature exists in solid or dissolved form). The difference between N and P is reflected in the retention, which in both cases is calculated based on regression methods. The retention rates of Helmes et al. (2012) are a fixed value in each interval (71.2 year⁻¹ if discharge < 0.0882 km³year⁻¹ ¹, 25 year⁻¹ if 0.4473< discharge <0.0882 km³year⁻¹, and 4.4 year⁻¹ if discharge > 0.4473 km³year⁻¹), while retention rates in our study are site-dependent, the average of which for these intervals are 92.1 year⁻¹, 32.4 year⁻¹, and 18.4 year⁻¹, respectively. Moreover, Helmes et al. (2012) did not consider domestic and industrial water consumption. Higher N retention rates together with domestic and industrial water consumption result in lower FFs in our study. Nonetheless, the distribution of low-to-high values of P cumulative FFs (Helmes et al., 2012) is consistent with our model on a global scale, especially for the hotspots in North America, Central Asia, and Turkey. Furthermore, our model's spatial differentiation of dominant removal processes is similar to Helmes et al. (2012).

2.4.2 Uncertainties

Current FF studies, including the one presented here, do not estimate sub-year variability, and thus ignore seasonal information. de Andrade et al. (2021) assessed the temporal and spatial variability of phosphorus FFs for freshwater in Bahia, Brazil, and concluded that FFs do not intensely vary monthly, although they recommended distinguishing two periods of higher and lower water availability. In contrast, their analysis suggests that FFs are highly site-dependent, thus it is important to regionalize eutrophication indicators. FFs in temporal regions, however, may be subject to much more pronounced seasonality, and thus the level of temporal variability on a global scale requires further study. Furthermore, in contrast to water consumption affecting only a few extreme grids cells, our results show that the cumulative FFs are related to hydrological features, retention, and other biogeochemical processes. Given that IMAGE-GNM and PCR-GLOBWB control these aspects, some uncertainties for these models are presented below.

2.4.2.1 Advection

The assumptions of the hydrological model PCR-GLOBWB introduce uncertainties in the estimation of advection. On the one hand, the reservoirs in PCR-GLOBWB are designated for hydropower generation and therefore it maximizes the available potential energy (Beusen et al., 2015), which can overestimate the real reservoir volume and could lead to an overestimation of FFs. On the other hand, PCR-GLOBWB divides multi-cell water bodies (i.e. lakes and reservoirs) by splitting the volume and combines multiple water bodies located within the same grid cell, ignoring the small water bodies if their total water volumes are lower than the volume of the river channel. This results in an underestimation of FFs due to an assumed lower water volume (774 out of a total of 6369 reservoirs were omitted in the year 2000) (Beusen et al., 2015).

Further improvement in simulating global gridded hydrological parameters in PCR-GLOBWB would provide a better assessment of eutrophication impacts.

2.4.2.2 Retention

The retention rate, as an argument in the inverse proportional function of FF, tends to have higher values when the water depth is underestimated (Eq. 2.11). Due to the proportional relationship between water volume and depth, the overestimated real reservoir volume in PCR-GLOBWB leads to an underestimation of the retention rate. Furthermore, the exclusion of small water bodies leads to an overestimation of the retention rate. In that case, FFs are inversely affected by the inaccurate estimation of retention removal rate. The empirical equation of Wollheim et al. (2006) is based on a first-order degradation process, assuming retention follows an exponential function of net uptake velocity and hydraulic load. However, there are also other options for empirical retention equations. For instance, Behrendt and Opitz (1999) and Venohr et al. (2005) assumed retention is a power function of surface water area, De Klein (2008) assumes that discharge plays a role in the retention process; while Seitzinger et al. (2002) only related the hydraulic load to retention. Empirical equations are limited in that they quantify retention ignoring chemical-mechanistic processes such as the interaction among different nutrient forms. Nevertheless, studies such as Vilmin et al. (2020) are increasingly incorporating mechanistic geochemical dynamics to better understand nutrient transport in the hydrosphere. With such information, N fate can be more precisely estimated by including the transformations among different N forms, including ammonium (NH_4^+) , nitrate (NO_3^-) , nitrite (NO_2^-) , and organic nitrogen, together with increasing the temporal resolution of the model (Vilmin et al., 2020).

2.4.2.3 The exclusion of sludge

As we mentioned in the methods section, calculating the cumulative FFs for N only relates to the denitrification process in the soil and the hydrological conditions of the

water. Hence, the exclusion of sludge in IMAGE-GNM does not influence the calculation of the cumulative FFs. Nevertheless, the exclusion of sludge might affect the aggregation of regional FFs by underestimating the emission-weighting data for direct emissions to freshwater. This impact on the regional FFs is difficult to generalize as overestimation or underestimation due to the uneven distribution of the sludge's share of emission-weighting data.

2.4.3 Potential variation under the climate change

Despite increasing retention, throughout the 20th century, more nutrients have been exported to the coast (Beusen et al., 2016). Further into the future, this trend is set to continue due to increasing use of fertilizer and increasing population and wastewater discharge (Mogollón et al., 2018a; van Puijenbroek et al., 2019). However, under a warmer climate, more evaporation can lead to an acceleration of the hydrological cycle, which may lead to a higher water advection rate and more nutrient transport. Together with the stronger advection rate, predicted additional water extraction from surface and groundwater (Wada and Bierkens, 2014) may counteract the effect of more intensive nutrient emissions. More research into future scenarios is required to assess future FFs.

2.4.4 Implications for LCIA modeling

LCIA methods seek to characterize the fate of human emissions. Cosme et al. (2018) have shown that FFs contribute much more to the spatial variability of CFs than exposure or effect factors, which demonstrates the importance of regionalizing especially the FFs, as presented here. The application of a gridded FF model may improve LCIA methods with regards to previous spatially resolved models, as it includes more details of intra-basin heterogeneities. Additionally, this work complements existing P-related LCIA models, and thus both the N and P fate can be used to better assess global eutrophication. Our analysis shows the strong relationship between FFs and N removal processes, which is crucial to designing more sustainable

site selections for N emitting activities and to raise awareness on the potential environmental impacts of globalized manufacture, trade, and consumption in terms of the N cycle. For such implementation in LCIA, the FFs can be aggregated from the original half-degree resolution to an arbitrary regional scale by weighing according to the emissions, or in case of erosion, using the land use area. This will allow LCA practitioners to obtain the final fate for nutrients emitted during production in any region matching their inventory data.

2.5 Conclusion

We introduced N into the assessment of the environmental impacts on the global freshwater system as a co-limiting nutrient for eutrophication to complement present analyses based on P. Our spatially explicit approach provides global FFs of nitrogen for grid-based emissions both from the soil and directly to freshwater systems. Moreover, our study emphasizes the quantitative analysis of the connection between hydrological conditions and FFs. Our study revealed that FFs show conspicuous spatial heterogeneity because of differences in hydrological conditions and provided regionalized FFs which serve as midpoint indicators and can help LCA practitioners choose more sustainable production sites or suppliers.

Supporting Information

Figure S2.1 Global emission-weighting data of direct emission of N to freshwater

Figure S2.2 Global emission-weighting data of N of diffuse source to the soil

Figure S2.3 Global natural land area

Figure S2.4 Global grassland area

Figure S2.5 Global arable land area

Figure S2.6 Global contribution of advection

Figure S2.7 Global contribution of retention

Figure S2.8 Global contribution of water consumption

Figure S2.9 Relative standard deviation (RSD) of FFs between 1998 and 2000 (RSD is dimensionless shown as (-) in the legend.)

Figure S2.10 Global discharge

Figure S2.12 Global aridity index (AI, dimensionless shown as (-) in the legend.)

Table S2.1 Coutry-level fate factors (FFs)

