Global riverine nitrous oxide emissions: the role of small streams and large rivers

Alessandra Marzadri[a,*], Giuseppe Amatulli[b,c,**], Daniele Tonina[d], Alberto Bellin[a], Longzhu Q. Shen[e], George H. Allen[f], Peter A. Raymond[b]

[a]*Department of Civil, Environmental and Mechanical Engineering, University of Trento, Trento, 38123, Italy*
[b]*School of Forestry and Environmental Studies, Yale University, New Haven, CT 06520, USA*
[c]*Center for Research Computing, Yale University, New Haven, CT 06520, USA*
[d]*Center for Ecohydraulics Research, University of Idaho, Boise, ID 83702, USA*
[e]*Spatial Ecology, 35A, Hazlemere Road, Penn, Bucks HP10 8AD, UK*
[f]*Department of Geography, Texas A&M University, College Station, Texas 77843-3147, USA*

Abstract

Nitrous oxide, N_2O, is the leading cause of stratospheric ozone depletion and one of the most potent greenhouse gases (GHG). Its concentration in the atmosphere has been rapidly increasing since the green revolution in the 1950s and 1960s. Riverine systems have been suggested to be an important source of N_2O, although their quantitative contribution has been estimated with poor precision, ranging between 32.2 and 2,100 $GgN_2O - N/yr$. Here, we quantify reach scale N_2O emissions by integrating a data-driven machine learning model with a physically-based upscaling model. The application of this hybrid modeling approach reveals that small streams (those with widths

*Corresponding author
**Corresponding author
Email addresses: alessandra.marzadri@unitn.it (Alessandra Marzadri), giuseppe.amatulli@yale.edu (Giuseppe Amatulli)

less than 10 m) are the primary sources of riverine N_2O emissions to the atmosphere. They contribute nearly 36 $GgN_2O - N/yr$; almost 50% of the entire N_2O emissions from riverine systems (72.8 $GgN_2O - N/yr$), although they account for only 13% of the total riverine surface area worldwide. Large rivers (widths wider than 175 m), such as the main stems of the Amazon River (~ 6 $GgN_2O - N/yr$), the Mississippi River (~ 2 $GgN_2O - N/yr$), the Congo River (~ 1 $GgN_2O - N/yr$) and the Yang Tze River (~ 0.7 $GgN_2O - N/yr$), only contribute 26% of global N_2O emissions, which primarily originate from their water column. This study identifies, for the first time, near-global N_2O emission and NO_3 removal hot spots within watersheds and thus can aid the development of local- to global-scale management and mitigation strategies for riverine systems with respect to N_2O emissions. The presented framework can be extended to quantified biogeochemical, besides N_2O emissions, processes at the global scale.

Keywords: Nitrous oxide emissions, River networks, Hybrid model, Global scale

1. Introduction

Streams and rivers are natural receptors of solute-laden water delivered by the hydrological cycle. Human activities have had a widespread impacts on Earth's streams through flow and morphology modifications, land use and land cover changes, as well as water quality alteration [54, 23, 49, 59], exemplified by augmented reactive Dissolved Inorganic Nitrogen (DIN mainly in the form of ammonium, NH_4^+, and nitrate, NO_3^-) loads to otherwise nutrient limited riverine environments [24, 67, 62] due to fertilizers used for

agricultural production. Loading of nitrogen into lotic fresh waters also caused supersaturation of nitrous oxide, N_2O, with respect to the atmosphere [68, 15, 9, 12, 74]. As a GHG, N_2O is long-lived, with atmospheric residence time >100 yr, and has a warming capacity 300 times higher than carbon dioxide (CO_2) [34]. Nitrous oxide is also the leading gas that depletes stratospheric ozone, resulting in an increasing amount of harmful solar ultraviolet radiation reaching the Earth's surface [17, 58]. In streams and rivers, N_2O emissions are mainly controlled by microbially-mediated nitrification and denitrification processes [49, 9]. Denitrification, a biogeochemical process that converts NO_3 to N_2O and N_2 (dinitrogen) under anaerobic conditions, is considered the major pathway responsible for N_2O emissions to the atmosphere [42]. These processes occur within the surface (e.g. water column and benthic zone) and subsurface (e.g. hyporheic zone, the shallow groundwater system mainly saturated with stream water) environments of the riverine corridor [49, 9, 43]. This causes that estimating N_2O emissions from riverine systems has been challenging and somewhat elusive with global predictions spanning nearly three orders of magnitude [32, 42, 57]. This high uncertainty is mainly due to the inherent dependence of N_2O emissions on stream morphology, hydraulics (among and within the different riverine environments) and water quality changes, which are driven by regional land use, biomes and climatic conditions [66, 9, 12] and hyporheic processes. All of these quantities vary along river networks, with highly complex and interdependent relationships. Due to this system complexity, traditional regression methods that are based on relatively few emission measurements, are unable to accurately estimate the N_2O emission variability. This is reflected by the high variabil-

ity of N_2O emission factors (EFs) used by the IPCC [34] to estimate N_2O emissions from concentrations or loads of some (e.g. nitrate, ammonium or DIN) or all (Total Nitrogen, TN) compounds of nitrogen. The EF approach has limited accuracy because N_2O emissions from riverine systems do not depend only on available reactants (nitrate, ammonium or TN), but they also are highly dependent on stream hydromorphology and hyporheic processes. Some information of surface hydraulics have started to be included in recent N_2O global models [42, 77, 72] but benthic and hyporheic contributions are not well characterized in these models although they may be key sources depending on stream size [43]. Advances on hyporheic modeling have provided relationships that link its processes to stream hydromorphological characteristics such that the shallow subsurface contribution of N_2O could be predicted from reach-scale hydromorphologic data. Efforts to spatially constrain these hydromorphological variables worldwide have accelerated recently [60, 2] with an increasing contribution from machine learning methodologies (hereafter ML) in environmental applications (e.g. artificial intelligence and deep learning algorithms) that enabled the generation of high-resolution data with a large spatial extent, such as the prediction of annual stream flow [5, 40] and freshwater-specific variables [18]. Despite these encouraging results, the use of available ML algorithms to directly predict N_2O emissions from riverine networks may lead to poor forecasts due to the lack of hydromorphological and biogeochemical descriptors and the scarcity of available N_2O data.

Here we quantify 1) the global riverine N_2O emissions and 2) their spatial distributions along riverine networks at reach scale ($\sim$ 1 km length) resolution, which within the watershed context is important for water quality

management as it allows identifying N_2O emissions hot spots by accounting for N_2O production within both surface and subsurface environments of the river corridor. We proposed a novel hybrid framework that leverages the strengths of ML in spatially and temporally distributing hydromorphological quantities at the reach scale and a physically-based scaling model that accounts for in-stream, benthic and hyporheic processes and is fully informed by the ML modeled quantities to predict N_2O emissions. Our approach combines a data-driven Random Forest (hereafter RF) ML model [14] with the three-equation model for reach-scale riverine N_2O emissions proposed by Marzadri et al. [43, 44]. The data-driven model produces hydrologic and water quality information along the HydroSHEDS river network [38] at 30 arc-second grid-cell resolution (representing "the stream reach unit" within $\sim$ 1 km cell size at the equator). This results in the finest-resolution of spatial N_2O emissions to date [6, 12]. A strength of data-driven models [70, 5] is the use of stream-variable predictors quantified with a hierarchical-nesting approach [18, 19] that accounts for the influence of up-stream environmental conditions, (e.g. land cover information on agriculture and urban areas) carried down onto the downstream water quantity (discharge) and quality (nutrient concentration). The data-driven models use gauging station observations to provide distributed estimates of: (1) mean annual water discharge (from the FLO1K discharge dataset [5]) and (2) in-stream DIN species concentrations [18]. We applied hydromorphological equations onto a discharge and topographic dataset to quantify stream features at each stream network grid-cell, such as riverine morphology (median grain size, averaged bedform type and size [47, 37]), flow width, depth and velocity (with regime equations

[62]) (see Materials and Methods and Appendix sections for a detailed description) to inform the physically-based upscaling model of N_2O emissions developed by Marzadri et al. [43, 44]. The three-equation model captures local scale physical and chemical processes responsible for N_2O emissions within the hyporheic, benthic and water column zones and upscale them to predict reach-scale N_2O mass fluxes. Its advantage is that it does not require parameter fitting and it can be applied to any stream and river by knowing their hydromophologic and chemical quantities, measured in the field or estimated similarly to our framework.

2. Materials and Methods

The N_2O emission model is fully characterized by the following reach-scale hydromophological quantities: depth, velocity, width, bedform type and size, median grain size, streambed hydraulic conductivity, water quality DIN concentrations and the denitrification uptake rate constant (v_{fden}) [43]. All these variables depend on land-use and land cover, climatic and biomes conditions that inform the N_2O model predictions. They may not be available at the global scale, but they can be reduced to a sub-set of key variables, which are discharge (both the discharge of interest and bankfull discharge), and streambed slope for hydromorphology and DIN concentrations for water quality (Table 1). These necessary and sufficient input data were predicted with the data-driven models (Subsection 2.2). The other input variables, whose statistics (minimum, mean and maximum values) are summarized in Table 2, were derived from these basic inputs data with available hydromorphological and biochemical relationships (Subsection 2.4).

Grid-cell resolution of the data-driven models was of 30 arc-second, thereby resulting in squared cells with side of $1\,km$ at the Equator and each cell was uniquely identified by the georeferenced position of its center. Each grid's cell is assumed to contain a reach, which sets the finest resolution of the present investigation. This resulted in 16,450,188 grid-cells (or riverine reaches) composing the near-global river network from 60^oN to 5^oS latitude and 145^oW to 180^oE longitude from N_2O emissions from estimated.

We applied the model with mean annual discharge. This is not by any means a limitation of the model and finer temporal resolutions (e.g. seasonal, monthly, weekly or daily) may be used when data at the proper temporal resolution are available. However, given the global nature of the estimates proposed in the present work, and in line with previous evaluations (see e.g. Hu et al. [32], Maavara et al. [42]), we considered the annual scale a proper modeling choice.

Table 1: Sufficient and necessary model input parameters to capture N_2O emissions from streams and rivers at different spatial scales.

Parameter description	**Units**	**References**
Lat : cell latitude WGS84	[Decimal Degree]	-
$Long$: cell longitude WGS84	[Decimal Degree]	-
Q_{max} : maximum water discharge	[m^3/s]	Barbarossa et al. [5]
Q : water discharge	[m^3/s]	Barbarossa et al. [5]
S : stream slope	[$-$]	Amatulli et al. [3]
$[NO_3^-]$: nitrate concentration	[μmol/L]	Shen et al. [70]
$[NH_4^+]$: ammonium concentration	[μmol/L]	Shen et al. [70]

Table 2: Main statistics (minimum, mean and maximum) of model input parameters derived by hydro-morphological relationships. (W is the channel width, D is the mean flow depth, V is the stream velocity, W_{bf} is the bankfull width, D_{bf} is the bankfull depth, S is the stream slope, d_{50} is the median grain size of streambed sediments, K_h is the streambed hydraulic conductivity, L_{st} is the stream length, SA is the channel surface area, v_{fden} is the uptake rate of denitrification, τ_{50} is the median hyporheic residence time, τ_D is the time of denitrification evaluated according to Mulholland et al. [49] and Böhlke et al. [11], respectively. t_m is the time of turbulent vertical mixing, Da_{DHZ} and Da_D are the denitrification Damköhler numbers from the benthic-hyporheic zone and the water column, respectively and $FDIN$ is the in-stream flux of DIN species).

Parameter	**Minimum**	**Mean**	**Maximum**	**Reference**
$W\,(m)$	0.55	20.55	2376.12	Raymond et al. [62]
$D\,(m)$	0.05	0.48	15.29	Raymond et al. [62]
$V\,(m/s)$	0.02	0.22	6.20	Raymond et al. [62]
$W_{bf}\,(m)$	1.38	43.26	2894.13	Raymond et al. [62]
$D_{bf}\,(m)$	0.09	0.83	17.53	Raymond et al. [62]
$S\,(-)$	2.16×10^{-5}	0.037	0.568	Amatulli et al. [3]
$d_{50}\,(m)$	1.0×10^{-5}	8.08×10^{-2}	0.300	Lee and Julien [37]
$K_h\,(m/s)$	1.97×10^{-4}	1.01×10^{-2}	3.70×10^{-2}	Salarashayeri et al., [65]
$L_{St}\,(m)$	1017.11	1295.08	1438.32	Calculated
$SA\,(m^2)$	776	27000	2.83×10^{6}	Calculated
$v_{fden}\,(m/s)$	1.319×10^{-7}	5.757×10^{-7}	2.281×10^{-6}	Mulholland et al. [49]
				Böhlke et al. [11]
$\tau_{50}\,(s)$	9.05	4.25×10^{5}	8.43×10^{6}	Elliott and Brooks [20]
				Marzadri et al. [45]
$\tau_D\,(s)$	0.00	8.89×10^{5}	1.85×10^{7}	Mulholland et al. [49]
				Böhlke et al. [11]
$t_m\,(s)$	2.46	33.1	14400	Rutherford [64]
$Da_{DHZ}\,(-)$	1.24×10^{-5}	0.861	61.7	Marzadri et al. [43]
$Da_D\,(-)$	1.50×10^{-6}	3.86×10^{-5}	4.41×10^{-3}	Marzadri et al. [43]
$FDIN\,(\mu mol/m^2/h)$	1.07×10^{6}	4.31×10^{7}	2.16×10^{9}	Calculated

2.1. N_2O emission model

We selected the three-equation model proposed by Marzadri et al. [43, 44] to predict the dimensionless N_2O emissions flux (F^*N_2O):

$$F^*N_2O = \begin{cases} F^*N_2O_{HZ} = 1.55 \times 10^{-7}(\mathrm{Da_{DHZ}})^{0.43},\ Zone\,1:\ W \leq 10m \\ F^*N_2O_{BZ} = 1.91 \times 10^{-8}(\mathrm{Da_{DHZ}})^{0.58},\ Zone\,2:\ 10m < W \leq 175m \\ F^*N_2O_{WC} = 4.56 \times 10^{-6}(\mathrm{Da_D})^{0.72},\ Zone\,3: W > 175m \end{cases} \tag{1}$$

as a function of two Damköhler numbers, one for the hyporheic and benthic zones, Da_{DHZ}, and one for the water column, Da_D, given by the ratios between the characteristic times of solute transport and denitrification (see subsection 2.4 and Appendix A). Equations (1) account for the shift in the main source of N_2O from the hyporheic zone to the water column as the reach size increases from small streams (stream widths, W, less than 10 m, Zone 1) to intermediate (W =10-175 m; Zone 2) and large rivers (W >175 m, Zone 3). The model accounts for the lumped effect of nitrification and denitrification processes within the river corridor (e.g., hyporheic, benthic and water column zones) and for reactants, dissolved reactive nitrogen (DIN), availability due to climatic, hydrological, and land use variables from any source to the reach. It does not consider N_2O contributions transported to the river from large scale (valley and deep aquifer) groundwater circulation. Accounting for direct groundwater input is challenging as N_2O is not a final gas but it can be consumed to N_2 depending on residence time and biogeochemical characteristics of the groundwater path [57].

The model was validated on more than 1100 reaches across the world (See

Appendix B). Reach scale mass flux, FN_2O is quantified by multiplying F^*N_2O by the reach DIN mass flux ($FDIN = VDIN$, where V is the mean flow velocity). The morphological parameters needed to evaluate Da_{DHZ} and Da_D are presented in the following sections.

Total N_2O emissions (TE_{N2O}, ($kg\,N_2O - N/yr$)) are then computed as follows:

$$\begin{aligned} TE_{N2O} &= \sum_{i=1}^{NC} ERN2O, i = \\ & c \sum_{i=1}^{NC} FDIN_i \cdot F^*N_2O_i \cdot W_i \cdot L_{St,i} = \\ & c \sum_{i=1}^{NC} FN_2O_i \cdot SA_i \end{aligned} \tag{2}$$

where $FDIN_i$ ($\mu mol\,m^{-2}h^{-1}$) is the in stream DIN flux per unit stream surface, $ER_{N2O,i}$ is the N_2O emission rate ($kg\,N_2O - N/yr$) from the $i-th$ reach (see Figure 9), and W_i (m) and $L_{St,i}$ (m) are the stream width and length, respectively. The product between W_i and $L_{St,i}$ represents the surface area SA_i (m^2) of the $i-th$ reach. The summation is extended to all the NC streams identified in the analyzed portion of the Earth's surface (e.g., a catchment, a continent or part of the globe), and $c = 2.453 \times 10^{-4}$ is a units transformation factor.

2.2. Data-driven models

2.2.1. Hydrologic data

The FLO1K dataset [5] reports minimum, mean and maximum annual discharge (m^3/sec) at 1 km resolution from 1960 through 2015 (3 parame-

ters x 55 year-layers). We calculated the overall mean from the yearly mean (identified as for Q, representing the mean water discharge) and overall maximum from the yearly maximum (Q_{max}). The latter was used to represent the bankfull discharge $Q_{max} = Q_{bf}$, which was used as an index for riverine morphological characterization, and bedform size estimation for gravel bed rivers. We overlaid the stream network HydroSHEDS to obtain the mean and maximum overall discharge for each grid cell. The HydroSHEDS network was filtered by removing streams located in the major desert regions [52] or having a FLO1K mean discharge [5] smaller than $\approx 6 \times 10^{-4}\ m^3/sec$ (0.6 liter per second) (see Supplementary Figure S1). These streams are most likely ephemeral, and thus rarely contribute to the emissions of N_2O. In fact, the spatial distribution of Q reflects the global climatic characteristics with small values in correspondence of the most arid regions of the world where potential evapotranspiration exceeds precipitation and high values in temperate areas at medium latitude where precipitation is larger than potential evapotranspiration (Supplementary Figure S1).

2.2.2. DIN data

We developed a data-driven model to estimate yearly mean concentration of ammonium, NH_4^+, and nitrate, NO_3^-, along the HydroSHEDS river network [38] at each 30 arc-second cell, which provided the near global spatial distribution of DIN (see Figure 1). The model was built upon a RF framework [14] using: 1) the observed concentration values from the GLObal RIver CHemistry database (GLoRiCh) [28] and the Water Quality Portal (WQP - `https://www.waterqualitydata.us/portal`) [63], and 2) the stream-variables that account for up-stream environmental conditions

[18]. Full details of model calibration and validation are reported in Shen et al. [70].

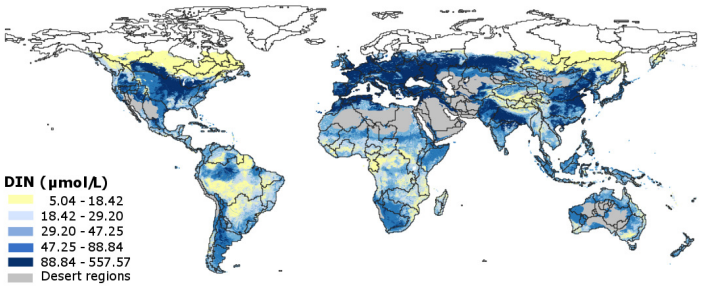


Figure 1: Map of average annual Dissolved Inorganic Nitrogen (DIN) concentration distribution obtained by the data-driven (Random Forest) model [70]. DIN map obtained as the combination of the in-stream load of ammonium and nitrate DIN = $[NH_4^+] + [NO_3^-]$, (μmol/L).

To balance geographical distribution of observations between the WQP and GLoRiCh databases, we considered only WQP gauge stations with more than 20 observations per station. Based on the analysis of the data-driven model, we found the estimated ammonium and nitrate concentrations are driven by the land-cover and the environmental conditions of the upper catchments. The RF model prediction was less accurate for NH_4^+ than for NO_3^- concentrations. The Pearson coefficients between the model prediction and measurement at gauge station are 0.3 and 0.8 for NH_4^+ and for NO_3^-, respectively. The errors on NH_4^+ caused small N_2O prediction errors because NO_3^- concentration is typically one order of magnitude larger than that of NH_4^+ given the predominance of the former with respect to the latter in synthetic fertilizers (see Results and Discussion section for details).

2.3. Streambed slope

The streambed slope (S), which has key riverine hydro-mophological implications, is estimated by superimposing the stream network to the 1 km slope map derived from aggregation-mean of the Geomorpho90m dataset slope [3] (see Supplementary Figure S2).

2.4. Reach scale stream hydromorphology and streambed hydraulic conductivity

Reach-scale mean width W (m), depth D (m), and velocity V (m/s) were estimated as a function of the mean annual discharge, Q (m^3/s) (see maps in Supplementary Figures S3, S4 and S5, respectively), with the power law relationships proposed by Raymond et al. [62]:

$$\begin{cases} W = 12.936 \cdot Q^{0.423} \\ D = 0.408 \cdot Q^{0.294} \\ V = 0.194 \cdot Q^{0.285} \end{cases} \tag{3}$$

The resulting D and V are used in the computation of Da_{DHZ} ($= \tau_{50}/\tau_D$) and Da_D ($= t_m/\tau_D$) that require the evaluation of the characteristic time of denitrification, τ_D, and the median hyporheic residence time, τ_{50}, or the time of turbulent vertical mixing, $t_m = D/(0.067\sqrt{gDS}$ assuming uniform flows and shallow water hypothesis [64]. On the other hand, τ_{50} depends on stream hydro-morphology and is evaluated with the formulation proposed by Elliott and Brooks [20] and Marzadri et al. [45] for reaches with dune-like and pool-riffle morphologies, respectively (see Appendix A). For small stream reaches characterized by step-pool or cascade morphology, the pool-riffle hyporheic model [31, 21, 30] was used when deemed applicable; reaches

with inconsistent or undefined morphology, for which there are no hyporheic flow models to provide a robust estimate of τ_{50}, are disregarded (note that these streams represent the 0.3% of the total stream network).

Stream morphology is classified according to Montgomery and Buffington [47] as a function of the median grain size d_{50} and stream slope S. The former (d_{50}) is estimated according to the formulation proposed by Lee and Julien [37] (see Appendix C section and Supplementary Figure S6). Note that the model used to estimate d_{50}, allows us to characterize, for the first time, the near-global distribution of the median grain size of streams and rivers sediments. This is per se an important result considering the important role of hyporheic exchange in controlling N_2O emissions along river networks.

Information of the d_{50} allowed us to map the estimation of the streambed hydraulic conductivity (K_h, $[m/d]$). This was done with the expression provided by Salarashayeri and Siosemarde [65] which provides K_h as a function of streambed d_{50} (see Appendix Figure C2 and the near-global spatial distribution of K_h in Supplementary Figure S7).

Another important input parameter of the N_2O emission model is the stream surface area (SA, m^2) that represents the surface through which gasses exchange between water and air. Supplementary Figure S8 shows the near-global spatial distribution of SA, which depends on W and the steam length (L_{St}) spatial distribution. The layers (FLO1K, HydroSHEDS, stream-variables, etc.) are represented under the Plate Carrée projection on a WGS84 Ellipsoid (EPSG 4326). Any kind of quasi-spherical surface represented in 2D (map format) can lead to a distortion. The distortion is produced when the spherical coordinates of WGS84 are treated as a rectilin-

ear grid (e.g. Plate Carrée or Universal Transverse Mercator projections). It increases with a latitudinal gradient, with very high areal/length distortion values in the Sub-Arctic/Antarctic zones. To account for this distortion, the length L_{St} of the cell diagonal is computed as a function of the effective cell area (A_{cell}, see equation (4)). Hence, stream length is larger for the cells at the Equator compared to those in both Northern and Southern hemispheres. Moreover, stream sinuosity is accounted for by multiplying L_{St} by a sinuosity coefficient of 1.286 [22].

The resulting near-global spatial distribution of L_{St} with latitude reflects in the estimation of SA (and consequently affects N_2O emissions) as for the same width, a channel in the Equator region is characterized by a higher SA than a channel far from the Equator region [22]. To obtain consistent values of $SA = L_{St} \times W$ we use the following expression for L_{St} as a function of W:

$$\begin{cases} L_{St} = 0.5 \cdot (\sqrt{2}+1)\sqrt{A_{cell}}, \ if\, W \leq 1000\, m \\ L_{St} = \sqrt{A_{cell}}, \quad if\, W > 1000\, m \end{cases} \tag{4}$$

2.5. Reaction rate constant

For each reach, τ_D, the inverse of the reaction rate constant for denitrification, is evaluated as the ratio between water depth D provided by equation (3) and the uptake rate of denitrification (v_{fden}, m/s), which is estimated according to the following two formulations depending on the reach size:

$$\begin{cases} v_{fden} = 2.7 \times 10^{-7}(10.4[NO_3^-]^{-0.493}) \ \ W \leq 175m \\ v_{fden} = 2.7 \times 10^{-7}(17.0[NO_3^-]^{-0.490}) \ \ W > 175m \end{cases} \tag{5}$$

where $[NO_3^-]$ ($\mu mol/L$) is the in-stream nitrate concentration [18]. The first equation was derived by Mulholland et al. [49] for small headwater streams

surveyed in the Second Lotic Intersite Nitrogen Experiment (LINXII) and is used here for reaches in the Zones 1 and 2. Whereas the second one was proposed by Böhlke et al. [11], who combined and weighted data from the LINXII streams with those collected in the Upper Mississippi River basin, and is used for reaches in the Zone 3. According to these formulations the total v_{fden} from riverine systems is equal to 9.3921 m/s that correspond to 2.58 TgN/yr denitrified (see also the main statistics in Table 2 and the pattern of variation in Supplementary Figure S9).

2.6. DIN flux

The interaction between stream hydrology and biogeochemistry controls the near-global spatial distribution of the annual mean in-stream mass flux of DIN ($FDIN$, $mol\,m^{-2}h^{-1}$) (2). $FDIN = V([NH_4] + [NO_3])$ takes into account both the effects of stream hydraulics (through V) and land use and cover (through DIN whose near-global spatial distribution is shown in Figure 1 and is responsible to a total DIN load to riverine systems of 22.89 TgN/yr).

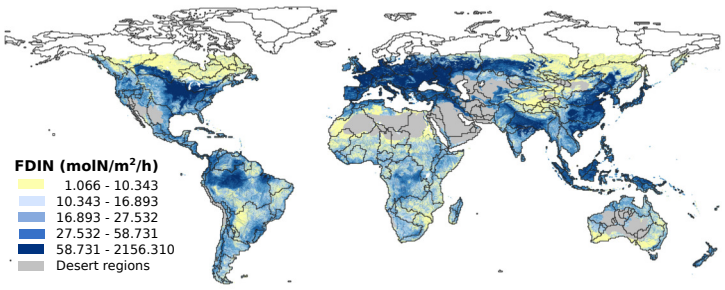


Figure 2: Map of the annual average of in stream mass Flux of Dissolved Inorganic Nitrogen ($FDIN$, $mol\,m^{-2}h^{-1}$) species according to data of DIN concentration calculated following the methodology described in Shen *et al.* [70] and the data of stream velocity [62].

3. Results and Discussion

3.1. Global distribution of denitrification Damköhler numbers

The hydromophological and water quality information quantified globally at reach-scale and annual resolution in the previous section allows mapping the distribution of riverine denitrification potential expressed through the denitrification Damköhler numbers (Figure 3). Their values are chiefly less than 1 (Figure 3a), which suggests that reaction rate is the main limiting factor rather than advection [29]. As expected due to the relatively small time scales of the vertical turbulent mixing, this is especially evident in large rivers where the water column is the main source of N_2O emissions (c.f., Figures 3b and 3c) than in small and intermediate streams where advection in the hyporheic and benthic zones defines larger transport time scales (Figure 3b). Values of Da_{DHZ} larger than 1 generally occur in areas with small streams (Figures S1, S3, S4 and S5 in Supplementary materials) in gentle sloped valleys (Figure S2 in Supplementary materials) and with fine sediments (Figure S6 in Supplementary materials). These hydro-morphological characteristics lead to low streambed hydraulic conductivity (Figure S7 in Supplementary materials), which in turn causes large hyporheic time scales and so large Da_{DHZ} values.

3.2. Global distribution of dimensionless N_2O *flux:* F^*N_2O

The application of the 3-equation N_2O model (equation (1)) informed by the Damköhler number spatial distributions provides the estimation of N_2O dimensionless flux F^*N_2O, which can be considered an equivalent of the IPCC Emission Factor (EF) [34]. However, contrarily to the IPCC EF,

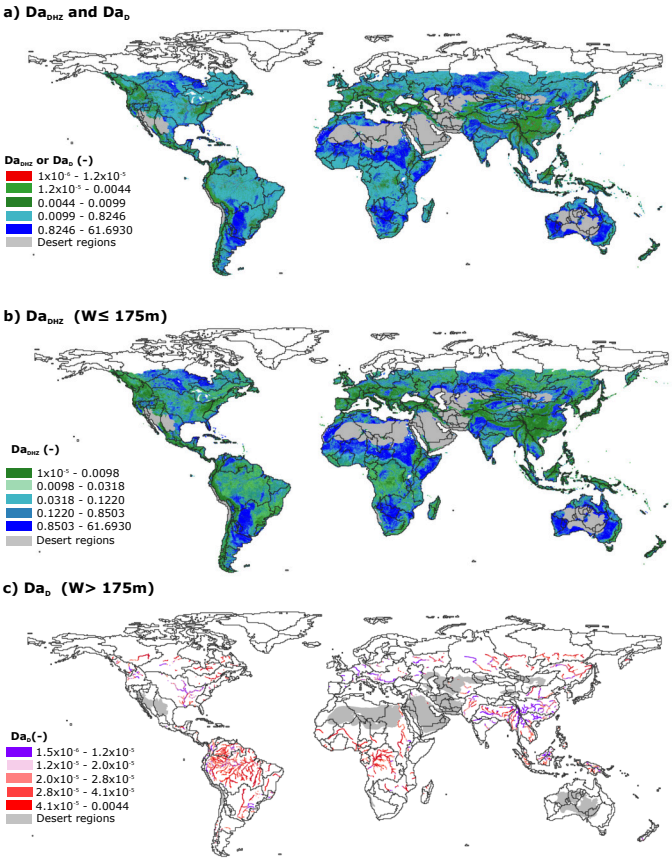


Figure 3: Map of the denitrification Damköhler numbers Da_{DHZ} and Da_D from: (a) the near-global river network, (b) the sediments (e.g. benthic and hyporheic zone Da_{DHZ}) and (c) the water column (Da_D) of the near-global river network.

which typically quantifies N_2O emissions as a percent of the reactants concentrations or fluxes (e.g., DIN, total nitrogen or nitrate), our proposed EF, F^*N2O, accounts for hydromorphological variations through Da_{DHZ} and Da_D and thus it is spatially and temporally variable. Values of F^*N_2O changes along the river network ranging 3 orders of magnitude from 10^{-10} to 10^{-7} (see Figure 4 and Table 4). High F^*N_2O values are chiefly located in small streams and their values decrease as the stream size increases (Figure 4). In small streams, F^*N_2O originates from the hyporheic and benthic zones, whose role in controlling the total emissions reduces in favor of the water column, as the stream size increases. This is why dark blue tones are located close to desert areas, where streams are small, and light blue and yellow tones along the main stems of the major river networks of the world (e.g. Amazon, Mississippi, Congo, Yang-Tze, etc.).

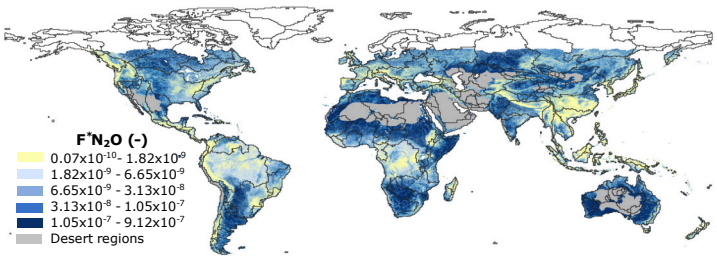


Figure 4: Map of dimensionless N_2O flux (F^*N_2O) along the world river network analyzed. Gray areas represent desert regions not accounted in the calculation.

3.3. Global distribution of N_2O *flux per unit area:* FN_2O

While DIN concentrations are important to explain N_2O emissions, they are not the only key variable. Accordingly, the pattern of variation of N_2O

fluxes per unit stream area, FN_2O, do not show a consistent association with DIN concentrations (compare Figure 5 and 1).

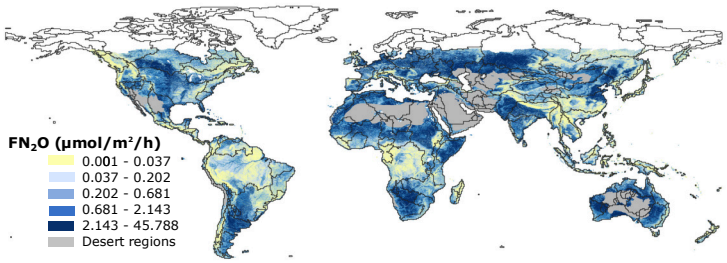


Figure 5: Map of N_2O flux per unit area (FN_2O, $\mu mol\, m^{-2} h^{-1}$). Gray areas represent desert regions not accounted in the calculation.

Focusing on the Mississippi River basin, very high FN_2O occurs in small streams even though their DIN concentrations are smaller than that of the main stem (Figure 6). Similarly, within the Congo River basin, FN_2O does not correlate with DIN concentration distributions. Conversely in other systems, like the Danube, Ganges and Yang-Tze, intensive agricultural practices produce DIN concentrations that match the areas with high FN_2O. This different behavior is in line with the loading capacities observed among the different regions in the world due to their different climatic and hydrological conditions and the associated fertilization strategies and practices [51]. Regardless of the location, N_2O emissions are modulated by the local hydromorphology. Accordingly, FN_2O reach-scale distribution has higher values in zones dominated by small streams than large rivers and FN_2O values decrease as the stream/river dimension increase (c.f., the pattern of channel width, W, in Supplementary Figure S3 and Figure 5). This behavior is in line with data of N_2O emissions from streams and rivers worldwide that shows a

reduction in FN_2O as the average channel width increases [26, 43].

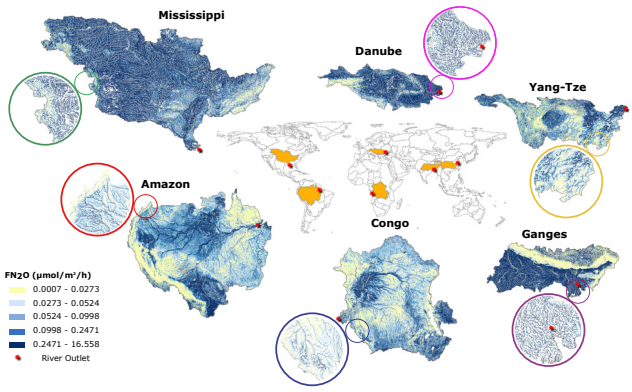


Figure 6: Detailed maps of N_2O emission from selected catchments. Spatial patterns of the N_2O fluxes per unit stream area (FN_2O) from several large river basins (fine-scale zoom-in shown in the circled areas). Catchment N_2O total emissions are available in Table 5.

The previous N_2O regression-based model of Hu et al. [32] provides similar global distribution of N_2O flux to our model because it accounts for spatial distribution of the emissions factors (Figure 7 and Table 3). However, our model provides estimates at much higher spatial resolution than any previous modeling attempt. More importantly, F^*N_2O values, our EF equivalent, are physically estimated from hydromorphological and water quality characteristics and are not regression coefficients like in previous works.

3.4. Global distribution of N_2O *emission rate:* ER_{N2O}

The high resolution of our proposed hybrid model maps global N_2O emissions at an unprecedented resolution (Figure 8), which illustrates the role of the river size and the importance of the river surface area in controlling the

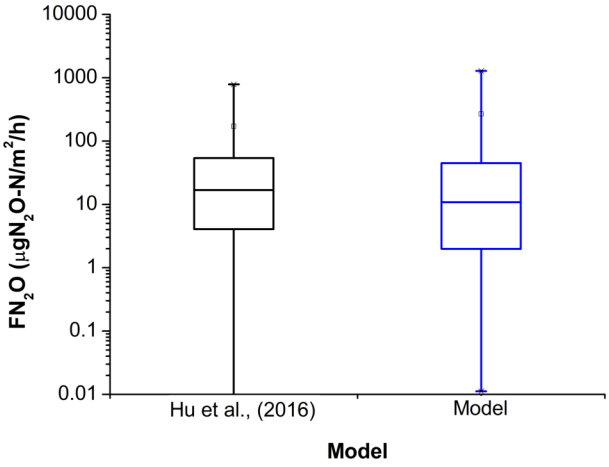


Figure 7: Boxplot of FN_2O ($\mu gN_2O - N\, m^{-2}h^{-1}$) reported by Hu et al. [32], black box, and obtained by model (1) under suitable hydro-morphodynamics and biogeochemical conditions [18, 5], blue box. Streams in desert areas are removed.

Table 3: Statistics of the N_2O flux per unit area FN_2O, ($\mu gN_2O - N/m^2/h$)) reported by Hu et al. [32] (Figure S3) and those obtained with our model by neglecting emissions from streams in desert areas (Min is the minimum value, $1^{st}Qu.$ is the first quantile, $3^{rd}Qu.$ is the third quantile, Max is the maximum value and IQR is the interquantile range).

Statistics of FN_2O	**Hu et al. (2016)**	**Proposed Model**
Min	-3.955	0.0112
1^{st} Qu.	4.084	1.980
Median	16.8	10.814
3^{rd} Qu.	53.849	44.863
Max	786.744	1282.076
IQR	49.765	42.883
Mean	30.849	40.320

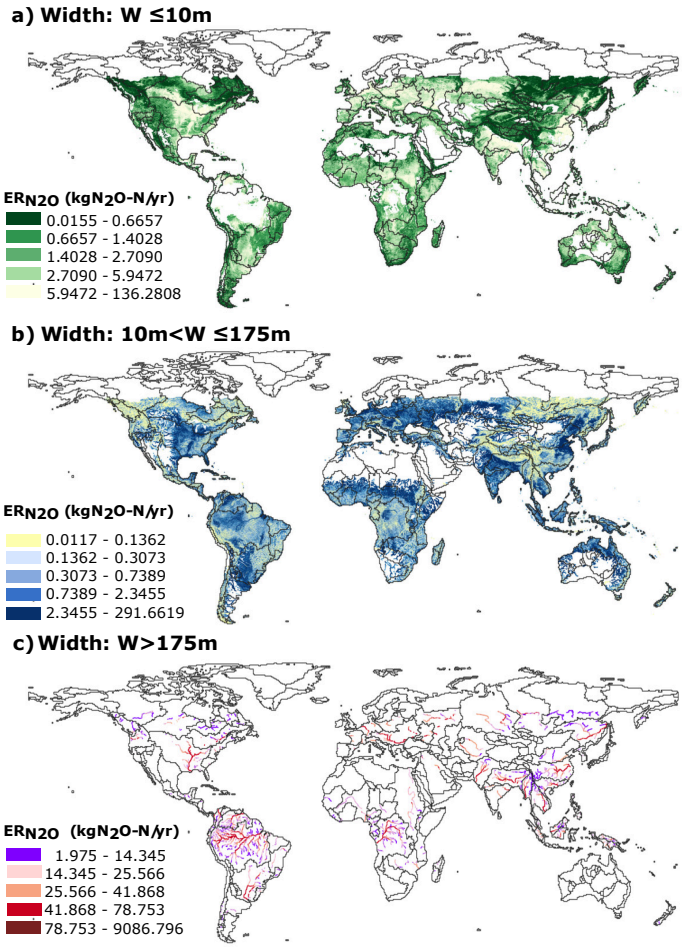


Figure 8: Maps of N_2O emission rate as a function of stream/river size. Patterns of N_2O emission rate ($ER_{N2O}, kgN_2O - N/yr$) as a function of the stream width: (a) small streams ($W \leq 10m$), (b) intermediate size rivers ($10m < W \leq 175m$) and (c) large rivers ($W > 175m$).

emission rate of N_2O from the different riverine environments (hyporheic zone 8a, benthic zone 8b and water column 8c) that characterize river networks.

Table 4: Main statistics (minimum, mean and maximum) of model output parameters. (F^*N_2O is the dimensionless flux of N_2O, FN_2O is the flux of N_2O per unit area and ER_{N2O} is the emission rate of N_2O.)

Parameter	**Minimum**	**Mean**	**Maximum**	**Reference**
$F^*N_2O\,(-)$	0.72×10^{-10}	6.37×10^{-8}	9.12×10^{-7}	Marzadri et al. [43, 44]
$FN_2O\,(\mu molN/m^2/h)$	0.00	1.44	45.79	Calculated
$ER_{N2O}\,(kgN_2O - N/yr)$	1.17×10^{-2}	4.65	9086.79	Calculated

At the reach scale, high ER_{N2O} are observed for small streams (e.g. zones close to the desert areas of Africa and Australia) where the model, in agreement with observations, predicts high FN_2O values, but also in correspondence of large rivers (e.g. Amazon) where, despite low FN_2O values, the large surface areas increase the emission rates. The effect of agricultural practice also impacts ER_{N2O} distribution as exemplified by high emission rates from the US Corn Belt (North America), one of the most intensive agricultural areas in the world with large $FDIN$ leading to large emission rates (Figure 2).

3.5. From local to near-global scale

The reach scale is the unit building block of our hybrid model from which ER_{N2O} can be aggregated to evaluate emissions at larger spatial scales (from the reach to the entire globe). An important intermediate scale is that of the COSCAT regions (Coastal Segmentation and related CATchments [46]). Table 5 reports the N_2O emissions (ER_{N2O}) from these region and Figure

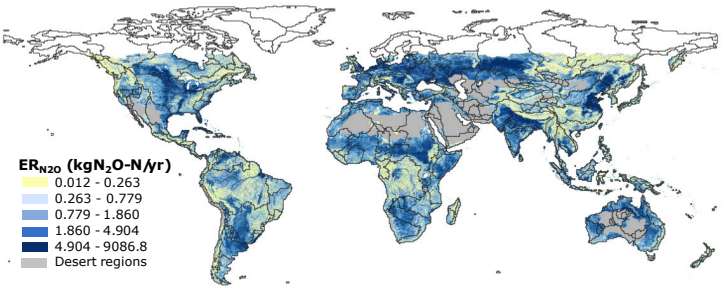


Figure 9: Map of the N_2O emission rate (ER_{N2O}, $kgN_2O - N/yr$). Gray areas represent desert regions not accounted in the calculation.

10 their relative contribution to the total emission. The major contribution of the Amazon and the Mississippi river basins are noticeable, such as the high emissions from the COSCAT regions of the the other major rivers of the world: the Paraná and the Orinoco rivers in South America; the Congo, the Niger and the Nile rivers in Africa; the Volga and the Danube rivers in Europe; and the Ganges, the Yang-Tze, the Huai and the Ob rivers in Asia (see Table 5 for major details).

On the other hand, COSCAT regions close to the deserted areas of North Africa and Asia show a small relative contribution, despite characterized by high ER_{N2O} at the reach scale (Figure 9). This is due to differences in the river network structure, which is characterized by a lower density and higher fragmentation in the vicinity of major deserts.

The fine scale resolution of our model shows that small streams with widths less than 10 m, where the bio-geochemical transformations occurring in the hyporheic zone control N_2O emissions, account for nearly 50% of near-global N_2O emissions ($TE_{N2O,\%}$ in Table 6 and TE_{N2O} in Figure 11). The contribution of small streams is significant despite representing only $\sim$ 13%

Table 5: Total nitrous oxide emission $ER_{N2O,COSCAT}\,(GgN_2O-N/yr)$ along the major world river basin and percentage of river contribution $(ER_{N2O,COSCAT}/TE_{N2O}\,(\%))$ on N_2O emission with respect to the total global emission $TE_{N2O}\,(GgN_2O-N/yr)$.

River	$\mathbf{ER_{N2O,COSCAT}}$	$\mathbf{ER_{N2O,COSCAT}/TE}$
–	$\mathbf{GgN_2O-N/yr}$	%
Amazon	7.279	10.01%
Mississippi	4.576	6.29%
Ob	3.847	5.29%
Ganges	2.693	3.70%
Amur	2.283	3.14%
Volga	2.218	3.05%
Danube	2.134	2.93%
Congo	1.895	2.60%
Nile	1.779	2.45%
Parana	1.752	2.41%
Yang-Tze	1.661	2.28%
Nelson	1.543	2.12%
Dnepr	1.164	1.60%
Orinoco	1.111	1.53%
Indus	1.069	1.47%
Niger	0.951	1.31%
St. Lawrence	0.881	1.21%
Don	0.812	1.12%
Lake Chad	0.783	1.08%
Huai	0.751	1.03%

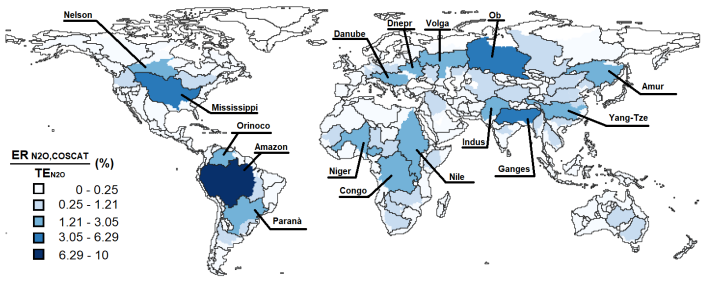


Figure 10: Percentage of N_2O emission ($ER_{N2O,COSCAT}/TE_{N2O}$ (%)) from the major streams and river network worldwide. The major watersheds are characterized according to the COSCAT regions [46] and Table 5 report the emission details.

of the entire riverine surface area (at the mean annual discharge within the 60 ^{o}North and 5 ^{o}South parallels). Their disproportionate contribution is because the hyporheic zone is a powerful and effective biogeochemical reactor, which accommodates both aerobic and anaerobic processes. Although, on average, small streams have higher concentrations of DIN (about 70% more) than large rivers, their mean DIN mass fluxes are about 90% lower than those of the large rivers (see Table 6) due to their lower discharge. Interestingly, the importance of streambed processes is epitomized in all continents except for South America, whose emissions are dominated by the Amazon River (see Figure 11).

In the Amazon River basin, intermediate and large rivers account for most ($\sim$94%) of the N_2O emissions (Figure 8). This is the consequence of the low DIN mass fluxes in small streams, which are mainly within the Amazon forest, and thus pristine, and the large amount of surface area occupied by large rivers (Figure 8). As a river gets larger, the concentration of DIN increases as the river crosses agricultural lands. The Amazon River basin

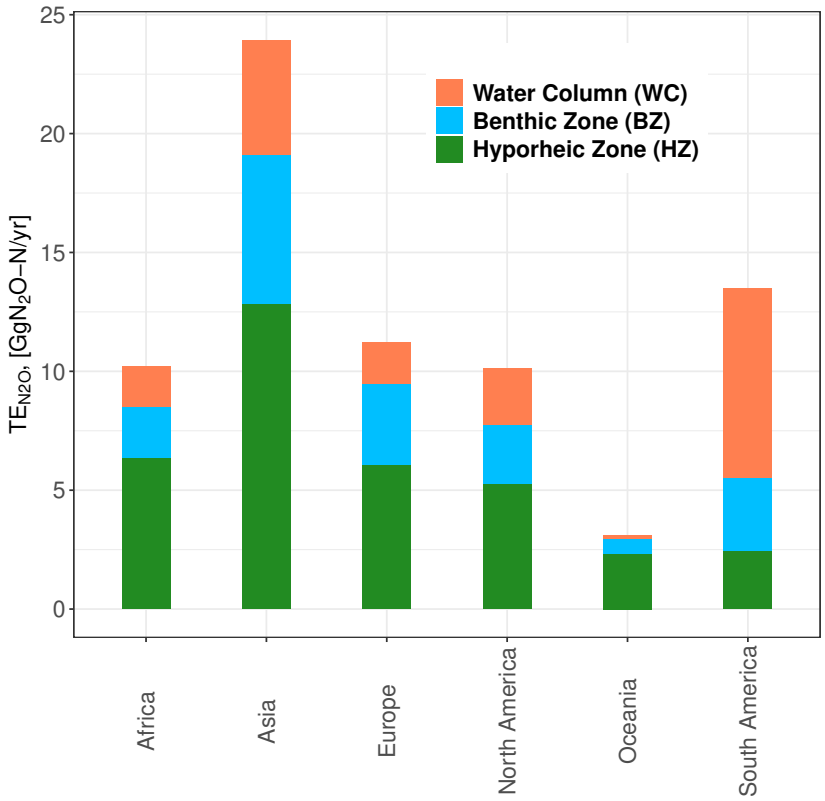


Figure 11: Total emission rate of N_2O from continents ($TE_{N2O}, GgN_2O - N/yr$). Column bars represent the overall emission rate as TE_{N2O} from the hyporheic zone (green), the benthic zone (light blue) and the water column (orange). The size of the three zones changes among the continents, according to the prevalent interaction of stream features and DIN concentration.

as a whole has comparable mean DIN mass fluxes as the highly agricultural Mississippi river basin, which has a fraction of the Amazon discharge, and consequently DIN concentrations are larger in the Mississippi (as reported in Shen et al. [70]) (Figure 1). Conversely, the Tibetan plateau has low DIN fluxes and mainly small streams and shows low N_2O emissions. The large

Table 6: Characteristic values of total surface area ($SA\,[km^2]$), total emission rate of N_2O ($TE_{N2O}\,[GgN_2O-N/yr]$), average stream velocity ($\overline{V}\,[m/s]$), average Dissolved Inorganic Nitrogen ($\overline{DIN}\,[\mu molN/L]$) and average DIN mass flux ($\overline{FDIN}\,[molN/m^2/h]$) within different range of channel width ($W\,[m]$). Values of total surface area and total emission rate of N_2O are represented also in terms of percentage respect to their total global values ($SA_{\%}$ and $TE_{N2O,\%}$).

Width Range	$\sum SA$	$\sum SA_{\%}$	TE_{N2O}	$TE_{N2O,\%}$	$\overline{V}$	$\overline{DIN}$	$\overline{FDIN}$
$[m]$	$[\times 10^3 km^2]$	$[\%]$	$[\frac{GgN_2O-N}{yr}]$	$[\%]$	$[m/s]$	$[\frac{\mu molN}{L}]$	$[\frac{molN}{m^2h}]$
$W \leq 10$	56.419	12.80	35.65	49.0	0.10	71.42	26.80
$10 < W \leq 175$	287.778	65.31	18.25	25.10	0.29	51.97	53.87
$W > 175$	96.460	21.89	18.85	25.9	1.62	46.26	263.76

rivers fed by the Tibetan plateau (e.g. Ganges and Yang-Tze) have large agricultural activities and consequently large N_2O emissions (Figure 8).

By spatially aggregating the total emissions quantified using mean annual quantities (discharge and DIN values), we calculated a near-global riverine N_2O emission of 72.8 GgN_2O-N/yr. This value is within the range of estimation generated from previous studies based on various methods to calculate global N_2O emission. Specifically, our estimate is larger than that predicted by Hu et al. [32] (between 12.4 and 66.9 GgN_2O-N/yr, see also Figures 7 and Table 3) but lower than Maavara et al. [42] (between 91.2 and 98.8 GgN_2O-N/yr) and Yao et al. [77] ($291.3 \pm 58.6 GgN_2O-N/yr$), and it is an order of magnitude smaller than those from earlier works (350-2,100 GgN_2O-N/yr) [36] based on emission rates, which were treated as spatially uniform. Our hybrid model captures local processes and values that previous works could not resolve as their resolution is coarse [32, 42] and did not

account for the effect of hydromorphology on N_2O emissions (e.g., Hu et al. [32], Maavara et al. [42] and Yao et al. [77]). Here, we explicitly characterized the role of the different riverine environments, e.g., hyporheic, benthic and water column zones in controlling N_2O emissions. Previous works did not separate the hyporheic and aquifer contributions (e.g., Yao et al. [77]), which are lumped together as a sole source. Using stream hydromophological information our model identifies the advective time scale of the hyporheic zone but does not quantify the N_2O aquifer contribution. This latter contribution is difficult to quantify, characterized by high uncertainty because of the inherent complexity of mapping groundwater fluxes, their upwelling locations in the stream and their water age. It is potentially small (e.g., Jurado et al. [35]) and much smaller than that of the hyporheic zone, because of long residence times which allow for full denitrification to N_2.

Our predictions constrain N_2O emissions from riverine systems between 0.3% and 0.6% of the global N_2O emissions (12.2 - 23.5 Tg $N_2O - N/yr$ [72]), excluding other inland water bodies like reservoirs, lakes and wetland, which may account of to 0.165 Tg $N_2O - N/yr$ [42]. This better qualifies the N_2O emission compared to previous prediction and allows to better prioritize N_2O management. The reach-scale resolution of our model allows identifying critical watershed or stream reaches.

3.6. Limitations and strengths

In Section 2.2, we reported that the error of the RF model of the NH_4^+ concentration is higher than that for NO_3^-. However, given that the in-stream NH_4^+ concentration is only 13% of NO_3^- concentration and its contribution on the Total Nitrogen (TN) entering the stream network is $\simeq$ 7%, the ex-

pected error due to the inaccurate estimates of NH_4^+ concentrations is deemed small. For this reason the estimates of NH_4^+ were included into the calculation of the DIN. Neglecting the role of NH_4^+ concentration does not impair the validity of the physically based model [43, 44] but would rather affect the estimation of N_2O emissions by reducing its global emissions of 14% ($\widetilde{TE_{N2O}} = 62.56\, GgN_2O - N/yr$) at the most. Note that the emissions obtained by excluding the load of NH_4^+ from the computation are still in the range of estimations reported in literature [32].

This global prediction may be potentially underestimated because the model does not directly account for N_2O contribution from the large scale groundwater flows, e.g., subsurface flow paths from far away agricultural fields, but it does account for the shallow groundwater processes within the hyporheic zone. Frequently, subsurface components such as hyporheic zone and groundwater are represented as a unique source of N_2O to the stream in field experiments, and is often referred to as N_2O from indirect denitrification [9, 77]. However, these two environments are hydrologically and biogeochemically disparate because of their different temporal and spatial scales of interaction with the riverine water [73]. We argue that the model captures the major contribution of the subsurface environment to streams and rivers by accounting for hyporheic processes. Furthermore, our model captures the land use effect via DIN concentrations and hydromorphological characteristics of the riverine reaches via the ML predictions such that contribution like sewage and wastewater treatment plants (e.g., urban impact) and agricultural and ranching practices are captured implicitly by the model.

Previous studies showed that HydroSHEDS misses at least one stream

order for some areas, and about ∼18-20% of the effective stream surface area [10, 60]. Unfortunately, with the current available dataset it is not possible to add this stream order into the analysis due to the difficulty of attributing the parameters essential to the model for N_2O emission (e.g., stream-variables [18] to model DIN concentration). Thus it is likely that our emission rate might be ∼18-20% higher if we accounted for this missing smallest stream order. In fact, due to the higher rate of emissions for small streams found here, and by empirical studies [75], it is likely that including this smallest streams order would result in a $> 20\%$ increase in the global emission rate. Moreover, these missed streams are those in which the additional source of N_2O from groundwater could be potentially greater [75]. The analysis also excludes a large portion of the Arctic and Antarctic areas, amounting to 22% of global river and stream surface area at mean flow [2]. However, due to cold temperatures, DIN metabolism is expected to be low in this area. Furthermore, this study was done at the annual time scale and the loading of nutrients, and their biogeochemical uptake is impacted by flow rates and temperature [61]. Future studies should therefore assess the effect of the temporal variability of these drivers on N_2O fluxes.

Knowing the locations of N_2O emissions is helpful to improve overall predictions of the importance of streams and rivers in the global N_2O budget. The ability to quantify N_2O emissions from spatially explicit quantities, which may also vary in time (e.g. DIN concentrations, flow discharges and stream water temperatures) provide predictive power to our model for evaluating alternative scenarios especially in studies concerning climate change. This can help to assess the status of atmospheric N_2O under different cli-

mate projections and strategies to reduce N_2O gas emissions. This novel framework could be extended to predict emissions of other GHGs (or fate of solutes) along riverine systems, hence supporting decision making that explicitly operates on the management of the variables responsible for those processes.

A key element of our approach is the merging of distributed quantities via RF-ML integrated with a set of hydromorphological relationships to inform a physically-based up-scaling to predict emissions of N_2O in our case. The approach can be used for many other processes, which may not be only riverine. The advantage of this approach is that changes in any supporting quantity due to human activities, e.g., DIN, discharge or stream size, and natural processes, e.g., climate, can be accounted within the model, making it a valuable alternative to fully statistical modeling or process based simulations.

4. Conclusions

We propose for the first time a conceptual framework able to take into account, together with dissolved inorganic nitrogen (DIN) concentrations and biochemical quantities, also the role of hydro-morphology in controlling N_2O emissions from streams and rivers worldwide. Our model integrates a data-driven Random Forest Machine Learning (RF-ML) model with a physically-based upscaling model to identify near-global (neglecting Arctic and Antarctic areas) N_2O emission hot spots at different spatial scales, from reach to global. The model framework can be used to predict other biogeochemical processes such as nitrate fates in riverine system.

The physically-based approach allowed to distributed estimates of important hydromophological and biogeochemical quantities globally at reach scale resolution for the first time. It allowed to characterize the near-global distribution of denitrification potential expressed with two denitrification Damköhler numbers (Da_{DHZ} and Da_D) that allow upscaling processes that occur at the local reach-scale both within surface and subsurface riverine environments to predict N_2O emissions at large scales (e.g. watershed or the global river network). It allowed us to map the median grain size (d_{50}), and hydraulic conductivity of streambed materials.

Our investigation constrains the near-global annual riverine N_2O emissions around 72.8 $GgN_2O - N/yr$, whereas the fine-spatial scale resolution allowed us to identify that nearly 50% of these emissions are from small streams (width $\leq$ 10m), where biogeochemical transformations are higher than previously reported [1]. Consequently, watersheds with high small stream density are likely to be hot spots of N_2O emissions especially in agricultural and urban areas, which DIN concentrations are expected to be large. In these systems, reducing excessive use of fertilizer may result in more effective reduction of N_2O than in areas with large rivers. We also recognize that small streams account for the majority of length of the fluvial network and thus regulating these water bodies may be challenging. We propose a new definition of the emission factor, EF, as our F^*N2O, whose values depend on reach scale hydro-morphological and DIN variables through our two Damköhler numbers. This new definition can be used in future IPCC predictions as it would help better distribute N_2O emissions sources.

Appendix A. Characterization of Damköhler numbers

The N_2O model (equation (1)) requires two denitrification Damköhler numbers: Da_{DHZ} and Da_D for the benthic-hyporheic zone and the water column, respectively. The former is defined as the ratio between the hyporheic median residence time τ_{50} and the characteristic time of denitrification τ_D, the latter as the ratio between the characteristic time of turbulent mixing t_m and again τ_D [43]. The hyporheic τ_{50} depends on stream hydro-morphology and is evaluated with the formulation proposed by Elliott and Brooks [20] for reaches with dune-like morphology and the model of Marzadri *et al.*, [45] for reaches with pool-riffle morphology:

$$\begin{cases} \tau_{50} = \frac{2\,cos^{-1}(0.5)}{\lambda^2\,h_m\,K_h}, \text{ dune} \\ \tau_{50} = \frac{0.21\,exp(1.22\,Y_D^*)\,L_{pr}}{C_z\,S\,K_h}, \text{ pool} - \text{riffle} \end{cases} \tag{A.1}$$

where $\lambda = 2\pi/L_d$ is the dune wavelength (with $L_d = 6D$ being the bedform length evaluated according to Yalin [76]), h_m is the amplitude of head variation, which depends on stream hydrodynamic parameters through the equation proposed by Shen et al. [69]. We assume that the dune height is a function of the mean flow depth: $H_d = 0.167\,D$ [76]. Furthermore, $L_{pr} = 6.5\,W$ is the bar length (m), C_z is the dimensionless Chezy coefficient quantifying streambed resistance, and $Y_D^* = D/H_{pr}$ is the dimensionless stream flow depth that depends on the pool-riffle height, H_{pr} (m), according

to the formulation proposed by Ikeda [33]:

$$\begin{cases} h_m = 0.28\,\frac{V^2}{2\,g}\left(\frac{0.167}{0.34}\right)^{3/8} \\ Y_D^* = [0.18\,d_s^{0.45}\beta^{1.45}]^{-1} \text{ valid for } 2 < \beta < 35 \\ C_z = 6 + 2.5\,ln\left(\frac{1}{2.5\,d_s}\right) \end{cases} \tag{A.2}$$

where $d_s = d_{50}/D$ is the grain relative submergence and β is the alternate bar aspect ratio (β, -).

Given the above hypothesis, the final expressions for the two Damköhler numbers are:

$$\begin{cases} Da_{DHZ} = 17.810\,g\frac{2.7\times10^{-7}(10.4[NO_3^-])^{-0.493}D}{K_h\,V^2}, \text{ dune} \\ Da_{DHZ} = \frac{1.365\,W_{bf}\,exp\left[\frac{1.22}{0.18\left(\frac{d_{50}}{D_{bf}}\right)^{0.45}\left(\frac{W_{bf}}{2\,D_{bf}}\right)^{1.45}}\right]}{\left[6{+}2.5\left(\frac{D}{2.5\,d_{50}}\right)\right]K_h\,S}\cdot\frac{2.7\times10^{-7}(10.4[NO_3^-])^{-0.493}}{D}, \text{ pool}-\text{riffle} \\ Da_D = 14.925\frac{2.7\times10^{-7}(17[NO_3^-])^{-0.49}}{\sqrt{g\,D\,S}}, \text{ surface water} \end{cases} \tag{A.3}$$

Appendix B. Upscaling model validation

Here, we used an updated version of the three-equation model proposed by Marzadri et al. [43] to characterize dimensionless emissions of N_2O. The use of the data collected by Turner et al. [74] along the main stem of the Upper Mississippi River allowed Marzadri et al. [44] to extend the validity of the intermediate scaling law to rivers with width up to 175m and to support the applicability of the scaling law model for large rivers with W> 175m.

The three-equation model (1) was previously calibrated using 291 data of N_2O emissions from different worldwide river reaches Marzadri et al. [43, 44].

The first two power law equations (1) ($F^*N_2O_{HZ}$ and $F^*N_2O_{BZ}$) were developed by analyzing N_2O emission data from 28 headwaters streams in the Kalamazoo River (Michigan, USA) watershed [7, 8] and 16 headwaters streams included in the second Lotic Intersite Nitrogen eXperiment (LINXII) study [49, 9]. The third equation ($F^*N_2O_{WC}$) was developed by analyzing 247 literature data collected on six large river networks in Africa [12] and the large tidal Hudson river (United States) [15, 16]. Then, each equation was validated separately with independent data ($n_{t,HZ}$=93, $n_{t,BZ}$=279 and $n_{t,WC}$=766) collected along two Midwestern US rivers (Tippecanoe and Manistee) [43], a mid-sized United Kingdom river (Swale-Ouse) [25, 53], the mid-sized reaches of six African river networks [12] and the main stem of the Upper Mississippi River [75, 41]. Results of this validation are summarized in Figure B1. The left panels, moving from top to bottom, show the fitting of the three power law scaling that predict F^*N_2O from the hyporheic zone (Figure B1a), the benthic zone (Figure B1c) and the water column (Figure B1e). Whereas the right panels show the validation of the three power laws for hyporheic zone: $F^*N_2O_{HZ}$ (Figure B1b), benthic zone: $F^*N_2O_{BZ}$ (Figure B1d), and water column: $F^*N_2O_{WC}$ (Figure B1f). Both fitting (left panels) and validation (right panels) have similar r^2 and NSE. The model fitting panels (a), (c) and (e) reveal the power laws that best fit with the data (see the values of the coefficient of determination: r^2=0.48, r^2=0.75 and r^2=0.54, respectively for HZ, BZ and WC power law); while the model validation panels (b), (d) and (f) confirm their performance to capture F^*N_2O (as underlined by the Nash-Sutcliffe efficiency index [50]: NSE=0.56, NSE=0.41 and NSE=0.57, respectively for HZ, BZ and WC power law) when applied

to streams not included in the calibration set.

Different aspects of the model's capability to reproduce the dominating mechanisms leading to N_2O emissions can be assessed by using the following metrics [48]:

1) the average value of the Absolute Error $\overline{AE}$:

$$\overline{\mathrm{AE}} = \frac{1}{\mathrm{n_{t,ZI}}} \sum_{\mathrm{i=1}}^{\mathrm{n_{t,ZI}}} |\mathrm{F^*N_2O_i^{obs}} - \mathrm{F^*N_2O_i^{sim}}|; \tag{B.1}$$

2) the Nash-Sutcliffe Efficiency index (NSE) [50]:

$$\mathrm{NSE} = 1 - \frac{\sum_{\mathrm{i=1}}^{\mathrm{n_{t,ZI}}} (\mathrm{F^*N_2O_i^{obs}} - \mathrm{F^*N_2O_i^{sim}})^2}{\sum_{\mathrm{i=1}}^{\mathrm{n_{t,ZI}}} (\mathrm{F^*N_2O_i^{obs}} - \overline{\mathrm{F^*N_2O_i^{obs}}})^2}; \tag{B.2}$$

the root mean square error ($RMSE$):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n_{t,ZI}} \sum_{\mathrm{i=1}}^{\mathrm{n_{t,ZI}}} (\mathrm{F^*N_2O_i^{obs}} - \mathrm{F^*N_2O_i^{sim}})^2}; \tag{B.3}$$

and the percentage of bias ($PBIAS$):

$$\mathrm{PBIAS} = \frac{\sum_{\mathrm{i=1}}^{\mathrm{n_{t,ZI}}} (\mathrm{F^*N_2O_i^{obs}} - \mathrm{F^*N_2O_i^{sim}})}{\sum_{\mathrm{i=1}}^{\mathrm{n_{t,ZI}}} (\mathrm{F^*N_2O_i^{obs}})} \cdot 100 \tag{B.4}$$

where $n_{t,ZI}$ is the number of validation data (e.g. $n_{t,ZI} = 93$, 279 and 766 data for ZI=HZ,BZ and WC, respectively), and the superscripts *obs* and *sim* represent measured and simulated values of F^*N_2O, respectively. Furthermore, $\overline{F^*N_2O^{obs}}$ is the mean of the observations. It must be emphasized that the model performance indexes were obtained by applying equation (1) (in the general form $F^*N_2O_{ZI} = a_{ZI}\,(Da)^{b_{ZI}}$, with ZI=HZ,BZ or WC and $Da = Da_{DHZ}$ or $Da = Da_D$ according to the channel width) to the validation data by using the coefficients a_{ZI} and the exponents b_{ZI} obtained by

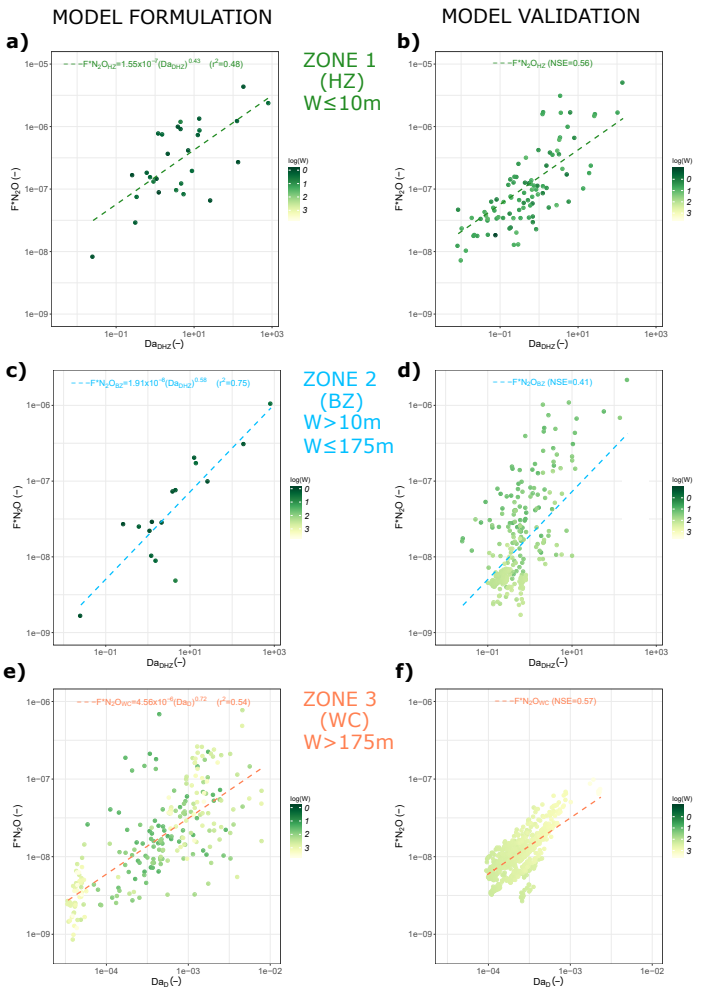


Figure B1: Upscaling models used for capturing the dimensionless flux of N_2O (F^*N_2O) as a function of the denitrification Damköhler numbers (Da_{DHZ} or Da_D) in the three different riverine environments (HZ, BZ and WC) as a function of the stream size (e.g. mean width W). Sub-panels (a), (c) and (e) show model formulation; while sub-panels (b), (d) and (f) show model validation (for HZ, BZ and WC, respectively). Dashes lines in sub-panels (a), (c) and (e) represent the three power law scaling obtained by Marzadri et al. [43, 44] by fitting available literature data (n_{HZ}=28, n_{BZ}=16, n_{WC}=247) on streams and rivers worldwide. Dashed lines in sub-panels (b), (d) and (f) shows the application of these three up-scaling equations to independent sets of available literature data ($n_{t,HZ}$=93, $n_{t,BZ}$=279, $n_{t,WC}$=766). Validation performances may be inferred from the Nash-Sutcliffe Efficiency index (NSE). Symbols are colored from green to light yellow according to the river width (log scale).

calibration on different and independent datasets as described above. This supports the use of the model (1) as a tool to predict global emissions.

The values assumed by the above metrics are shown in Table B1 and confirm the capability of the proposed model to reproduce the processes controlling the transformation of reactive nitrogen in N_2O in naturally reducing environments, as already evidenced in Figure B1.

Table B1: Main statistical parameters: average value of the Absolute Error $\overline{AE}$, Root-Mean Square Error (RMSE), Nash Sutcliffe Efficiency Index (NSE) and Percentage of bias (PBIAS) for the three power law scalings (equation (1)) that model N_2O emissions from different riveine environments (hyporheic zone: HZ or Zone1, benthic zone: BZ or Zone2 and water column: WC or Zone 3) that compose the stream networks.

Model	**Calibration Data**, $\mathbf{n_{ZI}}$	**Validation Data**, $\mathbf{n_{t,ZI}}$	$\overline{\mathbf{AE}}$	**NSE**	**RMSE**	**PBIAS**
[−]	[−]	[−]	[−]	[−]	[−]	[−]
HZ	28	93	0.3218	0.5650	0.4039	0.9022
BZ	16	279	0.3739	0.4051	0.5184	-0.3583
WC	247	766	0.1347	0.5737	0.1762	0.1753

Appendix C. Characterization of median grain size, hydraulic conductivity and stream morphology

Among the different formulations proposed in the literature for estimating the median grain size of the streambed material (d_{50} m), we selected that

proposed by Lee and Julien [37]:

$$d_{50} = \left(\frac{S}{4.981\, Q_{bf}^{-0.346}\, \theta_c^{0.966}} \right)^{1.047} \tag{C.1}$$

where Q_{bf} is bankfull discharge estimated according to the maximum value of the annual flow provided by FLO1K [5] ($Q_{bf} = Q_{max}$), S is the stream slope estimated by the geomorpho90m dataset slope [3], and θ_c is the critical Shields' number, which is set to 0.05 according to the analysis of Andrews [4] for streambeds with armor ratio (e.g. the ratio between the median grain sizes of the surface and subsurface materials) equal to 1.8, which is in the range of experimental values observed in real rivers [4]. We test the predictability of equation (C.1) by using 543 data from 3 different studies [39, 55, 56]. The result of this test is shown in Figure C1. Measured and predicted d_{50} are in a reasonable agreement, with a r^2 of 0.44. Notice that in this case r^2 is a measure of agreement between predicted and measured values without any fitting of the model parameters (equation (C.1)), which inherits the values suggested by Lee and Julien [37].

The map with the near-global distribution of d_{50} is showed in Figure S6 of Supplementary materials.

Several empirical expressions have been proposed to estimate the streambed hydraulic conductivity (K_h, m/d) from granulometric properties, such as d_{50} (mm). In the present work, we compared the values of K_h obtained by applying four models with the range of variability observed in the literature

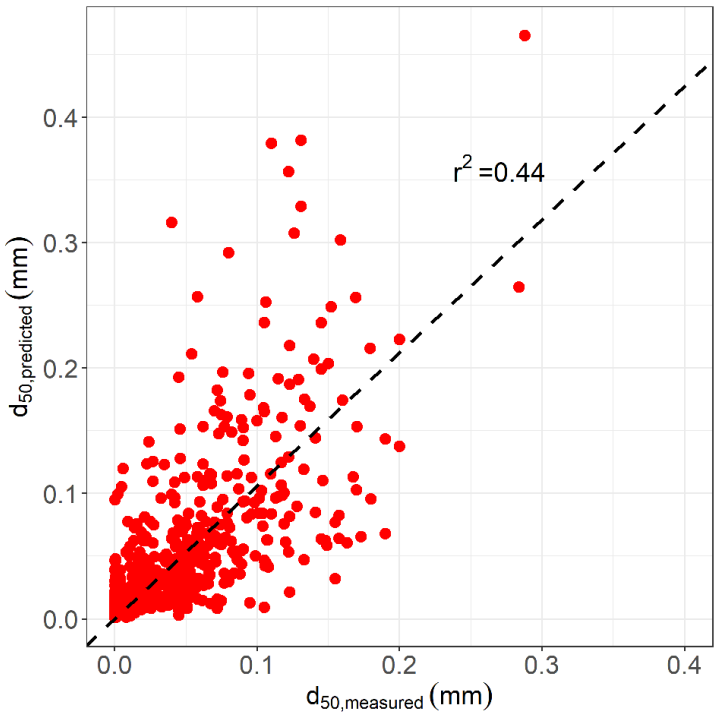


Figure C1: Median grain size (d_{50}) obtained by Eq. (C.1), without calibration, versus the measured ones obtained from previous studies [39, 55, 56]. The dataset is composed by 543 values.

over a wide range of d_{50} values:

$$K_h = \begin{cases} 16.88 + 10.6 \cdot d_{50} \\ 0.04075 \cdot 1014 (d_{50})^{1.5} \\ 0.04075 \cdot 3500 (d_{50})^{1.65} \\ 86400 \cdot 119.06 (d_{50}/1000)^{1.62} \end{cases} \tag{C.2}$$

which are presented in Salarashayeri and Siosemarde [65], Shepherd [71], Gomez-Velez et al. [27].

Figure C2 compares the values of $K_h\,(m/d)$ obtained by applying the four empirical models (C.2) against the envelope of minimum and maximum values proposed in literature for various soil types [13] (yellow band). Here, we adopted the empirical expression proposed by Salarashayeri and Siosemarde [65] ($K_h = 16.88 + 10.6 \cdot d_{50}$), which is the one that better falls within the yellow band [13]. The map of the near-global distribution of K_h along the analyzed streams and rivers is shown in Figure S7 of the Supplementary materials.

Values of d_{50} together with stream slope data, S, are used to classify the streambed morphology of the reaches by using the Montgomery and Buffington [47] classification:

- $S \leq 0.009$ and $d_{50} < 4\,mm$ for dune morphology
- $0.009 < S \leq 0.05$ and $d_{50} > 4\,mm$ for pool-riffle morphology
- $S > 0.05$ and $d_{50} > 4\,mm$ for step-pool and cascade morphology

At present, no hyporheic models are available for streams with step-pool or cascade morphologies. To overcome this limitation, the pool-riffle hy-

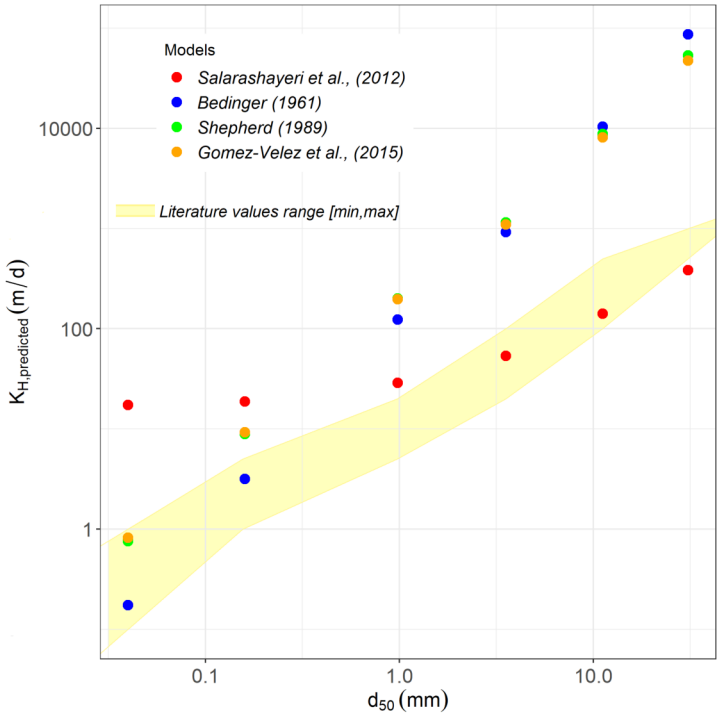


Figure C2: K_h versus d_{50} according to the four empirical expressions of Eq. (C.2). The yellow filled area represents the maximum and minimum range of values reported in literature [13].

porheic model [31, 21, 30] is applied when the aspect ratio, $\beta = W/(2D)$, is within the validity range of the formulation of Ikeda [33]. Reaches belonging to the last category and not respecting this condition are excluded from the computation.

For dune morphology, we assume that the bedform adjusts almost immediately with changes of Q due to the high mobility of the streambed material. Thus, dune size is assumed in dynamic equilibrium with flow, and dune morphology parameters are evaluated as a function of the mean annual discharge, Q [5]. For pool-riffle morphology, because bedform might take long time to adjust, we assume they do not change with discharge and consequently, the parameters that control bedform amplitude (W and D) are evaluated during bankfull conditions (W_{bf} and D_{bf}) according to equation (3) of Raymond et al. [62] by using the maximum annual discharge as a proxy of the bankfull discharge ($Q_{max} = Q_{bf}$ [5]).

CRediT authorship contribution statement

A.M, G.A., D.T., A.B. and P.A.R. designed the study. A.M. and G.A. equally contributed to the manuscript by developing and implementing the computational methodology and the processing chain for the full modelling procedure. L.S. designed the random forest implementation to estimate the N concentration and provided important input on the processing chain. D.T, A.B, G.H.A., P.A.R. provided expert knowledge in the field of hydrology and geochemistry. All authors contributed to the writing of the manuscript and interpretation of the results.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work benefited from discussions as part of the "Inland Water Global HydroBioGeoChemistry" working group. This study was supported in part by the facilities and staff at the Yale Centre for Research Computing (YCRC). It was supported by NASA award NNX17AI74G and by NSF award 1340749, 1344602 and 1457549. A.M. and A.B. acknowledge funding from the Italian Ministry of Education, University and Research (MIUR) in the frame of the Departments of Excellence Initiative 20182022 granted to the Department of Civil, Environmental and Mechanical Engineering of the University of Trento. Any opinions, conclusions, or recommendations expressed in this

material are those of the authors and do not necessarily reflect the views of the supporting agencies.

References

[1] Alexander, R.B., Smith, R.A., Schwarz, G.E., 2000. Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico. Nature 403, 758–761. doi:10.1038/35001562.

[2] Allen, G.H., Pavelsky, T.M., 2018. Global extent of rivers and streams. Science 361, 585–588. doi:10.1126/science.aat0636.

[3] Amatulli, G., McInerney, D., Sethi, T., Strobl, P., Domisch, S., 2020. Geomorpho90m - global high-resolution geomorphometry layers: empirical evaluation and accuracy assessment. Scientific data 7(1):162. doi:10.1038/s41597-020-0479-6.

[4] Andrews, E.D., 1983. Entrainment of gravel from naturally sorted riverbed material. Geol. Soc. Am. Bull. 94, 1225–1231. doi:10.1130/0016-7606(1983)94<1225:EOGFNS>2.0.CO;2.

[5] Barbarossa, V., Huijbregts, M.A., Beusen, A.H., Beck, H.E., King, H., Schipper, A.M., 2018. FLO1K, global maps of mean, maximum and minimum annual streamflow at 1 km resolution from 1960 through 2015. Nature, Scientific Data 5. doi:10.1038/sdata.2018.52.

[6] Baulch, H., Schiff, S., Maranger, R., P.J.Dillon, 2011. Nitrogen enrichment and the emission of nitrous oxide from streams. Global Biogeochem. Cycles 25, GB4013. doi:10.1029/2011GB004047.

[7] Beaulieu, J.J., Arango, C.P., Hamilton, S.K., Tank, J.L., 2008. The production and emission of nitrous oxide from headwater in the Midwestern United States. Global Change Biology 14, 878–894. doi:10.1111/j.1365-2486.2007.01485.x.

[8] Beaulieu, J.J., Arango, C.P., Tank, J.L., 2009. The effects of season and agriculture on nitrous oxide production in headwater streams. Journal of Environmental Quality 38, 637–646. doi:10.2134/jeq2008.0003.

[9] Beaulieu, J.J., Tank, J.L., Hamilton, S.K., Wollheim, W.M., Jr., R.O.H., Mulholland, P.J., Peterson, B.J., Ashkenas, L.R., Cooper, L.W., Dahm, C.N., Dodds, W.K., Grimm, N.B., Johnson, S.L., McDowell, W.H., Poole, G.C., Valett, H.M., Arango, C.P., Bernot, M.J., Burgin, A.J., Crenshaw, C.L., Helton, A.M., Johnson, L.T., O'Brien, J.M., Potter, J.D., Sheibley, R.W., Sobota, D.J., Thomas, S.M., 2011. Nitrous oxide emission from denitrification in stream and river networks. Proc. Natl. Acad. Sci. USA 108, 214–219. doi:10.1073/pnas.1011464108.

[10] Benstead, J.P., Leigh, D.S., 2012. An expanded role for river networks. Nature Geoscience 5, 678–679. doi:10.1038/ngeo1593.

[11] Böhlke, J.K., Antweiler, R.C., Harvey, J.W., Laursen, A.E., Smith, L.K., Smith, R.L., Voytek, M.A., 2009. Multi-scale measurements and modeling of denitrification in streams with varying flow and nitrate concentration in the upper Mississippi River basin, USA. Biogeochem. 93, 117–141. doi:10.1007/s10533-008-9282-8.

[12] Borges, A.V., Darchambeau, F., Teodoru, C.R., Marwick, T.R., Ta-

mooh, F., Geeraert, N., Omengo, F.O., Guérin, F., Lambert, T., Morana, C., Okuku, E., Bouillon, S., 2015. Globally significant greenhouse-gas emissions from African inland waters. Nature Geosci. 8, 637–642. doi:10.1038/NGEO2486.

[13] Bouwer, H., 1978. Groundwater hydrology. McGraw-Hill, New York, United States.

[14] Breiman, L., 2001. Random forests. Machine Learning 45, 5–32. doi:10.1023\%2FA\%3A1010933404324.

[15] Cole, J.J., Caraco, N.F., 2001. Emissions of nitrous oxide (n_2o) from a tidal, freshwater river, the hudson river, new york. Environmental Science Technology 35, 991–996. doi:10.1021/es0015848.

[16] Cole, J.J., Caraco, N.F., McDowell, W.H., Tranvik, L.J., Striegl, R.G., Duarte, C.M., Kortelainen, P., Downing, J.A., Middelburg, J.J., Melack, J., 2007. Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. Ecosystems 10, 171–184. doi:10.1007/s10021-006-9013-8.

[17] Crutzen, J.P., 1970. The influence of nitrogen oxides on atmosphere ozone content. Quarterly Journal of the Royal Meteorological Society 96. doi:10.1002/qj.49709640815.

[18] Domisch, S., G.Amatulli, Jetz, W., 2015. Near-global freshwater-specific environmental variables for biodiversity analyses in 1 km resolution. Nature, Scientific Data 2. doi:10.1038/sdata.2018.52.

[19] Domisch, S., Wilson, A., Jetz, W., 2016. Model-based integration of observed and expert-based information for assessing the geographic and environmental distribution of freshwater species. Ecography 39, 1078–1088. doi:10.1111/ecog.01925.

[20] Elliott, A.H., Brooks, N.H., 1997a. Transfer of nonsorbing solutes to a streambed with bedforms: Theory. Water Resour. Res. 33, 123–136.

[21] Endreny, T.A., Lautz, L., Spiegel, D., 2011. Hyporheic flow path response to hydraulic jumps at river steps: Hydrostatic model simulations. Water Resour. Res. 47, W02518. doi:10.1029/2010WR010014.

[22] Fekete, B.M., Vörösmarty, C.J., Lammers, R.B., 2001. Scaling gridded river networks for macroscale hydrology: development, analysis, and control of error. Water Resour. Res. 37, 1955 – 1967. doi:10.1029/2001WR900024.

[23] Galloway, J.N., Aber, J.D., Erisman, J.W., Seitzinger, S.P., Howarth, R.W., Cowling, E.B., Cosby, B.J., 2004. The Nitrogen Cascades. BioScience 53, 341–356. doi:0.1641/0006-3568(2003)053[0341:TNC]2.0.CO;2.

[24] Galloway, J.N., Cowling, E.B., 2002. Reactive Nitrogen and The World: 200 Years of Change. AMBIO: A Journal of the Human Environment 32, 341–356. doi:0.1641/0006-3568(2003)053[0341:TNC]2.0.CO;2.

[25] García-Ruiz, R., Pattinson, S.N., Whitton, B.A., 1999. Nitrous oxide production in the river Swale-Ouse, north-east England. Water Res. 33(5), 1231–1237. doi:10.1016/S0043-1354(98)00324-8.

790 [26] Garnier, J., Billen, G., Vilain, G., Martinez, A., Silvestre, M., Mounier,
791 E., Toche, F., 2009. Nitrous oxide (n_2o) in the Seine river and basin:
792 observations and budgets. Agr. Ecosyst. Environ. 133, 123–233. doi:10.
793 1016/j.agee.2009.04.024.

794 [27] Gomez-Velez, J.D., Harvey, J.W., Cardenas, M.B., Kiel, B., 2015. Den-
795 itrification in the Mississippi River network controlled by flow through
796 river bedforms. Nat. Geosci. 8, 941–945. doi:10.1038/ngeo2567.

797 [28] Hartmann, J., Lauerwald, R., Moosdorf, N., 2014. A brief overview of
798 the GLObal RIver Chemistry database, GLORICH. Procedia Earth and
799 Planetary Science 10, 23–27. doi:10.1016/j.proeps.2014.08.005.

800 [29] Harvey, J.W., Böhlke, J.K., Voytek, M.A., Scott, D., Tobias, C.R., 2013.
801 Hyporheic zone denitrification: Controls on effective reaction depth and
802 contribution to whole-stream mass balance. Water Resour. Res. 49,
803 6298–6316. doi:10.1002/wrcr.20492.

804 [30] Hassan, M.A., Tonina, D., Beckie, R.D., Kinnear, M., 2014. The effects
805 of discharge and slope on hyporheic flow in step-pool morphologies. Hy-
806 drol. Processes 29, 419–433. doi:10.1002/hyp.10155.

807 [31] Hester, E.T., Doyle, W.A., 2008. In-stream geomorphic structures as
808 drivers of hyporheic exchange. Water Resour. Res. 44, W03417. doi:10.
809 1029/2006WR005810.

810 [32] Hu, M., Chen, D., Dahlgren, R., 2015. Mmodeling nitrous oxide emission
811 from rivers: a global assessmen. Glob. Chang. Biol. 22, 3566–3582.
812 doi:10.1111/gcb.13351.

[33] Ikeda, S., 1984. Prediction of alternate bar wavelenght and height. Journal of Hydraulic Engineering 110(4), 371–386.

[34] IPCC, 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. Technical Report. Geneva, Switzerland.

[35] Jurado, A., Borges, A.V., Brouyère, S., 2017. Dynamics and emissions of n_2o in groundwater: A review. Science of The Total Environment 584-585, 207–218. doi:10.1016/j.scitotenv.2017.01.127.

[36] Kroeze, C., Dumond, E., Seitzinger, S.P., 2010. Future trends in emissions of N_2O from rivers and estuaries. Journal of Integrative Environmental Sciences 7, 295–308. doi:0.1080/1943815X.2010.496789.

[37] Lee, J.S., Julien, P.Y., 2006. Downstream hydraulic geometry of alluvial channels. J. Hydraul. Eng. 132, 1347–1352. doi:10.1061/(ASCE)0733-9429(2006)132:12(1347).

[38] Lehner, B., Verdin, K., Jarvis, A., 2008. New Global Hydrography Derived From Spaceborne Elevation Data. Eos, Transactions American Geophysical Union 89, 93–94. doi:10.1029/2008EO100001.

[39] Li, C., Czapiga, M.J., Eke, E.C., Viparelli, E., Parker, G., 2015. Variable Shields number model for river bankfull geometry: bankfull shear velocity is viscosity-dependent but grain size-independent. J. Hydraulic Res. 53, 36–48. doi:10.1080/00221686.2014.939113.

[40] Lin, P., Pan, M., Beck, H.E., Yang, Y., Yamazaki, D., Frasson, R., David, C.H., Durand, M., Pavelsky, T.M., Allen, G.H., Gleason, C.J., Wood, E.F., 2019. Global reconstruction of naturalized river flows at 2.94 million reaches. Water Resour. Res. 55. doi:10.1029/2019WR025287.

[41] Loken, L.C., Crawford, J.T., Dornblaser, M.M., Houser, J., Turner, P.A., Striegl, R.G., Stanley, E.H., 2018. Minimal nitrate retention in the Upper Mississippi River. not available in review.

[42] Maavara, T., Lauerwald, R., Laruelle, G.G., Akbarzadeh, Z., Bouskill, N.J., Cappellen, P.V., Regnier, P., 2018. Nitrous oxide emissions from inland waters: Are IPCC estimates too high? Global Change Biology 00, 1–16. doi:10.1111/gcb.14504.

[43] Marzadri, A., Dee, M.M., Tonina, D., Bellin, A., Tank, J.L., 2017. Role of surface and subsurface processes in scaling N_2O emissions along riverine networks. PNAS Proceeding of the National Academy of Sciences of the United States of America 114, 4330–4335. doi:10.1073/pnas.1617454114.

[44] Marzadri, A., Tonina, D., Bellin, A., 2020. Power law scaling model predicts N_2O emissions along the Upper Mississippi River basin. Science of the Total Environment 732, 138390. doi:10.1016/j.scitotenv.2020.138390.

[45] Marzadri, A., Tonina, D., Bellin, A., Vignoli, G., Tubino, M., 2010.

858 Semianalytical analysis of hyporheic flow induced by alternate bars. Wa-
859 ter Resour. Res. 46. doi:10.1029/2009WR008285.

860 [46] Meybeck, M., Dürr, H.H., Vörösmarty, C.J., 2006. Global coastal seg-
861 mentation and its river catchment contributors: A new look at land-
862 ocean linkage. Global Biogeochem. Cycles 20, GB1S90. doi:10.1029/
863 2005GB002540.

864 [47] Montgomery, D.R., Buffington, J.M., 1997. Channel-reach morphology
865 in mountain drainage basins. GSA Bulletin 109, 596–611.

866 [48] Moriasi, D.N., Arnold, J., Liew, M.V., Bingner, R., Harmel, R., , Veith,
867 T., 2007. Model evaluation guidelines for systematic quantification of
868 accuracy in watershed simulations. Transactions of the ASABE 50, 885–
869 900. doi:10.13031/2013.23153.

870 [49] Mulholland, P.J., Helton, A.M., Poole, G.C., Jr, R.O.H., Hamilton,
871 S.K., Peterson, B.J., Tank, J.L., Ashkenas, L.R., Cooper, L.W., Dahm,
872 C.N., Dodds, W.K., Findlay, S.E.G., Gregory, S.V., Grimm, N.B., John-
873 son, S.L., McDowell, W.H., Meyer, J.L., Valett, H.M., Webster, J.R.,
874 Arango, C.P., Beaulieu, J.J., Bernot, M.J., Burgin, A.J., Crenshaw,
875 C.L., Johnson, L.T., Niederlehner, B.R., O'Brien, J.M., Potter, J.D.,
876 Sheibley, R.W., Sobota, D.J., Thomas, S.M., 2008. Stream denitrifica-
877 tion across biomes and its response to anthropogenic nitrate loading.
878 Nature 452, 202–206. doi:10.1038/nature06686.

879 [50] Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through concep-

tual models part I - A discussion of principles. J. Hydrol. 10, 282–290. doi:10.1016/0022-1694(70)90255-6.

[51] Nishina, K., Ito, A., Hanasaki, N., Hayashi, S., 2017. Reconstruction of spatially detailed global map of NH_4^+ and NO_3^- application in synthetic nitrogen fertilizer. Earth Syst. Sci. Data 9, 149–162. doi:10.5194/essd-9-149-2017.

[52] Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R., 2001. Terrestrial Ecoregions of the World: a new map of life on earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. BioScience 51, 933–938. doi:10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2.

[53] Pattinson, S.N., García-Ruiz, R., Whitton, B.A., 1998. Spatial and seasonal variation in denitrification in the Swale-Ouse system: a river continuum. Sci Total Environ 210/211, 289–305. doi:10.1016/S0048-9697(98)00019-9.

[54] Peierls, B., Caraco, N., Pace, M., Cole, J., 1991. Human influence on river nitrogen. Nature 350, 386–387. doi:10.1038/350386b0D0.

[55] Pfeiffer, A.M., Finnegan, N.J., 2016. Basin-scale methods for predicting salmonid spawning habitat via grain size and riffle spacing, tested in a

California coastal drainage. Earth Surf. Process. Landform 42, 941–955. doi:10.1002/esp.4053.

[56] Pfeiffer, A.M., Finnegan, N.J., Willenbring, J.K., 2017. Sediment supply controls equilibrium channel geometry in gravel rivers. PNAS Proceeding of the National Academy of Sciences of the United States of America 114, 3346–3351. doi:10.1073/pnas.1612907114.

[57] Quick, A.M., Reeder, W.J., B.Farrell, T., Tonina, D., P.Feris, K., Benner, S.G., 2019. Nitrous oxide from streams and rivers: A review of primary biogeochemical pathways and environmental variables. Earth-Science Rev. 191, 224–262. doi:10.1016/j.earscirev.2019.02.021.

[58] Ravishankara, A., Daniel, J.S., Portman, R.W., 2009. Nitrous oxide (N_2O): The dominant ozone-depleting substance emitted in the 21st century. Science 326, 123–125. doi:10.1126/science.1176985.

[59] Raymond, P.A., Hamilton, S.K., 2018. Anthropogenic influences on riverine fluxes of dissolved inorganic carbon to the oceans. Limnology and Oceanography Letters 3, 143–155. doi:10.1002/lol2.10069.

[60] Raymond, P.A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Dürr, H., Meybeck, M., Ciais, P., Guth, P., 2013. Global carbon dioxide emissions from inland waters. Nature 503, 355–359. doi:10.1038/nature12760.

[61] Raymond, P.A., Saiers, J.E., Sobczak, W.V., 2016. Hydrological and

biogeochemical controls on watershed dissolved organic matter transport: Pulse-shunt concept. Ecology 97, 5–16.

[62] Raymond, P.A., Zappa, C.J., Butman, D., Bott, T.L., Potter, J., Mulholland, P.J., Laursen, A.E., McDowell, W.H., Newbold, D., 2012. Scaling the gas transfer velocity and hydraulic geometry in streams and small rivers. Limnology and Oceanography: Fluids and Environments 2, 41–53. doi:10.1215/21573689-1597669.

[63] Read, E.K., Carr, L., De Cicco, L., Dugan, H.A., Hanson, P.C., Hart, J.A., Kreft, J., Read, J.S., Winslow, L.A., 2017. Water quality data for national-scale aquatic research: The water quality portal. Water Resources Research 53, 1735–1745.

[64] Rutherford, J.C., 1994. River Mixing. John Wiley and Sons, Chichester, England.

[65] Salarashayeri, A.F., Siosemarde, M., 2012. Prediction of soil hydraulic conductivity from particle-size distribution. World Acad. Sci. Eng. Technol. 61, 454–458. doi:10.5281/zenodo.1055779.

[66] Seitzinger, S.P., 1988. Denitrification in freshwater and coastal marine ecosystems: Ecological and geochemical significance. Limnol. Oceanogr. 33, 702–724. doi:10.4319/lo.1988.33.4part2.0702.

[67] Seitzinger, S.P., Harrison, J.A., Böhlke, J.K., Bouwman, A.F., Lowrance, R., Peterson, B., Tobias, C., Drecht, G.V., 2006. Denitrification across landscapes and waterscapes: a synthesis. Ecological Applications 16, 2064–2090.

[68] Seitzinger, S.P., Kroeze, C., 1988. Global distribution of nitrous oxide production and n inputs in freshwater and coastal marine ecosystems. Global Biogeochemical Cycles 12, 93–113. doi:10.1029/97GB03657.

[69] Shen, H.V., Fehlman, H.M., Mendoza, C., 1990. Bed form resistances in open channel flows. J. Hydraul. Eng-ASCE 116, 799–815.

[70] Shen, L., Amatulli, G., Sethi, T., Raymond, P., Domisch, S., 2020. Estimating nitrogen and phosphorus concentrations in streams and rivers, within a machine learning framework. Scientific data 7(1):161. doi:10.1038/s41597-020-0478-7.

[71] Shepherd, R.G., 1989. Correlations of permeability and grain sizes. Groundwater 27, 633–638. doi:10.1111/j.1745-6584.1989.tb00476.x.

[72] Tian, H., Xu, R., Canadell, J.G., Thompson, R.L., Winiwarter, W., Suntharalingam, P., Davidson, E.A., Ciais, P., Jackson, R.B., Janssens-Maenhout, G., Prather, M.J., Regnier, P., Pan, N., Pan, S., Peters, G.P., Shi, H., Tubiello, F.N., Zaehle, S., Zhou, F., Arneth, A., Battaglia, G., Berthet, S., Bopp, L., Bouwman, A.F., Buitenhuis, E.T., Chang, J., Chipperfield, M.P., Dangal, S.R.S., Dlugokencky, E., Elkins, J.W., Eyre, B.D., Fu, B., Hall, B., Ito, A., Joos, F., Krumme, P.B., Landolfi, A., Laruelle, G.G., Lauerwald, R., Li, W., Lienert, S., Maavara, T., MacLeod, M., Millet, D.B., Olin, S., Patra, P.K., Prinn, R.G., Raymond, P.A., Ruiz, D.J., van der Werf, G.R., Vuichard, N., Wang, J., Weiss, R.F., Wells, K.C., Wilson, C., Yang, J., Yao, Y., 2020. A comprehensive

970 quantification of global nitrous oxide sources and sinks. Nature 586,
971 248–256. doi:10.1038/s41586-020-2780-0.

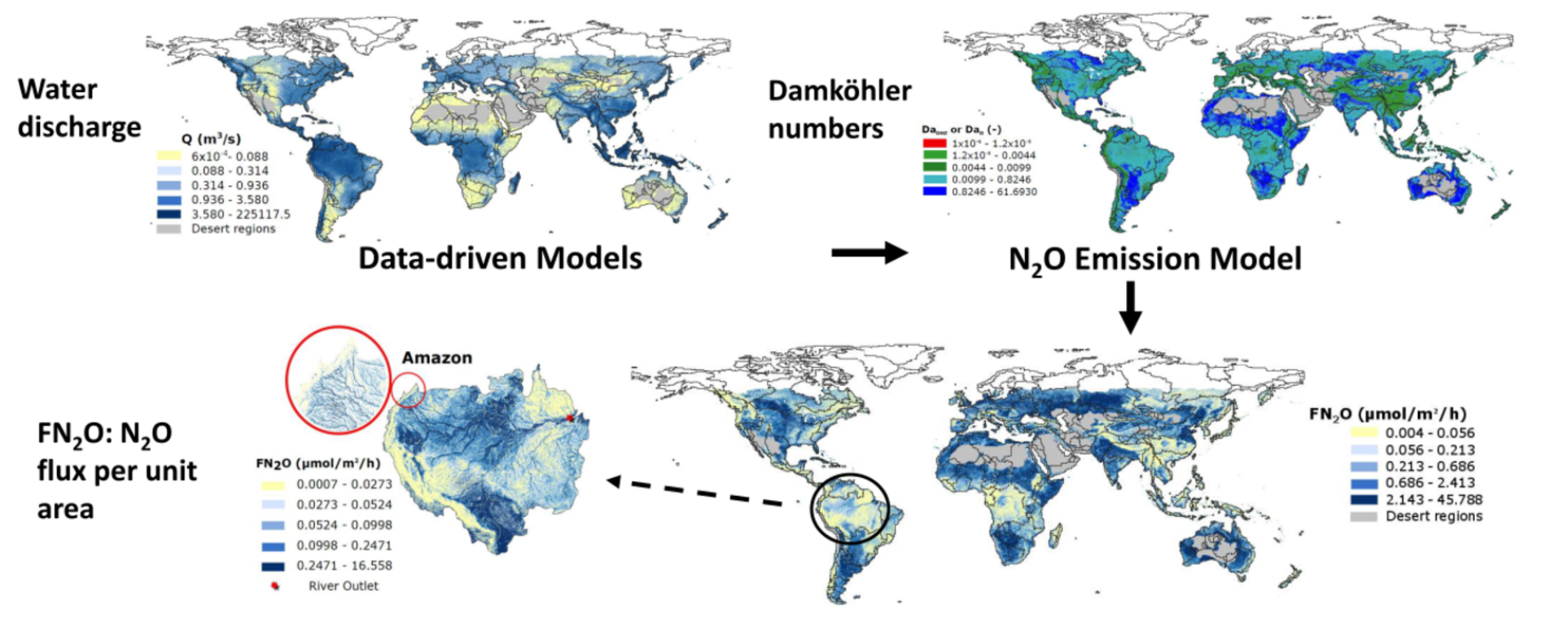
972 [73] Tonina, D., Buffington, J.M., 2009. Hyporheic exchange in mountain
973 rivers I: Mechanics and environmental effects. Geography Compass 3,
974 1063–1086. doi:10.1111/j.1749-8198.2009.00226.x.

975 [74] Turner, P.A., Griffis, T.J., Baker, J.M., Lee, X., Crawford, J.T., Loken,
976 L.C., Venterea, R.T., 2016. Regional-scale controls on dissolved nitrous
977 oxide in the Upper Mississippi River. Geophys. Res. Lett. 43, 4400–4407.
978 doi:10.1073/pnas.1503.

979 [75] Turner, P.A., Griffis, T.J., Lee, X., Baker, J.M., Venterea, R.T., Wood,
980 J.D., 2015. Indirect nitrous oxide emissions from streams within the US
981 Corn Belt scale with stream order. PNAS Proceeding of the National
982 Academy of Sciences of the United States of America 112, 9839–9843.
983 doi:10.1073/pnas.1503598112.

984 [76] Yalin, M.S., 1964. Geometrical properties of sand waves. J. Hydraul.
985 Div-ASCE 90, 105–119.

986 [77] Yao, Y., Tian, H., Shi, H., Pan, S., Xu, R., Pan, N., Canadell, J.,
987 2020. Increased global nitrous oxide emissions from streams and rivers
988 in the Anthropocene. Nat. Clim. Chang. 10, 138–142. doi:10.1038/
989 s41558-019-0665-8.


Water discharge
Q (m³/s)
6x10⁻⁴ - 0.088
0.088 - 0.314
0.314 - 0.936
0.936 - 3.580
3.580 - 225117.5
Desert regions
Damköhler numbers
Da or Da (-)
1x10⁻⁴ - 1.2x10⁻³
1.2x10⁻³ - 0.0044
0.0044 - 0.0099
0.0099 - 0.8246
0.8246 - 61.6930
Data-driven Models
N2O Emission Model
FN2O: N2O flux per unit area
Amazon
FN2O (µmol/m²/h)
0.0007 - 0.0273
0.0273 - 0.0524
0.0524 - 0.0998
0.0998 - 0.2471
0.2471 - 16.558
River Outlet
FN2O (µmol/m²/h)
0.004 - 0.056
0.056 - 0.213
0.213 - 0.686
0.686 - 2.413
2.143 - 45.788
Desert regions