Gridded dataset of nitrogen and phosphorus point sources from wastewater in Germany (1950–2019)

Fanny J. Sarrazin[1,2], Sabine Attinger[1,3], and Rohini Kumar[1]

[1]Department Computational Hydrosystems, UFZ - Helmholtz-Centre for Environmental Research, 04318 Leipzig, Germany
[2]now at Université Paris-Saclay, INRAE, UR HYCAR, 92160 Antony, France
[3]Institute of Environmental Science and Geography, University of Potsdam, 14476 Potsdam-Golm, Germany

Correspondence: fanny.sarrazin@inrae.fr

Abstract. Knowledge about the long history of the anthropogenic inputs of Nitrogen (N) and Phosphorus (P) is crucial to capture long-term N and P processes (legacies) and to investigate water quality and ecosystem health. These inputs include N and P point sources, that originate mainly from wastewater, and that are directly discharged into surface waters, thus having an immediate impact on ecosystem functioning. However, N and P point sources are challenging to estimate, due to the scarcity and uncertainty in observational data. Here, we contribute towards improved characterisation of N and P point sources from wastewater ,by providing a long-term (1950—2019), high resolution (0.015625° ≈ 1.4 km on average) dataset for Germany. The dataset includes both the domestic and industrial emissions treated in wastewater treatment plants and untreated domestic emissions that are collected in the sewer system. We adopt a modelling approach that relies on a large range of data collected from different sources. Importantly, we account for the uncertainties arising from different modelling choices (i.e., coefficients and downscaling approach). We provide 200 gridded N and P point sources realisations, which are constrained and evaluated using available (recent) observations of wastewater treatment plants outgoing loads. We discuss the uncertainties in our reconstructed dataset over a large sample of river basins in Germany, and provide guidance for future uses. Overall, by capturing the long-term spatial and temporal variations in N and P point sources and accounting for uncertainties, our dataset can facilitate long-term and large-scale robust water quality studies. The dataset is available at https://doi.org/10.5281/zenodo.10500535 (Sarrazin et al., 2024).

1 Introduction

Anthropogenic nitrogen (N) and phosphorus (P) enrichment in fresh and marine waters has been observed worldwide, with adverse effects on both human and ecosystem health (Millenium Ecosystem Assessment, 2005). Specifically, these N and P excesses can induce eutrophication of water bodies, resulting first in the excessive development of primary producers, such as phytoplankton and algae that can themselves be toxic to aquatic species and human, and second in a reduction in dissolved oxygen levels (hypoxia) that can impair aquatic ecosystems (Conley et al., 2009; Dodds and Smith, 2016; Lemley and Adams, 2019; Smith, 2003). In the most severe cases, a (almost) complete oxygen depletion can occur (anoxia), leading to the appearance of periodic or persistent 'dead zones' that are unsuitable to sustain most (aquatic) lives (Diaz and Rosenberg, 2008).

Further, inorganic N compounds also cause the acidification of water bodies, are toxic to aquatic species, and make water supply unsafe to drink (Camargo and Alonso, 2006; Lin et al., 2023; WHO, 2016). Notably, not only N and P excesses but also changes in their relative proportions compared to natural conditions have negative consequences on the structure, functioning and diversity of (aquatic) ecosystems (Penuelas et al., 2013, 2020).

N and P contamination of water bodies is a long-standing problem that has developed since the beginning of the twentieth century, and more intensively since the mid-twentieth century, due to an increase in anthropogenic N and P sources (Selman et al., 2008; Le Moal et al., 2019). Importantly, N and P inputs to the environment can affect the water quality status over decades because of the long timescales of the hydrologic transport, biogeochemical, and erosional processes controlling the discharge of nutrients to aquatic systems and the nutrient turnover within aquatic systems. Previous studies have shown that an accumulation (legacy) of N can occur in the soil (Jenkinson, 1991; Sebilo et al., 2013; Van Meter et al., 2016), vadose zone (Ascott et al., 2017; Wang et al., 2012) and groundwater (Puckett et al., 2011; Stuart et al., 2007), and of P in the soil (Parkhurst et al., 2022; Pavinato et al., 2020; Zhang et al., 2022) and sediments in rivers (Sharpley et al., 2013), lakes (O'Connell et al., 2020) and marine environments (Gustafsson et al., 2012; Herbert and Fourqurean, 2008; Kuliński et al., 2022). These N and P legacies can induce a delay between reductions in anthropogenic N and P sources and the corresponding response of water quality (Basu et al., 2022; Ehrhardt et al., 2019; Grimvall et al., 2000; Sharpley et al., 2013; Vero et al., 2018), eutrophication and hypoxia (Gustafsson et al., 2012; Kemp et al., 2009), as well as the response of aquatic ecosystems in terms of species abundance, biodiversity and functions (Jones et al., 2018; McCrackin et al., 2017). In addition, ecosystem recovery from anthropogenic disturbance, such as eutrophication, is a slow process that can take decades or even longer times, depending on the intensity and duration of the disturbance (Duarte et al., 2020; McCrackin et al., 2017; Moreno-Mateos et al., 2017, 2020).

To support nutrient management strategies, a quantification of long-term past sources of both N and P is crucial to drive water quality models over long timescales (decades to centuries), and thus to elucidate the temporal developments of water quality and nutrient legacies (e.g. Lee et al., 2016; Mittelstet and Storm, 2016; Van Meter et al., 2017, 2021). This is also needed to force water quality models over the period with output observations (such as N and P in-stream concentrations), thus allowing for model calibration and evaluation ~~against observations~~ (Sarrazin et al., 2022). N and P sources are separated into two categories, namely point and diffuse sources. Point sources correspond to "a stationary location or fixed facility from which pollutants are discharged; any single identifiable source of pollution" (EEA, 2023c). They encompass domestic, industrial and commercial wastewater that may or may not undergo treatment in wastewater treatment plants (WWTPs), urban runoff that is collected in sewers, as well as intensive livestock operations and fish farms (EEA, 2005; Macias Moy et al., 2022; OECD, 2017). Diffuse sources are "without a single point of origin or not introduced into a receiving stream from a specific outlet" (EEA, 2023c). They include N and P mineral fertilizers and manure application to soils, N and P atmospheric deposition, N biological fixation, P release through weathering, wastewater from households not connected to the sewer system nor to WWTPs, and additional sources in (impervious) urban areas (leaf fall and animal excrement) (Batool et al., 2022; Byrnes et al., 2020; Fuchs et al., 2010; Macias Moy et al., 2022). Both point and diffuse sources are important contributors to N and P levels in receiving waterbodies (e.g. Bouraoui et al., 2011; Sarrazin et al., 2022). The focus of this study is on point sources.

The estimation of long-term N and P point sources from wastewater is challenging, as direct measurements are scarce and uncertain. In Europe, the European Union (EU) makes available observations for the recent period only, namely N and P loads ~~for~~ from urban WWTPs and so-called "direct" industrial release. ~~Within~~ First, within the frame of the European Urban Waste Water Treatment Directive (EC, 1991), the EU publishes N and P loads for urban WWTPs from 2010 onward, including both domestic emissions and so-called "indirect" industrial emissions that are treated in urban WWTPs ~~, from 2010 onward (Waterbase; EEA, 2023b). Within~~ (Waterbase; EEA, 2023b). Second, within the frame of the European Pollutant Release and Transfer Register regulation (E-PRTR; EC, 2006) and the Industrial Emissions Directive (IED; EC, 2010), the EU provides, from 2007 onward, direct industrial N and P release to water (without treatment in urban WWTPs), as well as emissions from the ~~larger urban WWTPs from 2007 onward~~ largest urban WWTPs (EEA, 2020, 2023a). However, these observational datasets are not exhaustive. On the one hand, E-PRTR and IED data records may greatly underestimate industrial emissions to water, since they only cover large industrial activities. They also only include emissions above some prescribed threshold levels, which correspond to a small fraction of facilities subject to reporting (e.g. around 10 % in 2016 in E-PRTR), as discussed in EEA (2019). On the other hand, load data in the Waterbase dataset (EEA, 2023b) ~~,~~ are only reported for some WWTPs included in the record (Vigiak et al., 2020). Moreover, the Waterbase dataset contains records for only larger WWTP treating wastewater from agglomerations with more than 2000 population equivalent (PE; one PE being defined as the organic biodegradable load having a five–day biochemical oxygen demand - BOD5 of 60 g of oxygen per day, EC, 1991)~~and load data are only reported for some WWTPs included in the record (Vigiak et al., 2020)~~. Knowledge about the N and P discharge from smaller WWTPs is yet crucial to understand the water quality in smaller order streams (Yang et al., 2019b, a).
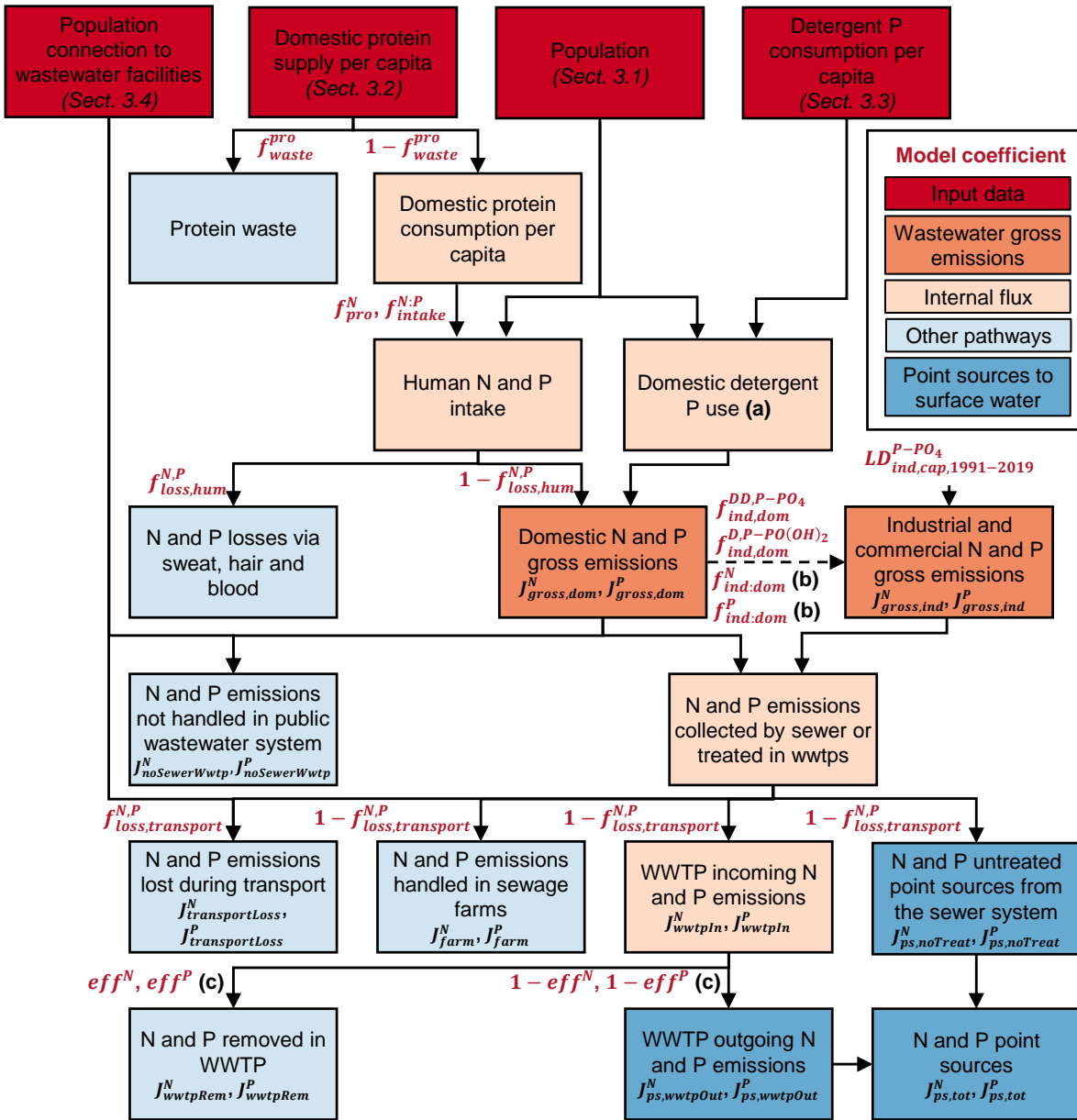
Given this data limitation, modelling strategies are typically used to quantify past N and P point sources. In this respect, on a global scale, Morée et al. (2013), assessed N and P point sources from wastewater for the period 1900–2000. Their approach, which builds on the study of Van Drecht et al. (2009), is based on data at the country level and proxy data when more precise data could not be collected (such as data of gross domestic product ~~(GDP)~~ that are used to estimate P detergent). On a European scale, Vigiak et al. (2020, 2023) provide N and P point sources from wastewater at river basin scale (CCM2 catchments; Jager and Vogt, 2007) for the period 1990–2016. For this, they used the WWTP data of Waterbase for the recent years (2014 and 2016) for EU countries, and a methodology similar to that of Morée et al. (2013) for the period 1990–2010 (further details in Grizzetti et al., 2022; Vigiak et al., 2020). These approaches to past data reconstruction that rely on country-level statistics are useful for large-scale water quality assessment. Yet, they do not account for the sub-national variability that can be large for countries ~~,~~ such as Germany, that was divided between a Western and an Eastern part between 1949 and 1990. Specifically, the performance of the wastewater handling system in East Germany was lagging behind that of West Germany before German reunification in 1990, and in the following years where efforts were put in place to improve wastewater treatment in East Germany (Rudolph and Block, 2001). ~~Yes, in~~ In Germany, N and P point sources data are provided by the German Environmental Agency at country level only for the period 1987–2016 (UBA, 2020). The methodology, which is described in multiple reports for the different years (Behrendt et al., 2000; Fuchs et al., 2010, 2017, 2022), is based on confidential microdata at the WWTP-level, as well as data on industrial direct emissions, in particular from the E-PRTR dataset. Modelling assumptions compensate for the paucity of data for the past period, where in particular WWTP microdata were lacking, as

documented for the year 1985 and 1995 in Behrendt et al. (2000). Importantly, despite the large uncertainties in the point

95   sources data arising from data scarcity and uncertainty, and modelling assumptions, uncertainties were not accounted for in the

data of Vigiak et al. (2020, 2023) and UBA (2020).

Our review of the literature reveals the need for a long-term dataset of past N and P point sources to support water quality assessment over Germany, that accounts for the spatial differences within Germany as well as uncertainties~~in the point sources estimates~~. This is crucial to inform water quality strategies in Germany where the majority of the national monitoring sites

100 for flowing surface water have shown nitrate and phosphorus concentrations above a limit that would ensure a good ecological status , (for instance 81 % for nitrate and 70 % for phosphorus in 2015~~(Arle et al., 2017), and where~~, Arle et al. (2017)). Furthermore, N and P emissions in Germany have contributed to the eutrophication of the North and Baltic Sea since the mid-twentieth century (EEA et al., 2019; Arle et al., 2017). To address this gap, we present here consistent N and P point sources estimates from wastewater over a 0.015625°resolution grid (around 1.4 km) in Germany for the period 1950–2019~~, accounting~~

105 . We account for the uncertainties arising from different modelling choices, namely coefficients and spatial disaggregation approaches to construct grid-level information based on state-level estimates derived for the 16 German federal states. Our state-level (in EU classification "Nomenclature of Territorial units for statistics" level 1 - NUTS-1) estimates are based on state-(NUTS-1) level statistical data, that we compile from different sources, or country- (NUTS-0) level statistics in absence of finer resolution data. We use a modelling approach that builds in particular on Morée et al. (2013), Van Drecht et al. (2009), Vigiak

110 et al. (2020) and IPCC (2019), while we make use of observational data of WWTP N and P emissions to constrain our modeled estimates and check their plausibility. As in previous studies (Morée et al., 2013; Van Drecht et al., 2009; UBA, 2020; Vigiak et al., 2020, 2 , we assess total N and P without distinction between the different forms of N and P. Our dataset encompasses emissions treated in urban WWTPs, including domestic and industial (indirect) emissions, as well as untreated domestic emissions collected in the sewer system. It contributes towards better characterisation of wastewater emissions and their impact on the surrounding

115 (aquatic) environment. In this paper, we discuss the uncertainties of our point source estimates at grid and river basin level to guide future uses of the dataset for water quality studies.

2   N and P point sources model

We estimate N and P gross emissions to wastewater and point sources to surface water (net emissions) building on the studies of Morée et al. (2013), Van Drecht et al. (2009), Vigiak et al. (2020), as well as IPCC (2019) for N emissions only. We consider

120 both domestic and industrial (commercial) indirect emissions that are collected in the sewer system and/or treated in WWTPs. Our approach, which is summarised in Fig.1, uses 18 parameters (or coefficients) defined in Table 1. We first estimate the total N and P point source emissions as well as their partitioning over urban and rural areas at NUTS-1 level (Sect 2.1). These NUTS-1 level data are then subsequently downscaled to grid level (Sect 2.2).

Population connection to wastewater facilities *(Sect. 3.4)*

Domestic protein supply per capita *(Sect. 3.2)*

Population *(Sect. 3.1)*

Detergent P consumption per capita *(Sect. 3.3)*

Model coefficient

- Input data
- Wastewater gross emissions
- Internal flux
- Other pathways
- Point sources to surface water

f_{waste}^{pro} $1 - f_{waste}^{pro}$

Protein waste

Domestic protein consumption per capita

$f_{pro}^{N}, f_{intake}^{N:P}$

Human N and P intake

Domestic detergent P use **(a)**

$LD_{ind,cap,1991-2019}^{P-PO_4}$

$f_{loss,hum}^{N,P}$ $1 - f_{loss,hum}^{N,P}$

N and P losses via sweat, hair and blood

Domestic N and P gross emissions $J_{gross,dom}^{N}, J_{gross,dom}^{P}$

$f_{ind,dom}^{DD,P-PO_4}$
$f_{ind,dom}^{D,P-PO(OH)_2}$
$f_{ind:dom}^{N}$ **(b)**
$f_{ind:dom}^{P}$ **(b)**

Industrial and commercial N and P gross emissions $J_{gross,ind}^{N}, J_{gross,ind}^{P}$

N and P emissions not handled in public wastewater system $J_{noSewerWwtp}^{N}, J_{noSewerWwtp}^{P}$

N and P emissions collected by sewer or treated in wwtps

$f_{loss,transport}^{N,P}$ $1 - f_{loss,transport}^{N,P}$ $1 - f_{loss,transport}^{N,P}$ $1 - f_{loss,transport}^{N,P}$

N and P emissions lost during transport $J_{transportLoss}^{N}, J_{transportLoss}^{P}$

N and P emissions handled in sewage farms $J_{farm}^{N}, J_{farm}^{P}$

WWTP incoming N and P emissions $J_{wwtpIn}^{N}, J_{wwtpIn}^{P}$

N and P untreated point sources from the sewer system $J_{ps,noTreat}^{N}, J_{ps,noTreat}^{P}$

eff^{N}, eff^{P} **(c)** $1 - eff^{N}, 1 - eff^{P}$ **(c)**

N and P removed in WWTP $J_{wwtpRem}^{N}, J_{wwtpRem}^{P}$

WWTP outgoing N and P emissions $J_{ps,wwtpOut}^{N}, J_{ps,wwtpOut}^{P}$

N and P point sources $J_{ps,tot}^{N}, J_{ps,tot}^{P}$

Figure 1. Schematic representation of the N and P point sources model, including the model coefficients (defined in Table 1), the model input data (described in Sect. 3.1–3.4), and the model fluxes (the figure reports the notations for the main fluxes). The fluxes comprise the N and P wastewater domestic and industrial/commercial gross emissions, the model internal N and P fluxes, the N and P point sources to surface water and their two components, that is the treated emissions coming from WWTPs and the untreated emissions coming from the sewer system, and other pathways. We note that (a) the laundry detergent phosphate data we use includes both the domestic and the industrial/commercial components before 1990 (Sect. 3.3.1); (b) the fractions of industrial/commercial to domestic gross N and P emissions, denoted as $f_{ind:dom}^{N}$ and $f_{ind:dom}^{P}$, respectively, are estimated from the their values in 1950 and 2000 that are defined as model coefficients (see Table 1); (c) the efficiency of N and P removal, denoted as eff^{N} and eff^{P}, respectively, result from the combination of the efficiencies of removal for the different treatement types (defined in Table 1).

Table 1. Description of the model parameters and ranges used in this study

Parameter	Description	Unit	Lower value	Upper value
f_{waste}^{pro}	Fraction of protein supply wasted at the distribution and consumption level	(-)	0.14	0.23
f_{pro}^{N}	N content in protein	(kg N kg^{-1})	0.16	0.18
$f_{intake}^{N:P}$	N:P ratio for human intake	(kg N $(kgP)^{-1}$)	9.5	11.5
$f_{loss,hum}^{N,P}$	Fraction of human N and P intake lost via sweat, hair and blood	(-)	0.02	0.04
$f_{loss,transport}^{N,P}$	Fraction of N and P emissions lost during wastewater collection and transport (losses in the sewer system, and in cesspits and during transportation of wastewater collected in cesspits to WWTPs)	(-)	0.05	0.15
eff_1^{N}	Efficiency of N removal for primary treatment	(-)	0.1	0.25
eff_{23noN}^{N}	Efficiency of N removal for secondary and tertiary treatment without targeted denitrification	(-)	0.35	0.6
eff_{3N}^{N}	Efficiency of N removal for tertiary treatment with targeted denitrification	(-)	0.7	0.95
eff_1^{P}	Efficiency of P removal for primary treatment	(-)	0.1	0.3
eff_{23noP}^{P}	Efficiency of P removal for secondary and tertiary treatment without targeted P removal	(-)	0.45	0.65
eff_{3P}^{P}	Efficiency of P removal for tertiary treatment with targeted P removal	(-)	0.8	0.98
$f_{ind:dom,1950}^{N}$	Fraction of industrial/commercial (indirect release) to domestic human (physiological) N gross emissions in 1950	(-)	0.35	0.75
$f_{ind:dom,1950}^{P}$	Fraction of industrial/commercial (indirect release) to domestic human (physiological) gross P emissions in 1950	(-)	0.35	0.75
$f_{ind:dom,2000}^{N}$	Fraction of industrial/commercial (indirect release) to domestic human (physiological) gross N emissions for the period 2000–2019	(-)	0.05	0.35
$f_{ind:dom,2000}^{P}$	Fraction of industrial/commercial (indirect release) to domestic human (physiological) gross P emissions for the period 2000–2019	(-)	0.05	0.35
$f_{ind:dom,1950-2012}^{DD,P-PO_4}$	Fraction of industrial/commercial (indirect release) to domestic dishwasher detergent phosphate use for the period 1950-2012	(-)	0.15	0.35
$f_{ind:dom,1950-2015}^{D,P-PO(OH)_2}$	Fraction of industrial/commercial (indirect release) to domestic detergent phosphonate use for the period 1950-2015	(-)	0.15	0.5
$LD_{ind,cap,1991-2019}^{P-PO_4}$	Industrial/commercial (indirect release) laundry booster phosphate P use per capita for the period 1991–2019	(kg P $capita^{-1}\ yr^{-1}$)	0.005	0.020

Note: Explanations and references for the determination of the ranges are reported in Supplementary Table S2. The parameters are sampled from a uniform distribution.

N and P gross emissions to wastewater can follow other pathways than point sources to surface water, that are depicted in light blue color in Fig. 1. First, part of the N and P gross emissions are not handled in the public wastewater system, that is, they are not collected in the sewer system nor treated in WWTPs. Second, part of the emissions that are collected in the sewer system or treated in WWTPs are lost during collection and transportation (e.g. sewer losses). Third, part of the emissions that are collected in the sewer system can be applied to agricultural soils in sewage farms and are thus a diffuse source to surface water. Fourth, through treatment in WWTPs, part of the N and P emissions are removed from the wastewater according to the efficiency of N and P removal. Further details on these pathways and their quantitative estimation are reported in Supplementary Sect S2.

2.1 N and P emissions at NUTS-1 level

2.1.1 Domestic and industrial N and P gross emissions

Following IPCC (2019) and Morée et al. (2013), we estimate domestic human (physiological) N gross emissions to wastewater per capita, that correspond to the amount of N in human excreta, as a function of the protein supply per capita at the distribution level $Pro_{supply,cap}(t,i)$ ($kg\ capita^{-1}\ yr^{-1}$), the fraction of protein supply wasted at the distribution and consumption level f_{waste}^{pro} ($-$), the N content in protein f_{pro}^{N} ($-$), and the fraction of human N intake lost via sweat, hair, and blood $f_{loss,hum}^{N,P}$ ($-$). We derive the total ,as well as the urban and rural N gross emissions at time t and for the i-th NUTS-1 region denoted as $J_{gross,dom,X}^{N}(t,i)$ ($kg\ yr^{-1}$) based on the population count $Pop_X(t)$, where X, can be urb (urban component), rur (rural component), and all (total, that is, sum of urban and rural components):

$$J_{gross,dom,X}^{N}(t,i) = Pop_X(t,i)Pro_{supply,cap}(t,i)f_{pro}^{N}(1-f_{waste}^{pro})(1-f_{loss,hum}^{N,P}) \tag{1}$$

Domestic P gross emissions $\cancel{J_{gross,dom,X}^{N}(t,i)\ (kg\ yr^{-1})}\underrightarrow{J_{gross,dom,X}^{P}(t,i)\ (kg\ yr^{-1})}$ is estimated assuming a constant N:P ratio for human intake $f_{intake}^{N:P}$ ($kgN(kgP)^{-1}$) (Morée et al., 2013) and considering P emissions from the use of detergents:

$$J_{gross,dom,X}^{P}(t,i) = \underrightarrow{f_{intake}^{N:P}J_{gross,dom,X}^{N}(t,i)}\frac{J_{gross,dom,X}^{N}(t,i)}{f_{intake}^{N:P}} + Pop_X(t,i)\left(LD_{dom,cap}^{P-PO_4}(t,i) + DD_{dom,cap}^{P-PO_4}(t,i) + D_{dom,cap}^{P-PO(OH)_2}(t,i)\right)$$

$$\tag{2}$$

where $LD_{dom,cap}^{P-PO_4}(t,i)$ ($kg\ P\ capita^{-1}\ yr^{-1}$) is the domestic laundry detergent phosphate P use per capita, $DD_{dom,cap}^{P-PO_4}(t,i)$ ($kg\ P\ capita^{-1}\ yr^{-1}$) is the domestic dishwasher detergent phosphate P use per capita, and $D_{dom,cap}^{P-PO(OH)_2}(t,i)$ ($kg\ P\ capita^{-1}\ yr^{-1}$) is the domestic detergent phosphonate P use per capita. We note that $LD_{dom,cap}^{P-PO_4}(t,i)$ includes both domestic and industrial/commercial use until the year 1991 and 1992 for West Germany and East Germany, respectively, as the available data does not allow to separate these two parts (Sect. 3.3.1). While Morée et al. (2013) and Van Drecht et al. (2009) rely on a modelling approach using proxy data such as ~~GDP~~ gross domestic product (GDP) to reconstruct P emissions from detergents, our estimation is based on detergent data that we collect for Germany (Sect. 3.3).

We assess industrial/commercial gross emissions as a fraction of total human emissions following previous studies (IPCC, 2019; Morée et al., 2013; Vigiak et al., 2020). We consider a time-varying fraction of industrial/commercial to domestic gross

N and P emissions ($f^N_{ind:dom}(t,i)$ (-), $f^P_{ind:dom}(t,i)$ (-), respectively) that decreases linearly between 1950 and 2000, similar to Morée et al. (2013), and that remains constant after 2000 as follows:

when $t \leq 2000$:

$$f^N_{ind:dom}(t,i) = f^N_{ind:dom,1950} + \frac{f^N_{ind:dom,2000} - f^N_{ind:dom,1950}}{50}(t-1950) \tag{3}$$

when $t > 2000$:

$$f^N_{ind:dom}(t,i) = f^N_{ind:dom,2000} \tag{4}$$

where $f^N_{ind:dom,1950}$ $(-)$ and $f^N_{ind:dom,2000}$ $(-)$ are the fractions of industrial/commercial to domestic N gross emissions in 1950 and 2000, respectively. Similar equations are used to assess $f^P_{ind:dom}(t,i)$.

To estimate the N and P gross emissions to wastewater, we apply the above coefficients (Eq. 3–4) to the total human N and P emissions (Eq. 1–2), while also accounting for the ~~industrial/commercial~~ detergent P use in businesses and institution such as bars, restaurants, canteens, hotels, bakeries, butcher shops, schools, hospitals, retirement homes (see details on industrial/commercial detergents in Mehlhart et al., 2021). We also consider that industrial/commercial activities are ~~mainly~~ located in urban areas only, hence industrial/commercial rural N and P gross emissions are considered to be equal to ~~zeros~~zero. The industrial/commercial urban N and P gross emissions, denoted as $J^N_{gross,ind,urb}(t,i)$ $(kg\ yr^{-1})$ and $J^P_{gross,ind,urb}(t,i)$ $(kg\ yr^{-1})$, respectively, are calculated as follows:

$$J^N_{gross,ind,urb}(t,i) = f^N_{ind:dom}(t,i) J^N_{\underset{\sim}{gross,dom,tot}\ gross,dom,all}(t,i) \tag{5}$$

$$\begin{aligned} J^P_{gross,ind,urb}(t,i) = &\frac{f^P_{ind:dom}(t,i)}{f^{N:P}_{intake}} J^N_{gross,dom,all}(t,i) + \\ &Pop_{all}(t,i)\left(LD^{P-PO_4}_{ind,cap}(t,i) + DD^{P-PO_4}_{ind,cap}(t,i) + D^{P-PO(OH)_2}_{ind,cap}(t,i)\right) \end{aligned} \tag{6}$$

where $LD^{P-PO_4}_{ind,cap}(t,i)$ $(kg\ P\ capita^{-1}\ yr^{-1})$ is the industrial/commercial laundry detergent phosphate P use per capita, $DD^{P-PO_4}_{ind,cap}(t,i)$ $(kg\ P\ capita^{-1}\ yr^{-1})$ is the industrial/commercial dishwasher detergent phosphate P use per capita, and $D^{P-PO(OH)_2}_{ind,cap}(t,i)$ $(kg\ P\ capita^{-1}\ yr^{-1})$ is the industrial/commercial detergent phosphonate P use per capita. Due to a lack of data on industrial/commercial detergent use, we estimate these quantities from domestic detergent use and specific parameters using the following equations:

$$\begin{aligned} LD^{P-PO_4}_{ind,cap}(t,i) &= 0 && when\ t \leq 1990\ (West\ Germany)\ or\ t \leq 1991\ (East\ Germany) \\ LD^{P-PO_4}_{ind,cap}(t,i) &= LD^{P-PO_4}_{ind,cap,1991-2019} && when\ t > 1990\ (West\ Germany)\ or\ t > 1991\ (East\ Germany) \end{aligned} \tag{7}$$

$$DD^{P-PO_4}_{ind,cap}(t,i) = DD^{P-PO_4}_{dom,cap}(t,i) f^{DD,P-PO_4}_{ind:dom,1950-2012} \qquad when\ t \leq 2012$$

$$DD^{P-PO_4}_{ind,cap}(t,i) = DD^{P-PO_4}_{dom,cap}(2012,i) f^{DD,P-PO_4}_{ind:dom,1950-2012} \qquad when\ t > 2012 \tag{8}$$

$$D^{P-PO(OH)_2}_{ind,cap}(t,i) = D^{P-PO(OH)_2}_{dom,cap}(t,i) f^{D,P-PO(OH)_2}_{ind:dom,1950-2015} \qquad when\ t \leq 2015$$

$$D^{P-PO(OH)_2}_{ind,cap}(t,i) = D^{P-PO(OH)_2}_{dom,cap}(2015,i) f^{D,P-PO(OH)_2}_{ind:dom,1950-2015} \qquad when\ t > 2015 \tag{9}$$

where $LD^{P-PO_4}_{ind,cap,1991-2019}$ ($kg\ P\ capita^{-1}\ yr^{-1}$) is the industrial/commercial laundry phosphate P use per capita for the period 1991–2019, $f^{DD,P-PO_4}_{ind:dom,1950-2012}$ (-) is the fraction of industrial/commercial to domestic dishwasher detergent phosphate use for the period 1950-2012, and $f^{D,P-PO(OH)_2}_{ind:dom,1950-2015}$ (-) is the fraction of industrial/commercial to domestic detergent phosphonate use for the period 1950-2015. Equation 7 accounts for the fact that our data of laundry detergent phosphate include both the domestic and industrial/commercial components until the year ~~1991 and 1992~~ 1990 and 1991 for the NUTS-1 regions of West Germany and East Germany, respectively (Eq. 2). In Eq. 8 and 9, we consider the European Union Detergent Regulation (2012) that led to a reduction in dishwasher detergent phosphate use and a consequent increase in phosphonate detergent use for domestic detergent only ($DD^{P-PO_4}_{dom,cap}(t,i)$ and $D^{P-PO(OH)_2}_{dom,cap}(t,i)$). These changes occurred after 2012 for $DD^{P-PO_4}_{dom,cap}(t,i)$ and after 2015 for $D^{P-PO(OH)_2}_{dom,cap}(t,i)$ in Germany ~~, as documented in~~ (Supplementary Fig S3). Therefore, we no longer assess $DD^{P-PO_4}_{ind,cap}(t,i)$ ~~proportionally to $DD^{P-PO_4}_{dom,cap}(t,i)$ after 2012,~~ and $D^{P-PO(OH)_2}_{ind,cap}(t,i)$ proportionally to ~~$D^{P-PO(OH)_2}_{dom,cap}(t,i)$ after~~ the domestic amounts after 2012 and 2015 ~~,~~ respectively, but we ~~set~~ assume that these two quantities ~~constant to their 2012 value~~ remain constant.

2.1.2 N and P point source emissions from the public wastewater system

Since the equations to calculate the N and P point sources are the same, we denote as $Nutri$ any of the two nutrients N and P. We calculate total, urban, and rural N and P point sources (net) emissions coming from the public wastewater system, denoted as $J^{Nutri}_{ps,X}(t,i)$ ($kg\ yr^{-1}$) defined as:

$$J^{Nutri}_{ps,X}(t,i) = J^{Nutri}_{ps,noTreat,X}(t,i) + J^{Nutri}_{ps,wwtpOut,X}(t,i) \tag{10}$$

N and P point sources (Eq. 10) include both the N and P emissions that are collected in the public sewer system but that are not treated (denoted as $J^{Nutri}_{ps,noTreat,X}(t,i)$ ($kg\ yr^{-1}$), see Eq 11 and 15), and the N and P emissions that are treated in WWTPs via collection in the public sewer system, as well as in cesspits (sealed tanks) from which the wastewater is transported by trucks to WWTPs (denoted as $J^{Nutri}_{ps,wwtpOut,X}(t,i)$ ($kg\ yr^{-1}$), see Eq 13 and 16). We also account for the fraction of emissions lost during wastewater collection and transport through the coefficient $f^{N,P}_{loss,transport}$ $(-)$, including losses in the sewer system (Morée et al., 2013) and during transportation of the wastewater collected in cesspits to WWTPs. In the following, we detail the calculation of the urban and rural point sources. The total point sources is derived as the sum of the urban and rural components.

Urban point sources

We consider that all industrial/commercial emissions are collected in the sewer system ~~in~~ or treated in WWTPs in urban areas, and that they are treated following the same efficiencies as the urban domestic emissions. Morée et al. (2013) made a similar assumption, but considered the WWTP efficiency for the total domestic emissions (including both urban and rural emissions). We estimate the urban untreated N and P emissions as follows:

$$J_{ps,noTreat,urb}^{Nutri}(t,i) = \left(1 - f_{loss,transport}^{N,P}\right)\left(J_{gross,dom,urb}^{Nutri}(t,i)T0_{urb}(t,i) + J_{gross,ind,urb}^{Nutri}(t,i)\frac{T0_{urb}(t,i)}{T_{sewer-wwtp,urb}(t,i)}\right) \tag{11}$$

where $T0_{urb}(t,i)$ $(-)$ is the fraction of urban population connected to the sewer system but not to public wastewater treatment, and $T_{sewer-wwtp,urb}(t,i)$ $(-)$ is the fraction of urban population connected to the sewer system or to WWTPs, which includes the transportation of wastewater from cesspits to WWTPs. $T_{sewer-wwtp,X}(t,i)$ with respect to total $(X = all)$, urban $(X = urb)$ and rural $(X = rur)$ population is estimated as:

$$T_{sewer-wwtp,X}(t,i) = T0_X(t,i) + T_{farm,X}(t,i) + T1_X(t,i) + T23_X^{noNutri}(t,i) + T3_X^{Nutri}(t,i) \tag{12}$$

where $T0_X(t,i)$ $(-)$ is the fraction of population connected to the sewer system but not to public wastewater treatment, $TFarm_X(t,i)$ $(-)$ is the fraction of population connected to sewage farms (see Supplementary Sect. S2), $T1_X(t,i)$ $(-)$ is the fraction of population connected to primary (mechanical) treatment, $T23_X^{noNutri}(t,i)$ $(-)$ is the fraction of population connected to secondary (biological) treatment and tertiary (advanced) treatment without targeted N removal (denitrification, when $Nutri = N$) and P removal (when $Nutri = P$), and $T3_X^{Nutri}(t,i)$ $(-)$ is the fraction of population connected to tertiary treatment with targeted N removal (when $Nutri = N$) and P removal (when $Nutri = P$). Not all WWTPs that are equipped with tertiary treatment specifically remove N and P with very high efficiency (further details in Supplementary Sect. S6 and in particular Fig. S20). In cases without targeted N and P removal, we consider that the efficiency of N and P removal is similar to that of secondary treatment, hence our definition of the treatment classes reported above. To estimate the fraction of both urban and rural population connection to the different types of treatment (Eq. 12) from the data that refer to the total population (see Sect. 3.4), we assume that the wastewater handling system is more advanced in urban areas. Therefore, we consider that the connection of urban population to the more advanced types of treatment has precedence over rural population (see Supplementary Sect. S7 for further details).

Treated N and P emissions are assessed based on the fraction of population connected to the different types of treatment and the efficiency of nutrient removal for each treatment type (IPCC, 2019; Morée et al., 2013; Van Drecht et al., 2009). Here, specifically, we estimate the urban N and P treated emissions as:

$$J^{Nutri}_{ps,wwtpOut,urb}(t,i) = \left(1 - f^{N,P}_{loss,transport}\right)$$

$$\left(J^{Nutri}_{gross,dom,urb}(t,i)\left(T1_{urb}(t,i)(1-eff^{Nutri}_1) + T23^{noNutri}_{urb}(t,i)(1-eff^{Nutri}_{23noNutri}) + T3^{Nutri}_{urb}(t,i)(1-eff^{Nutri}_{3Nutri})\right) + \right.$$

$$\left. J^{Nutri}_{gross,ind,urb}(t,i)\frac{T1_{urb}(t,i)(1-eff^{Nutri}_1) + T23^{noNutri}_{urb}(t,i)(1-eff^{Nutri}_{23noNutri}) + T3^{Nutri}_{urb}(t,i)(1-eff^{Nutri}_{3Nutri})}{T_{sewer-wwtp,urb}(t,i)}\right)$$

(13)

235   where $eff^{Nutri}_1 \; (-)$, $eff^{Nutri}_{23noNutri} \; (-)$, and $eff^{Nutri}_{3Nutri} \; (-)$ are the efficiencies of N removal (when $Nutri = N$) and P removal (when $Nutri = P$) for primary treatment, for secondary and tertiary treatment without targeted N removal (when $Nutri = N$) and P removal (when $Nutri = P$), and for tertiary treatment with targeted N removal (when $Nutri = N$) and P removal (when $Nutri = P$), respectively. In Eq. 11 and 13, the fate of the industrial/commercial gross emissions is assessed by normalizing the population connection to the different treatment types by $T_{sewer-wwtp,urb}(t,i)$. This ensures that all

240   industrial/commercial gross emissions are collected in the sewer system or treated in WWTPs.

The incoming urban N and P emissions to WWTPs, denoted as $J^{Nutri}_{wwtpIn,urb}(t,i) \; (kg \; yr^{-1})$, which does not account for N and P removal in WWTPs, are calculated as:

$$J^{Nutri}_{wwtpIn,urb}(t,i) = \left(1 - f^{N,P}_{loss,transport}\right)\left(J^{Nutri}_{gross,dom,urb}(t,i)(T1_{urb}(t,i) + T23^{noNutri}_{urb}(t,i) + T3^{Nutri}_{urb}(t,i)) + \right.$$

$$\left. J^{Nutri}_{gross,ind,urb}(t,i)\frac{T1_{urb}(t,i) + T23^{noNutri}_{urb}(t,i) + T3^{Nutri}_{urb}(t,i)}{T_{sewer-wwtp,urb}(t,i)}\right)$$

(14)

Rural point sources

245

In our study, the rural N point sources originate from domestic emissions only and are assessed similar to the domestic components of the urban N point sources (Eq. 11 and 13). The rural untreated N and P emissions of Eq. 10 are calculated as follows:

$$J^{Nutri}_{ps,noTreat,rur}(t,i) = \left(1 - f^{N,P}_{loss,transport}\right)\left(J^{Nutri}_{gross,dom,rur}(t,i)T0_{rur}(t,i)\right)$$

(15)

250   The rural treated N and P emissions of Eq. 10 are estimated as:

$$J^{Nutri}_{ps,wwtpOut,rur}(t,i) = \left(1 - f^{N,P}_{loss,transport}\right)(J^{Nutri}_{gross,dom,rur}(t,i)(T1_{rur}(t,i)(1-eff^{Nutri}_1) + $$

$$T23^{noNutri}_{rur}(t,i)(1-eff^N_{23noNutri}) + T3^{Nutri}_{rur}(t,i)(1-eff^{Nutri}_{3Nutri}))$$

(16)

The incoming rural N emissions to WWTPs, denoted as $J_{wwtpIn,rur}^{Nutri}(t)$ $(kg\ yr^{-1})$ ~~, which does not account for N and P removal in WWTPs~~ are calculated as:

$$J_{wwtpIn,rur}^{Nutri}(t,i) = \left(1 - f_{loss,transport}^{N,P}\right) J_{gross,dom,urb}^{Nutri}(t,i)(T1_{rur}(t,i) + T23_{rur}^{noNutri}(t,i) + T3_{rur}^{Nutri}(t,i)) \qquad (17)$$

2.2 N and P point sources emissions at grid level

We adopt two different approaches to downscale the N and P point sources emissions from NUTS-1 level to grid level to account for uncertainty. First, similar to Morée et al. (2013) and Vigiak et al. (2020), we perform the disaggregation according to the population density. We use gridded urban and rural population counts to disaggregate the NUTS-1 level emissions of the i-th NUTS-1 region calculated from Eq. (11, 13, 15, and 16) to the k_i-th grid cell that belong to the i-th NUTS-1 region, as follows:

$$J_{ps,noTreat,grid}^{Nutri}(t,k_i) = J_{ps,noTreat,urb}^{Nutri}(t,i)\frac{Pop_{urb,grid}(t,k_i)}{Pop_{urb}(t,i)} + J_{ps,noTreat,rur}^{Nutri}(t,i)\frac{Pop_{rur,grid}(t,k_i)}{Pop_{rur}(t,i)} \qquad (18)$$

$$J_{ps,wwtpOut,grid}^{Nutri}(t,k_i) = J_{ps,wwtpOut,urb}^{Nutri}(t,i)\frac{Pop_{urb,grid}(t,k_i)}{Pop_{urb}(t,i)} + J_{ps,wwtpOut,rur}^{Nutri}(t,i)\frac{Pop_{rur,grid}(t,k_i)}{Pop_{rur}(t,i)} \qquad (19)$$

where $J_{ps,noTreat,grid}^{Nutri}(t,k_i)$ $(kg\ yr^{-1})$ and $J_{ps,wwtpOut,grid}^{Nutri}(t,k_i)$ are the untreated and treated point source emissions at grid level, respectively; $Pop_{urb,grid}(t,k_i)$ and $Pop_{rur,grid}(t,k_i)$ are the urban and rural population counts at grid level, respectively; and $Pop_{urb}(t,i)$ and $Pop_{rur}(t,i)$ are the urban and rural population counts at NUTS-1 level, respectively.

Second, we use WWTP observations to disaggregate the NUTS-1 level treated emissions of Eq. (13, 16). To derive spatially distributed estimates of domestic waste emissions to European waters, Vigiak et al. (2020) also used WWTP data, specifically entering load expressed in population equivalent (PE). In this study, since we have access to detailed WWTP data for Germany, we construct a gridded map of WWTP outgoing N and P loads around the year 2016 (see Sect.3.5.1) that we utilise to downscale the NUTS-1 level emissions to the k_i-th grid cell as follows:

$$J_{ps,wwtpOut,grid}^{Nutri}(t,k_i) = (J_{ps,wwtpOut,urb}^{Nutri}(t,i) + J_{ps,wwtpOut,rur}^{Nutri}(t,i))\frac{J_{ps,wwtpOut,grid,obs}^{Nutri}(2016,k_i)}{J_{ps,wwtpOut,obs}^{Nutri}(2016,i)} \qquad (20)$$

where $J_{ps,wwtpOut,grid,obs}^{Nutri}(2016,k_i)$ $(kg\ yr^{-1})$ and $J_{ps,wwtpOut,obs}^{Nutri}(2016,i)$ $(kg\ yr^{-1})$ are the observations of WWTP outgoing N and P ~~loaded~~ loads around the year 2016 at grid level and NUTS-1 level, respectively. This downscaling approach (Eq. 20) is more reliable for years that are closer to 2016, since it does not account for temporal changes in WWTPs. This approach assumes that the location of the WWTPs and their relative contribution to the total N and P treated point sources did not change substantially in time. To apply both disaggregation schemes based on population (Eq. 18–19) and ~~wwtp~~ WWTP data (Eq. 20), we derive gridded maps of population data (see Sect. 3.1) and WWTP data (see Sect. 3.5.1) at $0.015625° \times 0.015625°$ resolution, which is on average 1.74 km × 1.09 km over Germany.

Table 2. Description of the raw data used to reconstruct the input data for the estimation of N and P point sources at NUTS-1 level in Germany (DE), namely population count, protein, detergent, and population connection data (dark red boxes in Fig. 1).

Variable[a]	Category	Spatial feature		Temporal feature		Source
		Extent	Level	Extent	Frequency	
Pop_{urb}, Pop_{rur} [b]	Population	DE	grid (5'x5')	1950-2017	annual (1950-2000)/ decadal (2001-2017)	HYDE v3.2.1 (Klein Goldewijk et al., 2017, 2022)[e]
Pop_{all} [b]	Population	West DE	NUTS-1	1950–1957	annual	SO-WDE (1970)[f]
Pop_{all} [b]	Population	West DE	NUTS-1	1958–1989	annual	SO-DE (2022a)[f]
Pop_{all} [b]	Population	East DE	NUTS-1	1950–1989	annual	SO-EDE (1955–1990)[f]
Pop_{all} [b]	Population	DE	NUTS-1	1990–2019	annual	SO-DE (2022a)[f]
Pop_{all} [b]	Population	DE	NUTS-3	1995–2019	annual	SO-DE (2022b)[f]
$Pro_{supply,cap}$	Protein	DE	NUTS-0	1950	one value	FAO (1951)[g]
$Pro_{supply,cap}$	Protein	DE	NUTS-0	1961–2019	annual	FAO (2021, 2022)[g]
Pop_{all}	Detergent	Berlin	West and East	1990–2000	annual	SO-BE (2001)[h]
N_{house}^{dish}	Detergent	West DE	NUTS-0	1962–1988	5 years (from 1973)	SO-DE (Fachserie 15 Heft 1)[f]
N_{house}^{dish}	Detergent	DE	West and East	1993–2018	5 years	SO-DE (Fachserie 15 Heft 1)[f]
N_{house}^{dish}	Detergent	DE	NUTS-1	2008	one value	SO-DE (Fachserie 15 Heft 1)[f]
$LD_{tot}^{PO_4}$	Detergent	West DE	NUTS-0	1975–1988	variable	BMU (1989)
$DD_{dom}^{PO_4}$	Detergent	DE	NUTS-0	1994-2019	variable	IKW (2005, 2011, 2017, 2019)[i]
$D_{dom}^{PO(OH)_2}$	Detergent	West DE	NUTS-0	1989	one value	IKW (2017)[i]
$D_{dom}^{PO(OH)_2}$	Detergent	DE	NUTS-0	1994-2019	variable	IKW (2005, 2011, 2017, 2019)[i]
$sewer_{all}$ [c]	Connection	West DE	NUTS-1	1975–2019	3 or 4 years	
$sewer_{all}$ [c]	Connection	East DE	NUTS-1	1991–2019	3 or 4 years	
$wwtp_{tot,all}$ [c]	Connection	West DE	NUTS-1	1975–2019	3 or 4 years	
$wwtp_{tot,all}$ [c]	Connection	East DE	NUTS-1	1991–2019	3 or 4 years	Mostly statistical reports from the statistical offices of DE and the federal states (see details in Supplementary Sect. S6)
$T1_{all}$ [c,d]	Connection	West DE	NUTS-1	1975–2019	3 or 4 years	
$T1_{all}$ [c,d]	Connection	East DE	NUTS-1	1991–2019	3 or 4 years	
$T23_{all}^{noN}$, $T3_{all}^{N}$ [c,d]	Connection	West DE	NUTS-1	1987–2019	3 or 4 years	
$T23_{all}^{noN}$, $T3_{all}^{N}$ [c,d]	Connection	East DE	NUTS-1	1991–2019	3 or 4 years	
$T23_{all}^{noP}$, $T3_{all}^{P}$ [c,d]	Connection	West DE	NUTS-1	1987–2019	3 or 4 years	
$T23_{all}^{noP}$, $T3_{all}^{P}$ [c,d]	Connection	East DE	NUTS-1	1991–2019	3 or 4 years	

[a] Notations: Pop_{all} is the total population count; Pop_{urb} is the urban population count, Pop_{rur} is the rural population count; $Pro_{supply,cap}$ is the protein supply per capita at the distribution level; N_{house}^{dish} is the number of households having a dishwasher; $LD_{tot}^{PO_4}$ is the total (domestic and industrial/commercial) laundry detergent phosphate use; $DD_{dom}^{PO_4}$ is the domestic dishwasher detergent phosphate sale; $D_{dom}^{PO(OH)_2}$ is the domestic detergent phosphonate sale; $sewer_{all}$ is the fraction of population connected to the sewer system; $wwtp_{tot,all}$ is the fraction of population connected to wastewater treatment via the sewer system and cesspits; $T1_{all}$ is the fraction of population connected to primary wastewater treatment; $T23_{all}^{noN}$ and $T23_{all}^{noP}$ are the fractions of population connected to secondary or tertiary wastewater treatment without targeted N and P removal, respectively; $T3_{all}^{N}$ and $T3_{all}^{P}$ are the fractions of population connected to tertiary wastewater treatment with targeted N and P removal, respectively. [b] Population data are also used to downscale the NUTS-1 level point sources to grid level. [c] The actually data availability depends on the NUTS-1 region (details in Supplementary Sect. S6). [d] The raw data may refer to (1) **connection of the resident population** of each NUTS–region to the different treatment types, (2) **connection of the population treated** in each NUTS-1 region to the different treatment types, and (3) **volume of wastewater treated** in each NUTS-1 region following the different treatment types (details in Supplementary Sect. S6.1.2). [e] CC BY 4.0 license. [f] ©Statistisches Bundesamt (German Statistical Office). [g] ©FAO, CC BY-NC-SA 3.0 IGO license. [h] ©Statistisches Landesamt Berlin (Statistical Office of Berlin). [i] ©IKW (license at https://www.ikw.org/impressum).

3 Data

This section describes the raw data that we used to estimate the N and P point sources and our processing approaches to derive (1) the required input data to estimate the N and P point sources at NUTS-1 level, namely population counts, domestic protein supply per capita, detergent P consumption per capita and population connection to wastewater facilities (see Table 2 and dark red boxes in Fig. 1); (2) the data to disaggregate the NUTS-1 level N and P point sources estimates to grid level, namely the gridded population and WWTP data (as described in Sect. 2.2); and (3) the outgoing load data from WWTPs to infer the parameters (coefficients) of the N and P point sources model and to evaluate the model realisations, as explained in Sect. 4. ~~Details on WWTP data is reported in Table 3.~~

We adopt the NUTS regions map provided by the German Federal Agency for Cartography and Geodesy (BKG, 2020) at a scale of 1:250,000, which ~~refers~~ corresponds to the NUTS 2020 classification. Germany includes 16 NUTS-1 regions (~~states~~federal states; Supplementary Fig S1) that we consider as reference regions to construct the N and P point sources data, as explained in Section 2. ~~These NUTS-1 regions can be visualised in Supplementary Fig S1.~~

3.1 Population

We utilise the gridded History Database of the Global Environment dataset (HYDE; Klein Goldewijk et al., 2017) that provides urban ~~, rural, and total~~ and rural population counts at a spatial resolution of 5' every ten years in the period 1950–2000, and annually in the period 2000–2017. We adopt the gridded fractions of urban and rural population to total population (sum of urban and rural population) directly from the HYDE dataset. We derive the gridded total population counts ~~(sum of urban and rural population)~~ by adjusting the HYDE data to match annual statistical data from the Statistical Office of Germany available at NUTS-1 (state) level over 1950–2019 ~~and at NUTS-3 (county) level over 1995–2019. Details on this adjustment procedure are reported in Supplementary Sect. S4.2. The NUTS-1 level statistical data are adopted from the online data portal of the Statistical Office of Germany (~~(SO-DE, 2022a~~) that provides annual population counts for Germany over 1990–2019 and for West Germany over 1958–1989. In addition, we collect data of annual population counts for the period 1950–1957 from a yearbook from the Statistical Office of West Germany (;~~ SO-WDE, 1970~~) and for 1950–1989 from yearbooks from the Statistical Office of East Germany (;~~ SO-EDE, 1955–1990) .~~ For East Germany, we aggregate the population counts provided for former administrative regions to the current NUTS-1 regions following the NUTS-1 region creation act (LEinfG, 1990) . The~~ and at NUTS-3 ~~level statistical data are obtained from the online portal of the Statistical Office of Germany~~ (county) level over 1995–2019 (SO-DE, 2022b)~~and are processed as further detailed .~~ Details on the population data and the adjustment procedure can be found in Supplementary Sect. ~~S4.1~~S4.

We then downscale the adjusted HYDE data to our target resolution of 0.015625°using nearest neighbour resampling. We interpolate them linearly to fill the values for missing years in the period 1950–2017, and we assume the same values in 2018 and 2019 as in 2017. We only use one of the three scenario provided by the HYDE dataset (baseline), since after adjustment of the HYDE data, no appreciable differences are observed between the three scenarios.

The adjusted total, urban, and rural HYDE population data at NUTS-1 level are represented in Fig 2 (red lines and shaded areas). This figure also reports in blue the original HYDE data that show overall a good consistency with the adjusted data, although few discrepancies are noticeable for some NUTS-1 regions~~, such as~~. For instance, in Saxony-Anhalt, ~~where~~ the population decrease in the 1950s is more marked in the original than in the adjusted data, ~~or Bremen, where~~ and, in Bremen,

315 the relative bias between original and adjusted data is over 20 % ~~. A~~ (see Supplementary Sect. S4.3 for a detailed comparison between the original and adjusted HYDE data~~at NUTS-0, NUTS-1 and NUTS-3 level is reported in Supplementary Sect. S4.3, showing the value of adjusting the data.~~). The discrepancies between the original and adjusted HYDE data may result from inconsistencies between the maps of administrative regions used to construct the HYDE dataset and in this study, as well as uncertainties in the HYDE dataset. The corresponding gridded maps of urban and rural population can be visualised in

320 Supplementary Sect. S4.2.
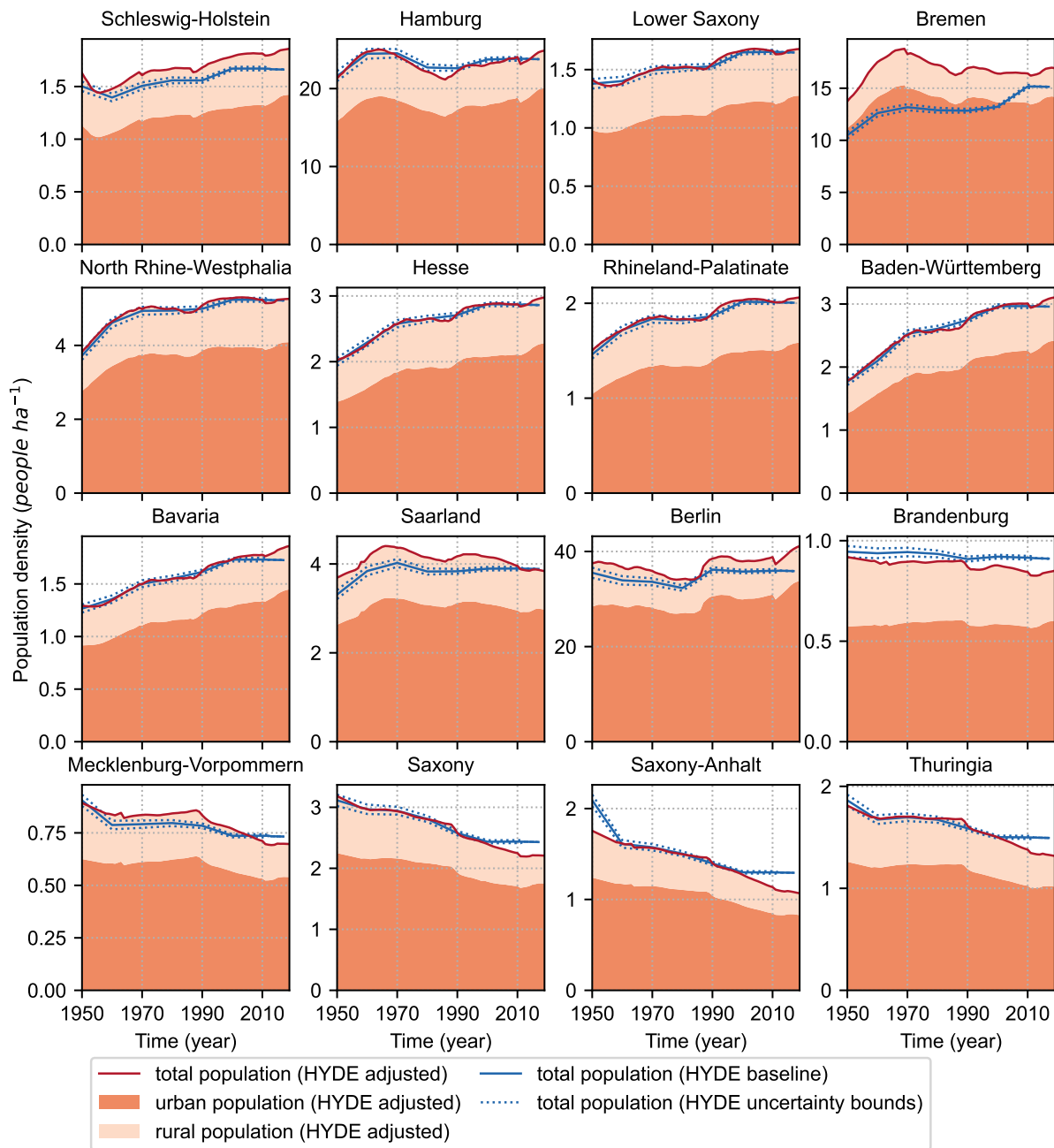
Figure 2. Population density estimated at NUTS-1 level from the original HYDE data for the period 1950–2017 (blue lines) and from the HYDE data that were adjusted to match statistical data for the period 1950–2019 (~~rd~~ red lines). The original HYDE data include a baseline scenario and uncertainty bounds (lower and upper scenarios, ~~dashed~~ dotted lines). ~~Regarding the adjusted HYDE data, the~~ The figure also shows the ~~total,~~ urban and rural population ~~counts~~ density for the adjusted HYDE data that were used to derive the N and P point sources data in this study (shaded areas). Data source: HYDE data (Klein Goldewijk et al., 2017, 2022) are under a CC BY 4.0 license. Statistical population data are from the German Statistical Office (details in Table 2).

3.2 Protein supply

We use data of protein supply at the distribution level in Germany available from the Food and Agriculture Organization (FAO) for the year 1950 (FAO, 1951) and annually in the period 1961–2019 (FAO, 2021, 2022). We set the value in 1950 to the population weighted average of the values provided for West and East Germany. Values for the years in the period 1951-1960 are filled using linear interpolation. The time series of the processed data at NUTS-0 (country) level is shown in Fig. 3a. Since no data are available at NUTS-1 level, we consider that the protein supply per capita is spatially uniform in Germany.
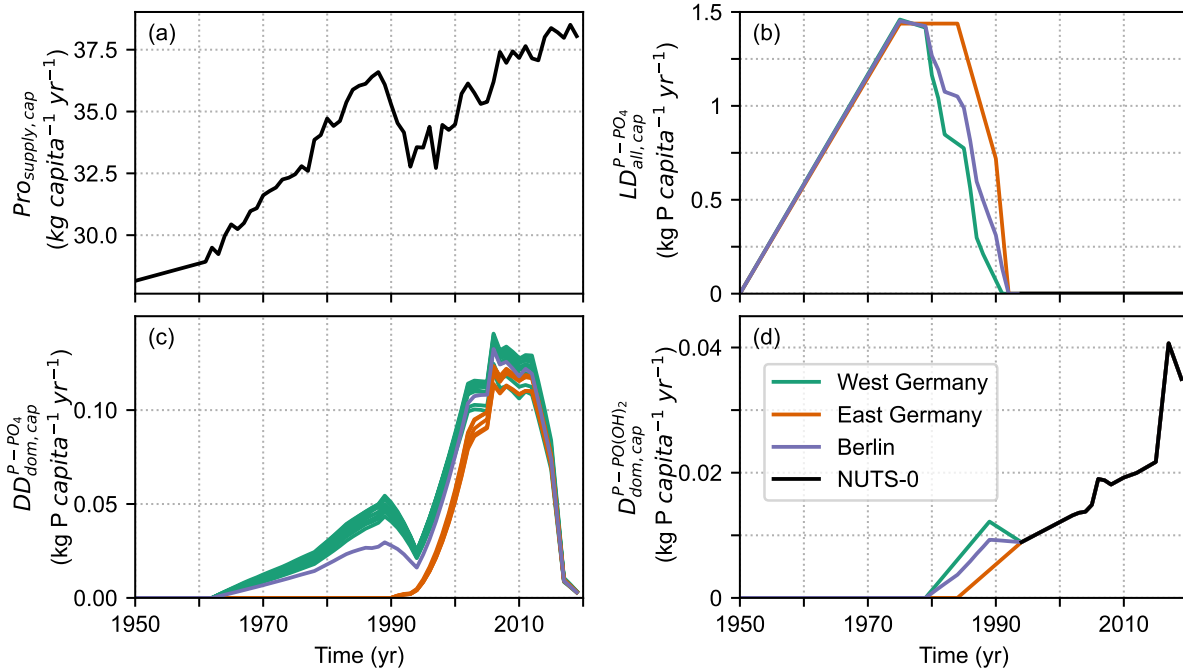
Figure 3. Processed protein and detergent data for the period 1950–2019: (a) protein supply per capita at NUTS-0 level ($Pro_{supply,cap}$), (b) total (domestic and industrial/commercial) laundry detergent phosphate P use per capita for West Germany, East Germany and Berlin ($LD_{all,cap}^{P-PO_4}$), (c) domestic dishwasher detergent phosphate P use per capita for each NUTS-1 region of West and East Germany and Berlin ($DD_{dom,cap}^{P-PO_4}$), and (d) domestic detergent phosphonate P use per capita for West Germany, East Germany and Berlin before 1992 and at NUTS-0 level from 1992 ($D_{dom,cap}^{P-PO(OH)_2}$). For the period after 1991, domestic laundry detergent phosphate P use is equal to zeros, while industrial/commercial laundry boosters contain phosphate, which we account for through the parameter $LD_{ind,cap,1991-2019}^{P-PO_4}$ that we vary within its uncertainty ranges (see Table 1). This is not reported in panel (b), which only shows the data before 1992. Data source: data in (a) come from ©FAO with a CC BY-NC-SA 3.0 IGO license. Data over the period 1961-2019 are provided in FAO (2021, 2022) and the 1950 value is estimated in this study as the average value over West and East from FAO (1951). Data in (b) are the results of our own calculation from BMU (1989). Data in (c) and (d) are the results of our own calculations from data provided in ©IKW (2005, 2011, 2017, 2019) (https://www.ikw.org/impressum) and dishwasher data from ©Statistisches Bundesamt (SO-DE, Fachserie 15 Heft 1).

3.3 Detergent data

We reconstruct annual detergent P use per capita at NUTS-1 level for the period 1950–2019 based on national and sub-national data, as well as qualitative information. We include the main P compounds present in detergents, namely inorganic phosphates
330 (BMU, 1989; Floyd et al., 2006; Glennie et al., 2002; Van Drecht et al., 2009) and phosphonates that are are poorly degradable organic compounds (Groß et al., 2012; Happel et al., 2021; IKW 2019; Jaworska et al. 2002). We consider laundry detergents (LD) and automatic dishwasher detergents (DD), that are main sources of phosphate (Floyd et al., 2006; Glennie et al., 2002; Van Drecht et al., 2009) and phosphonate (Groß et al., 2012; Happel et al., 2021; IKW 2019; Jaworska et al. 2002). Below, we describe the raw data and their processing ~~for~~ to estimate total use of LD phosphate until around 1991 (Sect. 3.3.1), domestic
335 DD phosphate use (Sect. 3.3.2) and domestic detergent phosphonate use (Sect. 3.3.3) ~~,~~ used in Eq. 2

. To construct the LD phosphate and detergent phosphonate data ~~(Sect. 3.3.1 and 3.3.3) at NUTS-1 level~~, we collect statistical population data for West and East Berlin ~~separately. For 1950–1989, these data are available from the statistical office of Germany (SO-DE, 2022a; SO-EDE 1955–1990; SO-WDE 1970), and for 1990–2000 from the 2001 statistical yearbook for Berlin from the Statistical Office of Berlin (SO-BE, 2001)~~ (see Table 2). To convert phosphate and phosphonate amounts to
340 corresponding P equivalents, we adopt a P content of 0.326 for phosphate (PO_4^{3-}) and of 0.382 for Phosphonate ($-PO(OH)_2$) (Mehlhart et al., 2021).

3.3.1 Laundry detergent (LD) phosphate

We construct separately the temporal development of total (domestic and professional) LD phosphate use per capita in West and East Germany, where different regulations were in place before the reunification of Germany in 1990. For the state of
345 Berlin, the phosphate LD use per capita is calculated as the population weighted average of the values for West ~~Germany~~ and East Germany that apply to the population of West and East Berlin~~)~~, respectively.

In West Germany, the 1980 Phosphate Ordinance sets an upper limit to the phosphate content in LD for both the domestic and professional sectors (PHochstMengV, 1980). In addition, industries in West Germany self-committed to produce phosphate-free domestic LD in 1985 (IKW, 2019). For the period 1975–1988, we adopt the data provided by the German Federal Ministry
350 of the Environment, Nature Conservation and Nuclear Safety (BMU, 1989). These data account for both the domestic and professional sector in West Germany only, and they are given for nine years in the period 1975–1988. We ~~divide these data by the corresponding annual population counts for West Germany to obtain the amount per capita. We~~ then fill the missing years in the period 1975–1988 using linear interpolation.

In East Germany, to our knowledge, no regulations were put into place to limit the phosphate amount in LD before the
355 Phosphate Ordinance (PHochstMengV, 1980) came into force in 1991 ~~,~~ following the German reunification (Kloepfer and Kröger, 1991). Since no data on detergent phosphate are available ~~for East Germany~~, we set the value of the consumption per capita to the average value for West Germany in the period 1975–1979. Further, we assume that this amount per capita remained unchanged until 1984. For the period 1985–1990, we consider that it followed similar temporal dynamics as that of other Eastern European countries. From Glennie et al. (Table 3.2 in 2002), we infer that between 1984 and 1990, LD phosphate

360 consumption decreased by around 50 % for Hungary, Czech Republic and Poland. Therefore, we set the 1990 value equal to 50 % of the 1984 value ~~for East Germany~~.

Since the early 1990s, domestic LD contain virtually no phosphate in Germany (IKW, 2011). "Normal" professional LD that are currently in use are also in general phosphate-free, but professional laundry boosters, utilized to wash heavily soiled textiles, can contain phosphate (Mehlhart et al., 2021). We consider that, in West Germany, household and "normal" professional LD were all phosphate-free from 1991, and that this occurred slightly after in East Germany (from 1992) because regulations were

365 implemented later in the Eastern part. ~~Therefore, we assume that, for the period 1991–2019 for West Germany and 1992–2019 for East Germany, LD phosphate only came from professional laundry boosters, that are estimated as shown in Eq. 6.~~ We then perform linear interpolation to fill the gap between 1988 and 1991 for West Germany and between 1990 and 1992 for East Germany.

370 For the period 1950–1975, only qualitative information on LD phosphate can be found. Although the first LD containing phosphate were developed in the early 1930s (Foroutan-Rad, 1981, as cited in Nork, 1992; ZEODET, 2000), the use of phosphate as detergent builder really started at the beginning of the 1950s (Berth et al., 1983, as cited in Nork, 1992). Therefore, we assume that the use of LD phosphate per capita was equal to 0 in 1950 in both West and East Germany, and we perform linear interpolation for the period 1950–1975. Figure 3b shows the processed data ~~, while the raw data can be seen in Supplementary and Supplementary~~ Fig. S11 the raw data.

3.3.2 Domestic dishwasher detergent (DD) phosphate

We derive the domestic DD phosphate consumption per capita for the period 1950-2017. For this, we disaggregate domestic DD phosphate data at NUTS-0 level, provided by the ~~the~~ German Cosmetic, Toiletry, Perfumery and Detergent Association IKW (2005, 2011, 2017, 2019) ~~at NUTS-0 level~~, to NUTS-1 level using data of household dishwasher ownership at NUTS-1

380 level ~~. We then divide the NUTS-1 level domestic DD phosphate data by the NUTS-1-level population to obtain the amount per capita. We explain~~ from the Statistical Office of Germany (SO-DE, Fachserie 15 Heft 1). We summarise below the different processing steps and ~~illustrate these steps in Supplementary~~ provide further details in section S5.2 ~~and Figure S12~~.

Regarding the detergent data at NUTS-0 level, IKW provides data on the sale of domestic detergent phosphate ~~and phosphonate~~ for 14 years in the period 1989–2019. The 1989 value corresponds to West German only, while the later data (1994–2019) is

385 available at NUTS-0 ~~(country)~~ level. Since domestic LD contain virtually no phosphate anymore since the early in 1990s in Germany (IKW, 2011), we consider that IKW household detergent phosphate data correspond exclusively to household DD phosphate over the period 1994–2019 ~~(1994 being the first year in the 1990s for which data are provided). We divide the~~. We estimate the use per household at NUTS-0 level by dividing the IKW phosphate data by the number of households owning a dishwasher ~~at NUTS-0 level~~.

390 For the period prior to 1994, no DD phosphate data are available, since the data provided by IKW (for year 1989) includes both LD and DD phosphate. We use qualitative information available on DD composition ~~to estimate the DD phosphate consumption prior to 1994.~~. A decline in DD phosphate occurred at the beginning of the 1990s due to the substitution of phosphate by metasilicates in DD (IKW, 2015). Owing to the corrosive nature of metasilicates, these were then removed from DD

formulations from the mid-1990s onward and replaced back by phosphate, hence an increase in the phosphate content after the mid-1990s, as per IKW (2015). After 2012, new formulations were introduced in household DD following the 2012 Detergent Regulation in the European Union, which limits the phosphate content in household DD (EC, 2012)~~, and~~. This resulted in a sharp decrease in DD phosphate consumption. Given these qualitative information, we assume that, for the period 1950-1989, the formulation of DD phosphate was similar to the one in the years 2002–2012. Hence, we set the amount of DD phosphate per household during 1950-1989 to the average 2002–2012 value. We verify that ~~the resulting~~ our estimated total amount of DD phosphate used in 1989 for West Germany (around 9950 $t\ yr^{-1}$) is lower than the value provided by IKW (20,000 $t\ yr^{-1}$) for West Germany which includes all domestic detergent phosphate (LD and DD). We fill using linear interpolation the values for the years in the period 1990-1993 as well as ~~for~~ the years without IKW inventory in the period 1994–2019.

~~We disaggregate the DD phosphate data at NUTS-0 data to~~ To assess the total amount of DD phosphate use at NUTS-1 level~~using NUTS-1 household ownership data. The latter are available from~~, we multiply our estimated NUTS-0 amount per household by household dishwasher ownership data at NUTS-1 level. We derive the latter from household dishwaher ownership data from the Statistical Office of Germany ~~, (SO-DE, Fachserie 15 Heft 1)~~ for West Germany for the period 1962-2018, for East Germany for the period 1993-2018 and at NUTS-1 level in 2008. We ~~consider that the data for West and East Germany adequately cover the relevant time period for dishwasher ownership in Germany. We~~ infer that dishwashers were introduced in West Germany in 1963, since the data report a household dishwasher ownership of 0 % in 1962, and in East Germany in 1991 following the German reunification, since the percentage of households having a dishwasher is low in 1993 (2.7 %). We ~~estimate~~ disaggregate the household dishwasher ownership ~~at~~ data for West and East Germany to NUTS-1 level ~~for~~ over the period 1950–2019~~combining these data and assuming~~. For this we use the 2008 NUTS-1 level data and we assume that all NUTS-1 regions that belong to West and East Germany, respectively, followed the same temporal dynamics. ~~Detailed explanation are provided in Supplementary Sect. S14.~~

We then derive the domestic DD phosphate consumption per capita by dividing the NUTS-1 level domestic DD phosphate data by the corresponding NUTS-1 level population, assuming a spatially uniform DD phosphate detergent use per capita within a given NUTS-1 region. ~~This assumption is supported by the fact that, from household dishwasher ownership data provided for seven region types having different degrees or urbanisation for the year 1998, we infer that the~~ The degree of urbanisation has a limited impact on the percentage of households and population owning a dishwasher (see ~~more detail in Supplementary Sec~~Supplementary Sect. S5.2.3). We report the final processed data at NUTS-1 level in Fig 3c.

3.3.3 Domestic detergent phosphonate

We estimate domestic detergent phosphonate use per capita ~~based on NUTS-0 data for the period 1994-2019 and based on data or qualitative information for West and East Germany~~ for ~~the period prior to 1994. Similar to LD phosphate, we calculate detergent phosphonate use~~ West and East Germany based on IKW data from 1989 for West Germany and from 1994 for East Germany (2005, 2011, 2017, 2019) and based on qualitative information for the earlier period. We calculate the amounts for Berlin as the population weighted average of the values for West ~~Germany~~ and East Germany.

~~We adopt the~~ The domestic detergent phosphonate data ~~available from IKW (2005, 2011, 2017, 2019)~~ from IKW are available for the year 1989 for West Germany and for the period 1994–2019 at NUTS-0 level. These quantities correspond to the total detergent phosphonate use in LD and DD (and possibly to a lesser extent other cleaning products). Domestic detergent phosphonate is an important component that showed large increases in the period 1994-2019. Notably, in 2019, it is equal to 7613 $t\ yr^{-1}$, which is almost ten times the domestic DD phosphate amount (829 $t\ yr^{-1}$).

From the IKW data for the period 1994-2019, we assess the detergent phosphonate use per capita at NUTS-0 (country) level. We assume a spatially uniform value across Germany due to a lack of further spatial information. In addition, for West Germany, we assess the detergent phosphonate use per capita based on the 1989 IKW inventory and we perform linear interpolation between 1989 and 1994.

We utilize qualitative information on detergent phosphonate content ~~to construct the phosphonate consumption~~ for the period before 1989 for West Germany and before 1993 for East German. Phosphate-free LD typically contain phosphate, while the phosphonate content in phosphate-based LD is between 0 to half that of phosphate-free LD ~~(see Table 2.2 in Glennie et al., 2002)~~ (Table 2.2 in Glennie et al., 2002). In addition, domestic DD phosphonate were likely to be very small before 1980 in West Germany, since the proportion of households having a dishwasher was limited (less than 15 % of ~~household~~ households before 1978) and no DD were used in East Germany before 1991. Consequently, we neglect the domestic phosphonate use before LD phosphate content started to be reduced in the 1980s, i.e. we set the phosphonate amount to 0 for the period 1950-1979 for West Germany and for the period 1950-1984 for East Germany. We then assume a linear increase in household detergent ~~phosphate~~ phosphonate use per capita for the period 1979-1989 for West Germany and 1984-1994 for East Germany. Figure 3d depicts the processed data and ~~the raw data are shown in~~ Supplementary Fig. S14 the raw data.

3.4 Total population connection to the sewer system and to wastewater treatment

We collect and process data at NUTS-1 level of total population connection to the sewer system and to public WWTPs with different types of treatment, as well as to foreign and industrial WWTPs when available. These data come from various sources, mainly from statistical reports, and are complemented by qualitative information from Seeger (1999). In general, data are available from 1975 onward for West Germany and 1991 onward for East Germany. Table 2 provides an overview of these data~~, while Supplementary Sect. S6 reports detailed information on the data availability and sources for each NUTS-1 region.~~

~~We estimate the variables that are required to assess the N and P emissions, namely the fraction of total population whose wastewater is not handled in the public wastewater system ($NoCon = 1 - T_{sewer-wwtp,all}$; see Eq.12), and the fraction of total population that is connected to the sewer system and to the different treatment types having different removal efficiencies ($T0_{all}, T1_{all}, T23_{all}^{noN}, T23_{all}^{noP}, T3_{all}^{N},$ and $T3_{all}^{P}$; see Eq.12).~~ Figure 4 reports the final processed data and shows how the total population of the different NUTS-1 region was progressively connected to more advanced types of wastewater treatment. ~~Supplementary Fig. S18 further reports on development of $T23_{all}^{noP}$ and $T3_{all}^{P}$, that take values similar as $T23_{all}^{noN}$ and $T3_{all}^{N}$, respectively, shown in Fig. S18.~~ The data are briefly described below, while Supplementary Sect. S6 reports further details, including an analysis on the value of our data compared to the readily available data provided by the statistical office of the European Union (Eurostat, 2016, 2023).

For the state of Berlin, we consider that all wastewater that is collected in the sewer system but not treated in WWTPs is applied to agricultural soils in sewage farms ($TFarm_{all}$ of Eq.12; see Fig. 4), given the importance of this type of wastewater treatment, as documented for instance in Lottermoser (2012). This occurs up to year 1982, as in later years the data indicate that all wastewater collected in the sewer system is treated in WWTPs. We do not consider sewage farms in other NUTS-1 regions, due to a lack of information.

We account not only for wastewater treated in public WWTPs via collection in the sewer system, but also, when the data is available, for wastewater collected in cesspits (sealed tanks) and transported by trucks to public WWTPs(as shown in Supplementary Fig. S17). We do not consider independent treatment in small wastewater treatment plants because of a lack of data on such treatment. We consider that the resident population of a given NUTS-1 region is connected to WWTPs located in that NUTS-1 region only. An exception is the state of Berlin for which a large part of the population is connected to WWTPs in the neighbouring state of Brandenburg (35–65 % of the population over the period 1991–2019). We estimate separately the population connection to WWTPs located in Berlin and Brandenburg, as detailed in Supplementary Tables S37–S38 and Fig S19. We consider that the part of the population of Berlin whose wastewater is handled in Brandenburg is all connected to WWTPs, while the part of the population of Berlin whose wastewater is handled in Berlin can be either connected to sewage farms or WWTPs, or not connected to the public wastewater system.

Regarding the data of population connection to the different types of wastewater treatment, most data refer to the population treated in WWTPs located in a given NUTS-1 region and we consider that the same relative proportions of the different treatment types apply to the resident population (apart from Berlin and Brandenburg, as explained above). In a few cases, we use data of the volume of wastewater treated following the different treatment types to derive temporal changes for the population connection to these treatment types. Details regarding the data on the types of wastewater treatment for each NUTS-1 region are reported in Supplementary Tables S27–S43.

For the earlier years where statistical data are not provided, we use qualitative information provided by Seeger (1999) to set a starting year of the different treatment types. Accordingly, we assume that sewer connection started in 1880, primary wastewater treatment in 1910, secondary treatment in 1950, and tertiary treatment with targeted N and P removal in 1980 in West Germany and 1990 in East Germany. We then consider a linear development up to the first year with available data. This is consistent with the study of Seeger (1999) in which the development of wastewater treatment in Germany is described as a "steady process". To ensure that the sum of the population data is equal to 1, we estimate $T23^{noN}$ and $T23^{noP}$ as the difference between 1 and the sum of the other population connection data. Regarding the population connected to WWTPs via cesspits, in absence of further information, we consider that it is equal to zero before the first year with data and after the last year with data. For all missing data, we We perform linear interpolation between the years for which data are provided. population connection data are not provided.

We examine the value of our processed data (reported in Fig. 4 compared to the readily available data provided by the statistical office of the European Union at NUTS-0 level (Eurostat, 2023) for the period 1991–2016 and at NUTS-2 level (Eurostat, 2016) for the year 2010 (shown in Supplementary Fig. S20). Notably, our data provide estimates of population connection at a higher spatial resolution (NUTS-1 level) and for a longer time period (1950–2019) than the Eurostat data . In

particular, our data allow to capture the differences between West and East Germany that can only be retrieved for the year 2010 from the Eurostat data. In addition, the Eurostat dataset provides data of population connection to total tertiary treatment and do not specify about the targeted N and P removal capacities as our data. This leads to an overestimation of the efficiency of N and P removal.
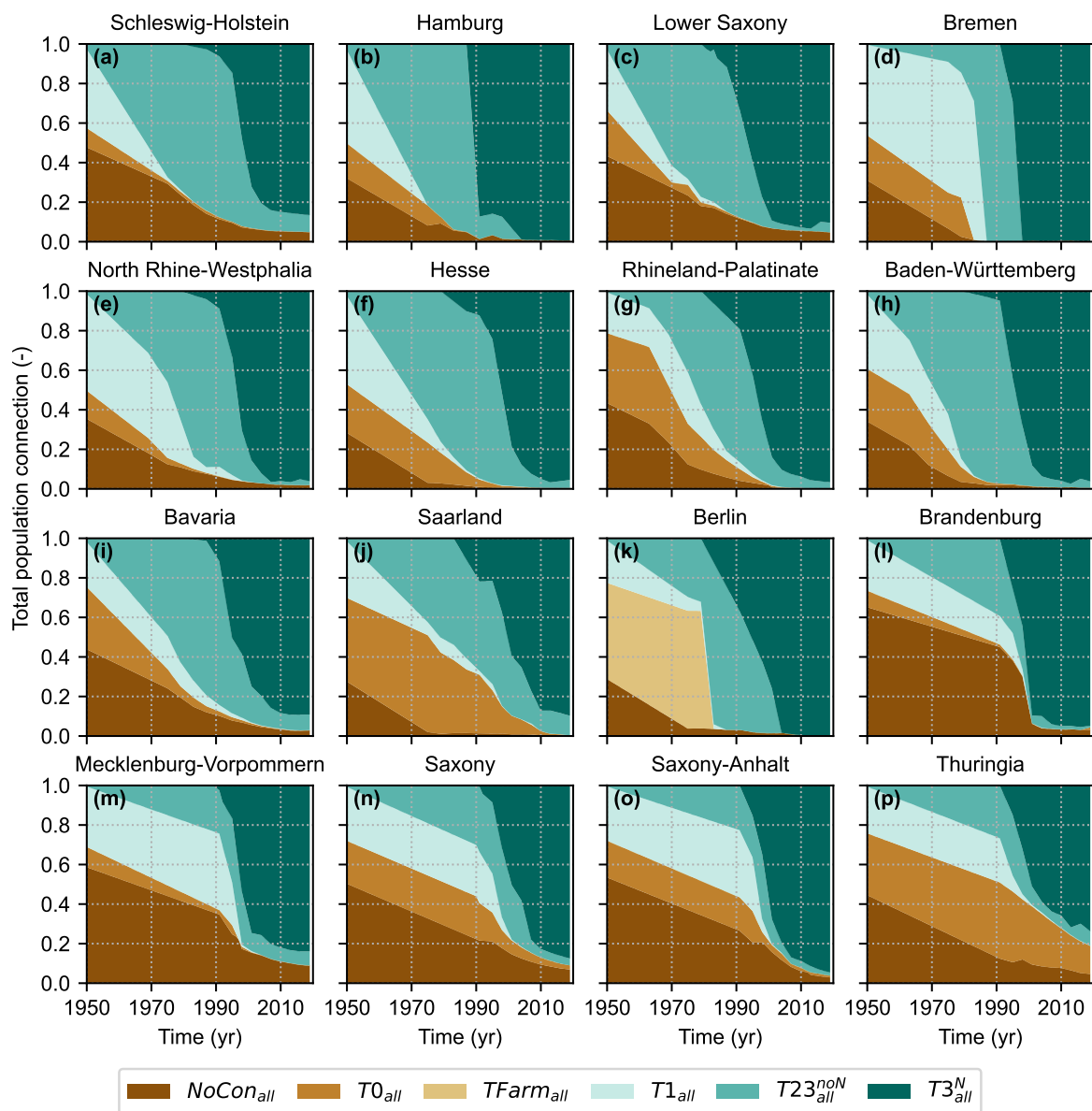
Figure 4. Processed fractions of total population connection at NUTS-1 level for the period 1950–2019, that were used to derive the N and P point sources data in this study. Specifically, the figure reports the fraction of total population that is not connected to the sewer system nor to WWTPs ($NoCon_{all}$) the fraction of total population connected to the sewer system but not to WWTPs ($T0_{all}$), the fraction of total population connected to sewage farms ($TFarm_{all}$), the fraction of total population connected to primary (mechanical) treatment ($T1_{all}$), the fractions of total population connected to secondary (biological) treatment and tertiary (advanced) treatment without targeted N removal ($T23_{all}^{noN}$), and the fraction of total population connected to tertiary treatment with target N removal ($T3_{all}^{N}$). Supplementary Fig. S18 further reports on development of $T23_{all}^{noP}$ and $T3_{all}^{P}$, that take similar values as $T23_{all}^{noN}$ and $T3_{all}^{N}$. Data source: the data shown in this figure were elaborated in this study building on data from the statistical offices of Germany and the different federal states as well as additional sources, as detailed in Supplementary Sect. S6.

Table 3. Description of the observational data of WWTPs incoming and outgoing N and P loads used for downscaling and model evaluation.

Nutrient species	WWTP type	Load type	Spatial feature[a]	Temporal feature		Source
			Level	Extent	Frequency	
total N, total P	capacities ≥ 2000 PE	incoming, outgoing	WWTP	2010–2018	2 years	EU dataset (Waterbase; EEA, 2023b)[f]
total N, total P	capacities < 2000 PE [b]	incoming, outgoing	WWTP	2015 or 2016 [c]	one value	DE dataset (Büttner, 2020)[g]
total N, total P, inorganic N [d]	capacities ≥ 2000 PE, all capacities [d]	incoming, outgoing [d]	NUTS-1	1987–2020 [e]	variable	Statistical reports of the statistical offices of Germany and the federal states, and situation reports of the authorities of the German federal states (see details in Supplementary Sect. S8)

Notations: PE is the Population Equivalent.
[a] The spatial extent is Germany for all data. [b] The DE dataset mostly includes WWTPs that have a design capacity lower than 2000 PE, but it also encompasses some WWTPs that have a design capacity higher or equal to 2000 PE. [c] For each WWTP, data are available either for year 2015 or for year 2016. [d] The nutrient species, WWTP types and load types vary across NUTS-1 regions and data sources (see details in Supplementary Sect. S8, in particular Tables S45 and S60). [e] Period over which some data are available. The temporal extent varies across NUTS-1 regions, nutrient species, WWTP types, load types (see details in Supplementary Sect. S8, in particular Tables S45 and S60). [f] ©European Environment Agency (EEA), CC BY 4.0 license. [g] Data collection was supported by research projects of UFZ under a confidentiality agreement from the German states' authorities.

3.5 Observations of WWTP incoming and outgoing N and P emissions

We construct datasets of observations of WWTP N and P emissions though the collection and processing of data from different sources, as ~~summarise~~ summarised in Table 3~~detailed below. We derived~~. We derive a gridded map of WWTP load from data at WWTP level (presented in Sect. 3.5.1) to downscale our NUTS-1 level N and P point sources estimates (as explained in Sect. 2.2)~~and data~~. Data at NUTS-1 level (presented in Sect. 3.5.2) are used to estimate the parameters/coefficients of the N and P point sources model ~~. Furthermore information on NUTS-1 level N and P emissions is also used~~ and to evaluate our N and P point sources estimates (as explained in Sect. 4).

3.5.1 Observations of WWTP incoming and outgoing N and P emissions at WWTP level

We merge two datasets of observations of WWTP incoming and outgoing N and P ~~emissions~~loads around the year 2016, namely (1) the Urban Waste Water Treatment Directive (UWWTD) Waterbase dataset of the European Environmental Agency (EEA, 2023b, hereafter referred to as the EU dataset), including 4302 WWTPs in 2016, and (2) data coming from the authorities of the German federal states and made available in the dataset of Büttner (2020) (hereafter referred to as the DE dataset)~~. The EU dataset reports data~~, including 6361 WWTPs for the year 2015 or 2016 depending on the plants. We could not derive WWTP

25

loads beyond the years 2016, since the EU dataset only (and not the DE dataset) provides data beyond 2016 (namely every two years in the period 2010–2018~~for a total of 4394 WWTPs. Data for some of these WWTPs are only reported for some years,~~

₅₁₅ ~~because some WWTPs were closed or opened during this time period or because of possible omissions in the reporting. These 4394~~). The EU dataset consists of WWTPs that treat wastewater in agglomerations larger than 2000 Population Equivalent (PE). ~~The DE dataset provides data for 6361 WWTPs for the year 2015 or 2016 depending on the plants. It complements the EU dataset, since it includes WWTPs treating wastewater in~~ This is why this dataset needs to be complemented by the DE dataset, that includes WWTPs for agglomerations smaller than 2000 PE.

₅₂₀ ~~We construct a dataset of WWTPs incoming and outgoing N and P loads around the year 2016, by combining the data reported for the year 2016 in the EU dataset, including 4302 WWTPs that were not indicated as being inactive in 2016, and the 2015–2016 data provided by the DE dataset. This results in a dataset encompassing~~ We refer to Supplementary Sect. S9 for details on the data processing, while a summary is reported below.

The combined dataset encompasses 9006 WWTPs, since we establish that 1657 WWTPs are present in both the EU and

₅₂₅ DE datasets. To avoid duplication, we merged the records for these WWTPs by selecting the coordinates from the DE dataset. We do not account for the uncertainty in the coordinates arising from the discrepancies between the EU and DE dataset, since the distance calculated between the coordinates reported in the two datasets is ~~lower than 1 km for 96 % of the WWTPs (see Supplementary Fig S32). The uncertainty on the coordinates is thus mostly within to~~ mostly within our target grid resolution of 0.015625°~~, which is on average 1.74 km x 1.09 km over Germany. We derive a lower and upper bound estimate of the incoming~~

₅₃₀ ~~and outgoing N and P loads for each WWTP. This allows to account for the uncertainties arising from the low precision of the load values reported in the EU dataset and from the discrepancies in the values of the load between the EU and DE dataset for some of the 1657 overlapping WWTPs. We find that the width of the uncertainty interval is in general contained, amounting to less than 25 % of the mean estimate (see Supplementary Table S78). The uncertainty interval is however larger for the outgoing N load for the NUTS-1 region of Thuringia, where it is around 72 % of the mean estimate. This is due to inconsistencies in the~~

₅₃₅ ~~load record in the EU (for year 2016) and DE dataset (for year 2015) for one of the WWTPs having large differences among these two datasets.~~

. We fill the missing values of the N and P load using values for the other years provided in the EU dataset when available or we perform extrapolation based on value of the WWTP entering load expressed in PE provided in both the EU and DE datasets. We ~~verify~~ find that these filled values account for less than 8% of the total load at NUTS-1 level~~(see Supplementary Fig. S34).~~

₅₄₀ . We also verify that the uncertainty in the loads of the combined dataset is generally contained. The uncertainty results from the low precision of the load values reported in the EU dataset, the discrepancies in the values of the load between the EU and DE dataset for the 1657 overlapping WWTPs and our gap-filling procedure. Our processed dataset comprises values of the incoming and outgoing N and P loads for over 99% of the 9006 WWTPs, while for only 39 WWTPs no values at all of the N and P loads can be estimated and for 16 WWTPs no value of the incoming N and P loads and the outgoing N loads

₅₄₅ can be derived~~(see Supplementary Fig. S33). Details on the procedure for identifying matching WWTPs in the EU and DE datasets, the derivation of the upper and lower bound estimates of the loads, as well as the gap-filling procedure are reported in Supplementary Sect. S9.~~.

Finally, we create a gridded map of WWTPs ougoing N and P emissions at a resolution of 0.015625°. For this, for each grid cell, we sum the mean values of the load (average of the lower and upper bound estimated) over all WWTPs that fall within that grid cell. We move the location of 137 WWTPs between 500 and 2500 m to ensure that they fall within the NUTS-1 regions reported in the EU and DE datasets.

3.5.2 Observations of WWTP incoming and outgoing N and P emissions at NUTS-1 level

We compile observational data of public WWTP incoming and outgoing N and P load at NUTS-1 level from different reports of the statistical offices of Germany and the federal states, as well as situation reports of the authorities of the German federal states (see details in Supplementary Sect. S8 and Fig. S28–S31). The data span over the period 1987-2020, depending on the NUTS-1 regions. While P data always correspond to total P, N load data are provided either as inorganic N or as total N. Some data only include WWTPs having a design capacity higher than 2000 PE and we do not further process these data. Supplementary Sect. S8 provide details on the processing, which is summarized in the following paragraph.

We find total N observations for half of the NUTS-1 regions, while inorganic N and total P observations are available for all NUTS-1 regions. Since our point sources estimation focuses on total N, for the NUTS-1 regions for which total N data are not provided at all or for a limited time period, we estimate total N from inorganic N observations for outgoing N load. For this, we apply a ratio of total N to inorganic N that we derive based on the relative value of these two quantities during their common periods of availability for the different NUTS-1 regions. We derive lower and upper bound estimates of the outgoing load when possible, accounting for the uncertainties in the total N to inorganic N ratio and from the fact that load measurements may not include all WWTPs (even for data referring to WWTPs of all design capacities).

We note that incoming N and P load data, that are mostly provided in the situation reports, may result from both measurements and estimates based on the PE and a constant coefficient of the N and P emissions per PE when observations are missing (LU-RP and LU-RP MKUEM, 2021; MUNLV-NW, 2020). However, this uncertainty cannot be quantified, since no

information on the relative proportion of the measured and estimated loads in the NUTS-1 level data is provided. In addition, data of incoming N loads cannot be obtained for all NUTS-1 regions.

4 Methods for parameter estimation and model evaluation at NUTS-1 level

To estimate the 18 model parameters (coefficients), we generate a prior parameter sample of size 100,000 using latin hypercube sampling and uniform distributions, from the ranges reported in Table 1, and using the SAFE toolbox (Pianosi et al., 2015). These ranges were either determined from previous studies, in particular Morée et al. (2013), or from our collection and processing of additional data –(details in Supplementary Sect. S3). Food waste data provided by FAO and SIK (2011) and Noleppa and Cartsburg (2015) are used to establish the range for the fraction of protein supply wasted at the distribution and consumption level (f_{waste}^{pro}). Data of industrial/commercial laundry and dishwasher detergent phosphate use in 2015 and industrial/commercial detergent phosphonate use in 2008 and 2015 from the German Industry Association for Hygiene and Surface Protection for Industrial and Institutional Applications (IHO) are adopted to set the range for detergent parameters ($f_{ind:dom,1950-2012}^{DD,P-PO_4}$, $f_{ind:dom,1950-2015}^{D,P-PO(OH)_2}$, and $LD_{ind,cap,1991-2019}^{P-PO_4}$). Further details on the determination of the ranges can be found in Supplementary Sect. S3.

For each of the 16 NUTS-1 regions of Germany, we perform Monte Carlo simulations for the period 1950–2019 using the prior parameter sample and we identify an ensemble of posterior (behavioural) simulations that are most consistent with both the N and P outgoing load observations described in Sect. 3.5.2. We thereby account for the uncertainty arising in particular from the observations that are uncertain and have a limited temporal coverage. Specifically, we select the 100 (behavioural) simulations that have the lowest value of the sum of the two root mean square errors for N and P between the observed and simulated WWTP outgoing load for N and P. The root mean square error for the i-th NUTS-1 region and j-th posterior simulation (realisation) is computer as follows:

$$RMSE(i,j) = \sqrt{\frac{\sum_{t \in E_{obs(i)}} err(t,i,j)^2}{N_{obs}(i)}} \frac{100}{\mu_{obs}(i)} \quad (21)$$

where $RMSE(i,j)$ (%) is the root mean square error, $err(t,i,j)$ ($kg\ yr^{-1}$) is the error between the simulated and observed (N or P) load at year t, $E_{obs}(i)$ is the ensemble of years for which observations are available, $N_{obs}(i)$ is the number of observations, and $\mu_{obs}(i)$ ($kg\ yr^{-1}$) is the mean of the observations. $RMSE$ is normalised by the mean of the observations to allow comparability of the realisation performance across nutrient species (N and P) and NUTS-1 regions. Since the observations can have an lower and upper uncertainty bound (as explained in Sect. 3.5.2), we calculate the error between simulations and observations as follows:

$$err(t,i,j) = max(X_{obs}^{lb}(t,i) - X_{sim}(t,i,j), 0) + max(X_{sim}(t,i,j) - X_{obs}^{ub}(t,i), 0) \quad (22)$$

where $X_{sim}(t,i,j)$ ($kg\ yr^{-1}$) is the simulated load, and $X_{obs}^{lb}(t,i)$ ($kg\ yr^{-1}$) and $X_{obs}^{ub}(t,i)$ ($kg\ yr^{-1}$) are the lower and upper bounds of the observations, respectively. From Eq.(22), the error is equal to zero when the simulation is within the observation uncertainty bounds.

We also calculate two additional metrics to evaluate the accuracy of the selected 100 posterior simulations with respect to the observations, namely: the mean absolute error for the *i*-th NUTS-1 region and *j*-th posterior realisation, denoted as $MAE(i,j)$ (%), and the distance of the observations to the posterior ensemble for the *i*-th NUTS-1 region and at time *t*, denoted as $DIST(t,i)$ (%), defined as follows:

$$MAE(i,j) = \frac{\sum_{t \in E_{obs(i)}} err(t,i,j)}{N_{obs}(i)} \frac{100}{\mu_{obs}(i)} \tag{23}$$

$$DIST(t,i) = 100 \frac{max(X_{obs}^{lb}(t,i) - X_{sim}^{ub}(t,i),0) + max(X_{sim}^{lb}(t) - X_{obs}^{ub}(t,i),0)}{\mu_{obs}(i)} \tag{24}$$

where $X_{sim}^{lb}(t,i)$ *(kg yr^{-1})* and $X_{sim}^{ub}(t,i)$ *(kg yr^{-1})* are the lower bound (minimum value over the 100 realisations) and upper bound (maximum value over the 100 realisations) of the posterior simulation ensemble, respectively. $DIST$ is equal to zero when the observation uncertainty interval intersects the simulation uncertainty interval, which corresponds to a high accuracy of the posterior ensemble.

We further evaluate the consistency of the posterior realisations by comparing ~~then~~ them visually with the additional observations of WWTP outgoing load available, that is observations of N inorganic load as well as N and P load for WWTPs that have a design capacity larger than 2000 PE (see Sect. 3.5.2). We also assess the performance metrics of Eq. (21, 23, and 24) for the WWTP incoming total N and P load described in Sect. 3.5.2, as an another plausibility check to our reconstructed ~~data-sets~~dataset. Due to the limited ~~of~~ coverage and uncertainty in data of WWTP incoming total N and P load, we only use these data for evaluation purposes and not for the identification of the posterior realisations.

~~Since a large part of the population of Berlin is connected to WWTPs in the neighbouring state of Brandenburg (as detailed in Sect. 3.4), we estimate separately the N and P emissions from the wastewater generated in Berlin, and handled in Berlin and in Brandenburg. To reflect this part, in~~ In our realisations, the N and P emissions for Berlin come from the wastewater that is both generated and handled in Berlin. For Brandenburg, N and P emissions include the part of wastewater that is generated in Berlin (which corresponds to a large part of the population of Berlin, as detailed in Sect. 3.4) along with its own (Brandenburg) contribution.

5   Results

5.1   Evaluation of the point source estimates at NUTS-0 and NUTS-1 level

The behavioural realisations show a good consistency with the observations of WWTP outgoing N and P loads ($J_{ps,wwtpOut,all}^{N}$ and $J_{ps,wwtpOut,all}^{P}$, respectively) for the 16 NUTS-1 regions. Minimum values of $RMSE$ and MAE across the 100 behavioural realisations, obtained for the best performing behavioural realisations, are lower than 17.5 % and 15.1 %, respectively, and maximum values lower than 24 % and 20 %, respectively (Table 4). We also find that the posterior realisations ensemble of the WWTP outgoing N load, which is bounded by their respective minimum and maximum values across the 100

posterior realisations, intersects all observation uncertainty bounds for the regions of Saarland, Brandenburg and Mecklenburg-Vorpommern (values of $DIST$ are equal to 0 in Table 4). In all other regions, the posterior realisations of the WWTP outgoing N and P load captures only some of the observational bounds, with positive value of $DIST$. Nonetheless, the mean value of $DIST$ across the observations is always lower than 12 %, which demonstrates an overall good accuracy of the WWTP

645 outgoing load realisations. From Fig. 5, 6, we qualitatively appreciate the consistency of ~~posterior realisation ensemble~~ the posterior realisation ensembles with the observational data that were used for parameter estimation (dark brown circles and crosses) and additional data, in particular the observations for WWTPs having a design capacity higher than 2000 PE for Saxony-Anhalt and Thuringhia (yellow and light green dots in Fig.5o,p and Fig.6o,p). The outgoing P load for the NUTS-1 region of Schleswig-Holstein shows relatively lower performance, with values of $RMSE$ and MAE in the range 83–95 % and

650 30–60 %, respectively (Table 4). From Fig. 6a, we see that the posterior realisation ensemble includes all observations except one in 1987 which is underestimated. Further observations around this time period would be required to determine whether this mismatch is due to errors in the observations or missing P contributions in our model and input data.

We evaluate quantitatively the plausibility of the posterior realisations of WWTP incoming N and P load ($J^N_{wwtpIn,all}$ and $J^P_{wwtpIn,all}$, respectively) ~~with observations when~~ when observations are available (Table 4~~). A visual depiction of the load~~

655 ~~time series is provided in~~, Supplementary Fig. S36,S37). For incoming N load, performance are overall satisfying with average values of $RMSE$ and MAE in the range 7.5–20.2 %, and 6.2–16.3 %, respectively, although the performance varies across the posterior ensembles with values up to 30–40 % for Bremen, Rhineland-Palatinate, as well as for Mecklenburg-Vorpommern where however only one observation point is available. $DIST$ also demonstrates good skill in modelled estimates, with values lower than 20 % for all NUTS-1 regions and always equal to zero for five out of the eight NUTS-1 regions with observations.

660 Additional observations for WWTPs with design capacity higher than 2000 PE confirm the plausibility of the realisations for Lower-Saxony and Saxony-Anhalt but suggest that our realisation underestimate the load in Saarland between 2008 and 2019 (Fig S36). For incoming P load, the posterior realisations tend to underestimate the observations~~and the performance are~~ . The performance is in general lower than for the incoming N load and the outgoing N and P load, with average values of $RMSE$ and MAE both in the range 12.7–40.1 % and average values of $DIST$ in the range 0–26 %.

665 ~~From~~ Fig. 5 ~~, 6 , we observe~~ and 6 depict a substantial uncertainty reduction in the posterior realisation ensemble (dark grey shaded areas) compared to the prior realisation ensemble (light grey shaded areas) of the WWTP outgoing N and P load, in particular over the most recent period. Specifically, the width of the posterior realisation ensemble (delimited by the minimum and maximum values across the posterior realisations) normalised by the mean estimate is, for most NUTS-1 regions, 1.17 to 7.7 time larger on average over the (past) period without observations compared to the (recent) period with observations~~(which~~

670 ~~starts~~ . The latter starts ~~starts~~ between 1987 and 2008, depending on the NUTS-1 regions and whether it is N or P~~) for the treated component of the point sources ($J^N_{ps,wwtpOut,all}$ and $J^P_{ps,wwtpOut,all}$; Fig. 5,6) and 1.1 to 7.2 times larger for the total point sources ($J^N_{ps,tot,all}$ and $J^P_{ps,tot,all}$; Supplementary Fig. S49,S50).~~ . Two NUTS-1 regions show a different pattern, namely Schleswig-Holstein where the normalised width is larger over the period with observations and North Rhine-Wesphalia where the normalised width is similar over the two time periods. Nevertheless, the absolute width of the posterior realisation ensemble

675 (expressed in $kg\ ha^{-1}\ yr^{-1}$) ~~for $J^N_{ps,wwtpOut,all}$ and $J^P_{ps,wwtpOut,all}$~~ ($J^N_{ps,tot,all}$ and $J^P_{ps,tot,all}$~~) is always~~ is always larger (1.8

to 63.5 times ~~larger (2.0 to 73.3 time larger) on average~~on average) over the period without observations~~compared to the period with observations~~. Specifically, the average width of the posterior ensemble for the ~~N point sources (total and treated part)~~ WWTP outgoing N (P) load is in the range 0.9~~–31.6~~–29 $kg\ ha^{-1}\ yr^{-1}$ (~~width normalised by the mean: 24.8–58.2 %~~0.11–4.3 $kg\ ha^{-1}\ yr^{-1}$) over the period without observations and in the range 0.03–2.2 $kg\ ha^{-1}\ yr^{-1}$ (~~width normalised by the mean: 6.3–43.6 %) over the period with observations. The average width of the posterior ensemble for the P point sources (total and treated part) is in the range 0.11–4.6 $kg\ ha^{-1}\ yr^{-1}$ (width normalised by the mean: 19.3–48.9 %) over the period without observations and in the range~~ 0.00–0.~~14~~.15 $kg\ ha^{-1}\ yr^{-1}$~~(width normalised by the mean: 5.6–65.8 %)~~ over the period with observations. Further details on the region-wise uncertainty intervals are reported in Supplementary Tables S80,S81 for both the WWTP outgoing load and the total point sources. For the other model variables beyond the total and treated point sources, the uncertainty reduction in the posterior compared to the prior ensemble is in general not as pronounced, as can be seen from Supplementary Fig. S36,S37 for WWTP incoming N and P load and from Fig. S51–S61 for ~~the other pathways~~of N and P gross emissions~~other pathways. This is likely related to the fact that no observational data of these components are used to constrain their respective parameters.

Table 4. Performance metrics calculated between the 100 posterior realisations of the WWTP incoming and outgoing N and P load and the observations for the 16 NUTS-1 regions of Germany.

NUTS-1 region	$J_{ps,wwtpOut,all}^N$			$J_{ps,wwtpOut,all}^P$			$J_{wwtpIn,all}^N$			$J_{wwtpIn,all}^P$		
	RMSE (%)	MAE (%)	DIST (%)	RMSE (%)	MAE (%)	DIST (%)	RMSE (%)	MAE (%)	DIST (%)	RMSE (%)	MAE (%)	DIST (%)
Schleswig-Holstein	$14.6_{10.0}^{21.8}$	$10.6_{5.3}^{18.8}$	$1.5_{0.0}^{11.2}$	$89.0_{83.0}^{94.7}$	$42.6_{30.1}^{60.1}$	$4.3_{0.0}^{47.5}$	n.a.	n.a.	n.a.	$16.9_{8.7}^{27.7}$	$14.2_{5.3}^{26.2}$	$4.7_{0.0}^{16.0}$
Hamburg	$10.9_{10.8}^{11.3}$	$8.8_{8.5}^{9.8}$	$6.3_{0.0}^{18.5}$	$7.5_{7.0}^{7.8}$	$6.9_{6.5}^{7.3}$	$3.3_{0.2}^{6.2}$	$14.0_{4.9}^{27.7}$	$13.0_{4.2}^{27.1}$	$0.1_{0.0}^{1.5}$	$18.7_{6.3}^{32.2}$	$17.5_{5.2}^{31.5}$	$1.9_{0.0}^{7.4}$
Lower Saxony	$9.7_{9.3}^{10.3}$	$6.2_{5.4}^{6.9}$	$2.9_{0.0}^{11.0}$	$11.7_{11.2}^{12.4}$	$9.2_{8.4}^{9.9}$	$4.7_{0.0}^{15.4}$	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Bremen	$13.0_{12.8}^{13.6}$	$8.5_{7.4}^{10.3}$	$4.2_{0.0}^{25.1}$	$15.8_{15.2}^{16.2}$	$13.1_{12.6}^{13.6}$	$9.4_{0.0}^{27.1}$	$20.2_{12.4}^{33.7}$	$16.3_{8.3}^{31.5}$	$3.9_{0.0}^{19.3}$	$36.5_{23.4}^{47.4}$	$34.7_{20.4}^{46.1}$	$19.6_{1.3}^{39.0}$
North Rhine-Westphalia	$19.8_{13.6}^{24.0}$	$15.9_{12.5}^{18.8}$	$2.1_{0.0}^{6.8}$	$15.1_{10.1}^{21.7}$	$12.6_{10.1}^{18.5}$	$1.6_{0.0}^{13.3}$	$7.5_{6.0}^{17.3}$	$6.2_{4.5}^{16.4}$	$0.0_{0.0}^{0.0}$	$30.7_{17.2}^{39.2}$	$30.3_{16.3}^{38.9}$	$16.1_{7.8}^{22.7}$
Hesse	$7.7_{4.5}^{10.3}$	$4.5_{1.9}^{7.2}$	$0.6_{0.0}^{7.6}$	$19.6_{17.5}^{23.1}$	$15.8_{13.8}^{19.9}$	$7.1_{0.0}^{43.5}$	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Rhineland-Palatinate	$8.9_{7.4}^{10.7}$	$6.0_{5.1}^{7.3}$	$1.3_{0.0}^{6.2}$	$9.4_{7.8}^{11.0}$	$7.0_{5.5}^{9.8}$	$1.7_{0.0}^{11.3}$	$15.8_{5.7}^{33.0}$	$14.8_{4.8}^{32.4}$	$0.0_{0.0}^{0.0}$	n.a.	n.a.	n.a.
Baden-Württemberg	$10.2_{9.9}^{10.8}$	$10.0_{9.7}^{10.4}$	$5.9_{5.0}^{7.0}$	$17.1_{16.2}^{17.6}$	$15.8_{15.1}^{16.3}$	$11.8_{1.9}^{25.3}$	$12.0_{5.8}^{25.9}$	$10.8_{5.4}^{25.5}$	$0.0_{0.0}^{0.0}$	$30.7_{10.8}^{41.5}$	$30.5_{10.0}^{41.3}$	$9.8_{1.4}^{13.8}$
Bavaria	$15.5_{14.4}^{16.5}$	$12.7_{11.3}^{13.9}$	$7.6_{0.0}^{21.0}$	$5.4_{4.7}^{6.6}$	$4.4_{3.4}^{6.2}$	$0.9_{0.0}^{3.9}$	$15.5_{8.4}^{27.2}$	$13.6_{7.5}^{26.1}$	$3.3_{0.0}^{12.4}$	$40.1_{26.2}^{49.7}$	$40.0_{26.0}^{49.6}$	$26.0_{22.5}^{30.0}$
Saarland	$0.2_{0.0}^{1.4}$	$0.1_{0.0}^{0.6}$	$0.0_{0.0}^{0.0}$	$2.5_{1.9}^{3.3}$	$2.1_{1.5}^{2.7}$	$0.1_{0.0}^{0.4}$	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Berlin	$7.2_{6.8}^{8.6}$	$4.6_{4.1}^{6.4}$	$1.5_{0.0}^{5.8}$	$9.4_{8.4}^{10.3}$	$7.5_{6.6}^{9.2}$	$2.9_{0.0}^{11.7}$	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Brandenburg	$4.6_{3.3}^{8.0}$	$2.7_{1.6}^{4.8}$	$0.0_{0.0}^{0.0}$	$10.9_{8.0}^{13.2}$	$8.6_{6.1}^{11.4}$	$0.8_{0.0}^{4.2}$	n.a.	n.a.	n.a.	$12.8_{0.0}^{26.7}$	$12.7_{0.0}^{26.6}$	$0.0_{0.0}^{0.0}$
Mecklenburg-Vorpommern	$0.9_{0.4}^{1.9}$	$0.5_{0.2}^{1.1}$	$0.0_{0.0}^{0.0}$	$11.2_{10.4}^{12.1}$	$8.7_{6.9}^{11.5}$	$3.7_{0.0}^{11.6}$	$10.9_{0.1}^{38.4}$	$10.9_{0.1}^{38.4}$	$0.0_{0.0}^{0.0}$	$31.5_{12.9}^{43.2}$	$31.5_{12.9}^{43.2}$	$12.9_{12.9}^{12.9}$
Saxony	$11.8_{11.6}^{12.2}$	$10.6_{10.0}^{11.2}$	$7.2_{0.0}^{14.8}$	$8.8_{8.4}^{9.2}$	$7.5_{6.9}^{8.0}$	$4.4_{0.0}^{10.1}$	n.a.	n.a.	n.a.	$23.8_{11.3}^{36.0}$	$22.8_{9.6}^{35.4}$	$8.8_{0.0}^{15.6}$
Saxony-Anhalt	$8.1_{7.3}^{9.5}$	$7.0_{6.1}^{8.4}$	$2.1_{0.0}^{5.6}$	$8.6_{7.7}^{9.6}$	$6.8_{5.6}^{9.0}$	$2.9_{0.0}^{8.9}$	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Thuringia	$13.9_{12.7}^{15.3}$	$10.1_{8.2}^{13.8}$	$3.8_{0.0}^{18.6}$	$6.7_{5.6}^{8.1}$	$5.6_{4.3}^{7.3}$	$1.2_{0.0}^{5.7}$	$9.6_{2.5}^{22.2}$	$9.2_{2.1}^{22.1}$	$0.0_{0.0}^{0.0}$	$19.6_{6.5}^{32.8}$	$19.3_{5.8}^{32.7}$	$5.7_{0.1}^{10.2}$

Note: $RMSE$ is the root mean square error (Eq. 21); MAE is the mean absolute error (Eq. 23); $DIST$ is the distance of the observations to the posterior realisation ensemble (Eq. 24); $J_{ps,wwtpOut,all}^N$ is the WWTP outgoing N load; $J_{ps,wwtpOut,all}^P$ is the WWTP outgoing P load; $J_{wwtpIn,all}^N$ is the WWTP incoming N load; $J_{wwtpIn,all}^P$ is the WWTP incoming P load; The numbers indicate the mean, minimum (min) and maximum (max) values: $mean_{min}^{max}$. For $RMSE$ and MAE the mean, maximum and minimum values are calculated over the posterior simulations, while for $DIST$ they are calculated over time.
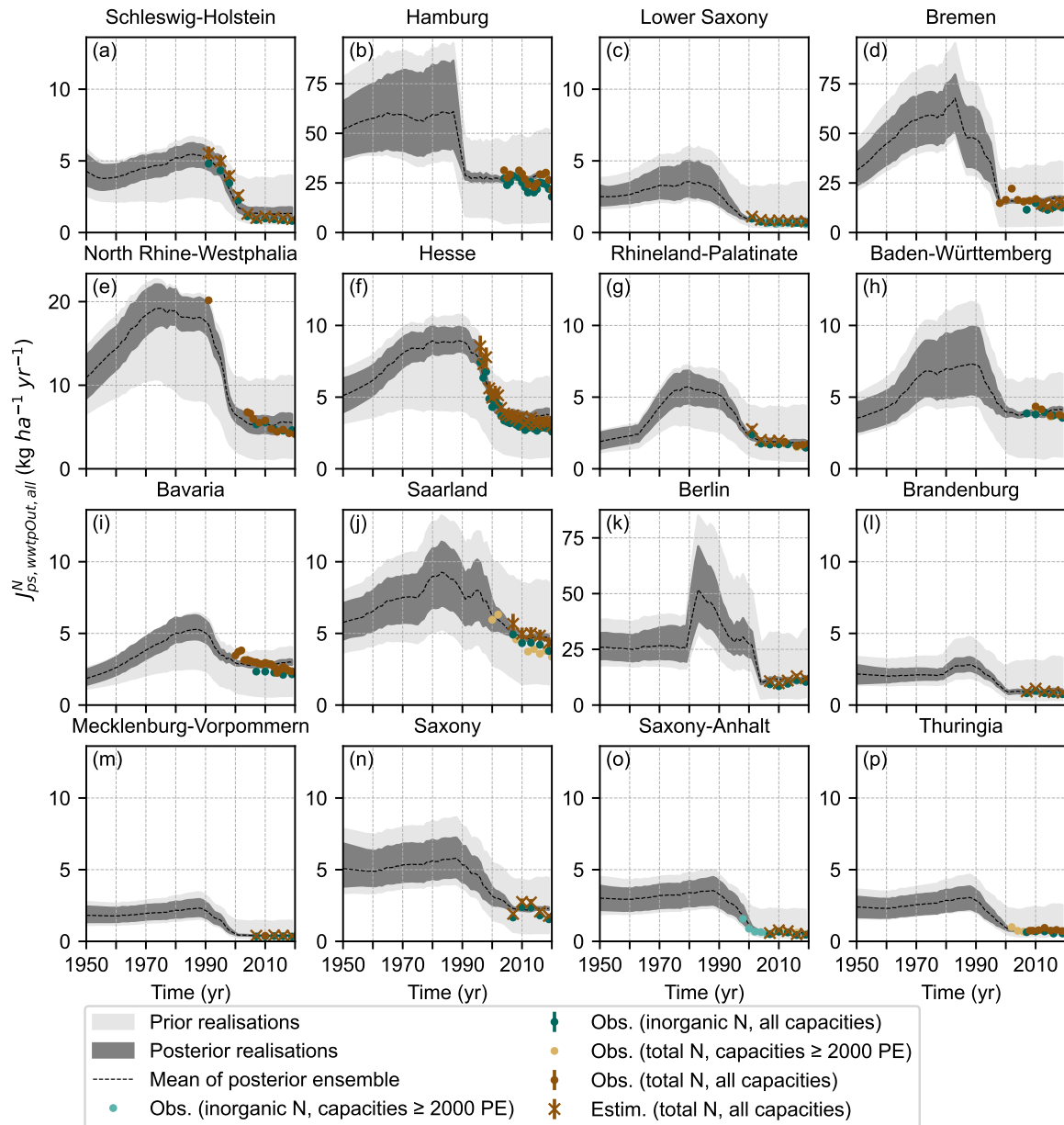
Figure 5. WWTP outgoing N loads $J_{ps,wwtpOut,all}^{N}$ $\left(kg\ ha^{-1}\ yr^{-1}\right)$ from the prior and posterior model realisations (grey shaded areas and dashed lines), observations (coloured circles), and our estimates of total N load from inorganic N load data (dark brown crosses) at NUTS-1 level for the period 1950–2019. The shaded areas delineate the minimum and maximum values of the 100,000 prior realisations (light grey) and the 100 posterior realisations (dark grey). Observational values are shown for both the ensemble of WWTPs having a design capacity higher than 2000 PE (light green and brown) and for all WWTPs (dark green and brown). Some observations are reported with error bars that depict the upper and lower bound estimates while the markers (circles or crosses) represent the average value. The scale on the y-axis is the same for all NUTS-1 regions but the city states (Hamburg, Bremen, and Berlin) and North Rhine-Westphalia for which the loads are much higher compared to the other NUTS-1 regions. Data source: observational data come from the statistical offices of Germany and the federal states and from the authorities of the federal states. The uncertainty intervals on the observations were processed in this study. The estimates of total N represented with brown crosses were derived in this study from inorganic N data. Details on NUTS-1 observational data sources and processing are in Supplementary Sect. S8.
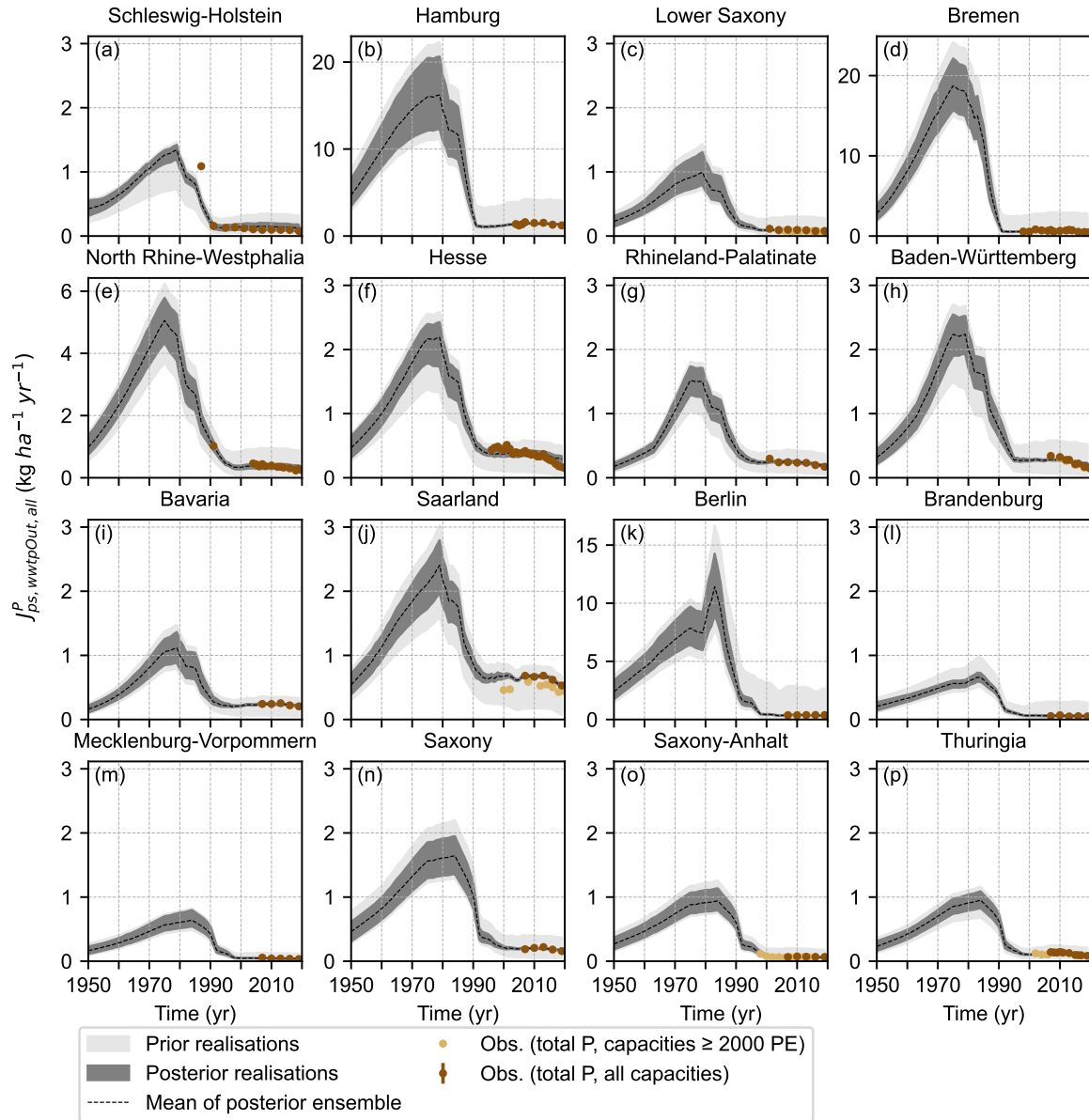
Figure 6. WWTP outgoing P loads $J^P_{ps,wwtpOut,all}$ $(kg\ ha^{-1}\ yr^{-1})$ from the prior and posterior model realisations (grey shaded areas and dashed lines) and observations (coloured circles) at NUTS-1 level for the period 1950–2019. The shaded areas delineate the minimum and maximum values of the 100,000 prior realisations (light grey) and the 100 posterior realisations (dark grey). Observational values are shown for both the ensemble of WWTPs having a design capacity higher than 2000 PE (light brown) and for all WWTPs (dark brown). Some dark brown circles are reported with error bars that depict the upper and lower bound estimates while the circles represent the average value. The scale on the y-axis is the same for all NUTS-1 regions but the city states (Hamburg, Bremen, and Berlin) and North Rhine-Westphalia, for which the loads are much higher compared to the other NUTS-1 regions. Data source: observational data come from the statistical offices of Germany and the federal states and from the authorities of the federal states. The uncertainty intervals on the observations were processed in this study. Details on NUTS-1 observational data sources and processing are in Supplementary Sect. S8.

~~A reason for such~~ These differences in uncertainty ~~bound~~ bounds between the periods with and without observations ~~is that~~ ~~some parameters, that have a limited~~ can be interpreted in terms of parameter sensitivity. Table 5 shows that the parameter estimation procedure allows to appreciably reduce the ranges of a few parameters, namely (1) the efficiencies of tertiary treatment with targeted N and P removal (eff_{3N}^N and eff_{3P}^P), and (2) the efficiencies of secondary and tertiary treatment without targeted N and P removal (eff_{23noN}^N and eff_{23noP}^P) to a lesser extent and for some NUTS-1 regions only. This also reflects the fact that these two types of wastewater treatment tend to be more prevalent during the (recent) period with observations (Fig. 4). We observe that the posterior parameter ranges for eff_{3N}^N and eff_{3P}^P vary across NUTS-1 regions. Regarding eff_{23noN}^N and eff_{23noP}^P, the higher values tend to be discarded in the posterior sample for some NUTS-1 regions. Notably, additional parameters have an impact on the simulated ~~WWTP outgoing~~ N and P ~~load over~~ outgoing load during the period with observations ~~,~~ for at least some of the NUTS-1 regions. This regards mostly parameters that control the magnitude of the human N and P gross emissions (namely f_{waste}^{pro}, f_{pro}^N, and $f_{intake}^{N:P}$), the fractions of industrial/commercial to domestic N and P gross emissions in 2000 (namely $f_{ind:dom,2000}^N$ and ~~therefore cannot be constrained during parameter estimation, may instead~~ $f_{ind:dom,2000}^P$) and the fraction of N and P emissions lost during wastewater collection and transport ($f_{loss,transport}^{N,P}$). Although the ranges of these parameters cannot be reduced in the posterior parameter sample because of parameter interactions, their distributions are constrained (the posterior distributions deviate from the prior distributions in Supplementary Fig. S35). During the period without observations, parameters that could hardly be constrained may have a significant impact on the simulations~~over the period without observations. Examples of such parameters are~~. This for example can be the fractions of industrial/commercial to domestic N and P gross emissions in 1950 (namely ~~$f_{ind:dom,1950}^N$~~ $f_{ind:dom,1950}^N$ and ~~$f_{ind:dom,1950}^P$, respectively) ,~~ $f_{ind:dom,1950}^P$) and the N and P removal efficiencies for primary treatment (eff_1^N and eff_1^P~~, respectively~~). Primary treatment was ~~more prevalent~~ indeed widespread before 1990 ~~while it is almost nonexistent after 2000, as shown in~~ (Fig. 4~~. The constraining of the parameter distributions in the posterior ensemble can be visualised in Supplementary Fig. S35.~~).

At NUTS-0 (country) level, we assess the consistency of the WWTP outgoing N and P load from the mean behavioural realisation (dark orange shaded areas in Fig. 7a,b) with the corresponding data provided by the Germany Environmental Agency for the period 1987–2016 (UBA, 2020) (~~dashed~~ dotted red line in Fig. 7a,b). The UBA data ~~, whose methodology is outlined in different reports for the different years~~ (Behrendt et al., 2000; Fuchs et al., 2010, 2017, 2022) ~~,~~ were constructed from WWTP data for the recent period, while they are based on further modelling assumptions for the earlier period to compensate for the paucity of data~~, as documented for the year 1985 and 1995 in Behrendt et al. (2000)~~. Hence, similar to our data, we expect that the UBA data have larger uncertainties for earlier years compared to the recent ones. The relative differences between our estimates and the UBA data with respect to the UBA data vary over time between -23.6 % and 6.8 % for N load and between -45.5 % to 2.5 % for P load, a negative value indicating that our realisations are lower then the UBA value, and vice-versa. A relatively good match is found for the recent period. Specifically, the error is always within ± 10 % from 2002 onward for N load and from 2006 onward for P load. The largest relative difference is found for the year 1987 for N (-23.6 %; -2 $kg\ ha^{-1}\ yr^{-1}$) and for the year 1994 for P (-45.5 %; -0.32 $kg\ ha^{-1}\ yr^{-1}$). These results confirms the higher uncertainty in WWTP outgoing load estimates before the 2000s compared to the 2000s and 2010s when more data are available to constrain

the calculations. Similar observations can be made when considering all 100 posterior realisations beyond the mean value , as

725 shown in (Supplementary Fig. S38,S39).

Table 5. Lower and upper values of the parameters in the prior sample and in the posterior sample for each NUTS-1 region, and percentage reduction in the parameter ranges (Δ) in the posterior compared to the prior distribution (we report only the four parameters for which Δ is higher than 20 % for at least one NUTS-1 region).

Parameter sample (prior or posterior for each NUTS-1 region)	eff_{23noN}^{N}			eff_{3N}^{N}			eff_{23noP}^{P}			eff_{3P}^{P}		
	Lower	Upper	Δ	Lower	Upper	Δ	Lower	Upper	Δ	Lower	Upper	Δ
	(-)	(-)	(%)	(-)	(-)	(%)	(-)	(-)	(%)	(-)	(-)	(%)
Prior distribution	0.35	0.60	-	0.70	0.95	-	0.45	0.65	-	0.80	0.98	-
Schleswig-Holstein	0.35	0.51	36.4	0.87	0.95	68.2	0.45	0.56	45.8	0.88	0.98	44.7
Hamburg	0.36	0.60	2.3	0.73	0.82	63.1	0.45	0.65	1.7	0.89	0.92	79.2
Lower Saxony	0.35	0.60	1.3	0.88	0.93	83.5	0.46	0.65	3.4	0.90	0.94	81.2
Bremen	0.35	0.60	1.2	0.76	0.84	68.4	0.45	0.65	1.4	0.92	0.95	86.8
North Rhine-Westphalia	0.35	0.43	69.2	0.80	0.88	69.0	0.45	0.65	1.0	0.88	0.93	67.6
Hesse	0.35	0.44	63.5	0.77	0.85	69.3	0.45	0.64	6.0	0.80	0.89	49.5
Rhineland-Palatinate	0.35	0.52	30.5	0.82	0.89	71.6	0.45	0.62	17.0	0.86	0.92	65.0
Baden-Württemberg	0.35	0.60	1.4	0.71	0.81	60.5	0.45	0.55	50.4	0.85	0.91	67.1
Bavaria	0.35	0.47	51.8	0.70	0.80	63.3	0.45	0.65	1.8	0.80	0.88	53.4
Saarland	0.35	0.60	1.7	0.76	0.86	59.3	0.45	0.65	2.2	0.80	0.83	80.9
Berlin	0.35	0.60	2.8	0.83	0.90	73.5	0.45	0.64	6.0	0.96	0.97	91.7
Brandenburg	0.36	0.60	3.5	0.85	0.91	76.6	0.48	0.64	17.7	0.94	0.96	86.4
Mecklenburg-Vorpommern	0.35	0.60	1.8	0.89	0.93	81.5	0.45	0.64	6.3	0.91	0.96	71.7
Saxony	0.36	0.60	3.0	0.75	0.84	65.3	0.45	0.65	4.0	0.83	0.89	65.2
Saxony-Anhalt	0.35	0.60	1.5	0.87	0.92	80.9	0.47	0.65	9.1	0.89	0.93	75.0
Thuringia	0.35	0.60	1.9	0.86	0.93	72.5	0.45	0.58	34.2	0.87	0.94	62.2

The parameters are defined in Table 1. The prior parameter sample is the same for each NUTS-1 region. For a given parameter, the percentage reduction in the range (Δ (%)) is calculated as $\Delta = 100(1 - \frac{Upper_{posterior} - Lower_{posterior}}{Upper_{prior} - Lower_{prior}})$, where $Upper_{posterior}$ and $Lower_{posterior}$ are the upper and lower values respectively in the posterior parameter sample, and $Upper_{priori}$ and $Lower_{prior}$ are the upper and lower values respectively in the prior parameter sample.
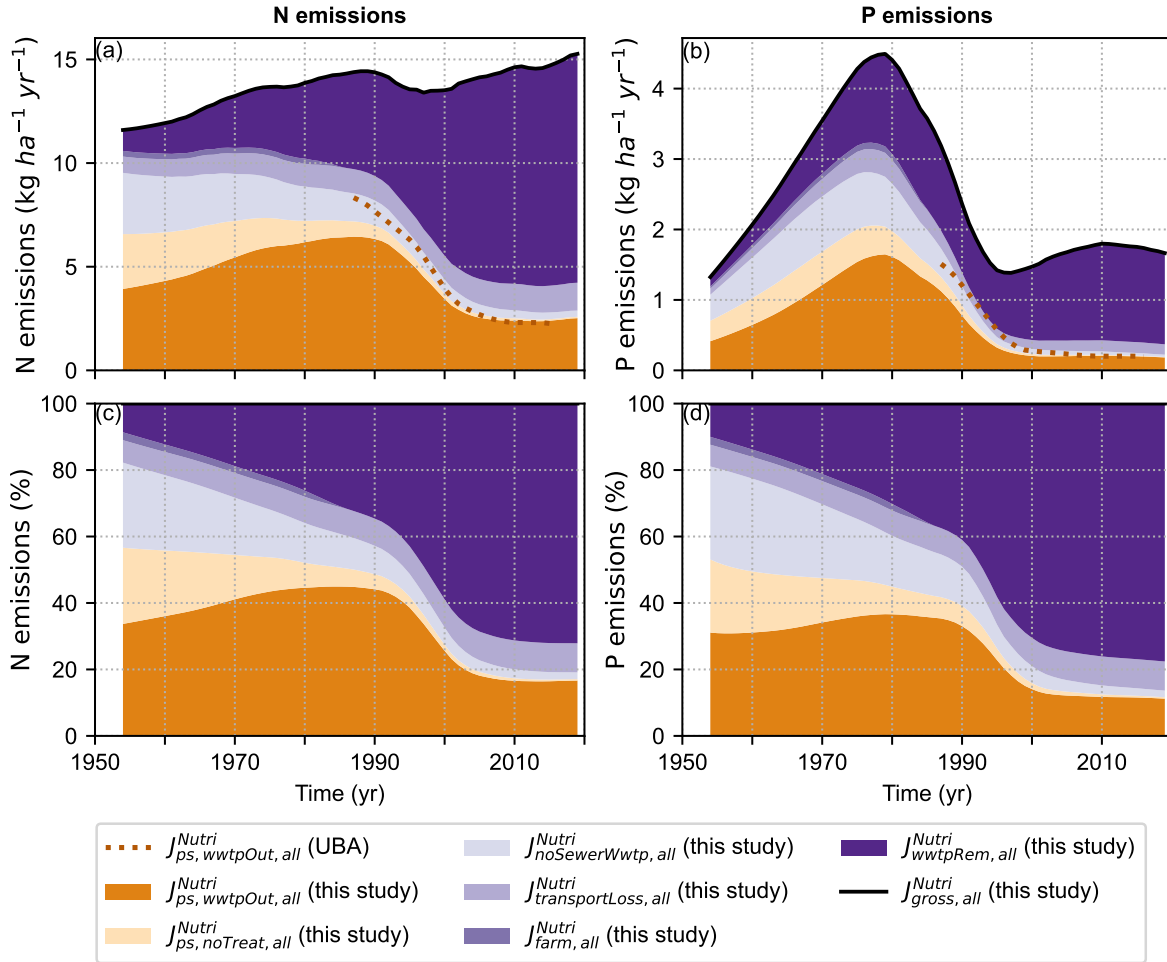
Figure 7. Fate of the N and P gross emissions in Germany at NUTS-0 level for the period 1954–2019 from the mean of our posterior realisations (black line and shaded areas in panels (a,b)), and independent estimates of WWTP outgoing loads provided by the German Environmental Agency (UBA, 2020) for the period 1987–2016 (~~dashed~~ dotted red line). All fluxes are shown are 5-year backward moving averages (the value in 1954 corresponds to the average of the values for 1950–1954). Panels (a,c) represent N emissions and panels (b,d) P emissions. Panels (c,d) report the N and P emissions normalised by the N and P gross emissions, respectively, while panels (a,b) show the absolute emissions. The pathways of the gross emissions are point sources (red colours) or other pathways (purple colours). Notations: $J_{gross,all}^{Nutri}$ are the N ($Nutri = N$) and P ($Nutri = P$) gross emissions; $J_{ps,wwtpOut,all}^{Nutri}$ are the N and P WWTP outgoing emissions (treated point sources); $J_{ps,noTreat,all}^{Nutri}$ are the N and P emissions collected in the public sewer system that are not treated in WWTPs (untreated point sources); $J_{noSewerWwtp,all}^{Nutri}$ are the N and P gross emissions that are not collected in the sewer system nor treated in WWTPs; $J_{transportLoss,all}^{Nutri}$ are the N and P emissions lost during wastewater collection and transport; $J_{farm,all}^{Nutri}$ are the N and P emissions applied to agricultural soils in sewage farms; $J_{wwtpRem,all}^{Nutri}$ are the N and P emissions that are removed during treatment in WWTPs.

Figure 7 presents the fate of the N and P gross emissions at NUTS-0 level from the mean of the posterior realisations. The dynamics in the different N emission pathways, and in particular the N point sources (reddish shaded areas in Fig. 7a, including the treated part - $J^N_{ps,wwtpOut,all}$, and the untreated part - $J^N_{ps,noTreat,all}$), are more controlled by the characteristics of the wastewater handling system that determine the partitioning of the gross emissions between the different pathways (shown in Fig. 7c) than the dynamics of the N gross emissions (black line in Fig. 7a). The later increase moderately by overall 32 % over the period 1950–2019, rising from 11.6 $kg\ ha^{-1}\ yr^{-1}$ on average over the period 1950–1954 to 15.3 $kg\ ha^{-1}\ yr^{-1}$ on average over the period 2015–2019. We can distinguish three temporal phases for the different N emission pathways. The first phase spans the period from the early 1950s to the end of the 1980s, where the point sources (sum of treated and untreated parts) are the dominant fates of the N gross emissions with a slowly decreasing contribution amounting to 56.6 % over the period 1950–1954 and to 48.7 % over the period 1986–1990 (Fig. 7c) and a slightly increasing magnitude from 6.6 $kg\ ha^{-1}\ yr^{-1}$ to 7.0 $kg\ ha^{-1}\ yr^{-1}$ over the same period (Fig. 7a). The importance of the non-treated part reduces strongly from 40.5 % of the total point sources in 1950–1954 to only 9.4 % in 1986–1990. During this phase, the second dominant fate in order of importance after the point sources is the emissions that are not collected in the sewer system nor treated in WWTPs ($J^N_{noSewerWwtp,all}$) at the beginning of the 1950s to then progressively the removal in WWTPs ($J^N_{wwtpRem,all}$). The second phase covers the 1990s and is characterised by a sharp decrease in the N point sources that drop to 2.8 $kg\ ha^{-1}\ yr^{-1}$ (20.1 % of total emissions) in 2000–2004. This is due in particular to an improvement in wastewater treatment, hence the considerable increase in the N removal in WWTP ($J^N_{wwtpRem,all}$) that almost doubles between 1986–1990 and 2000-2004 (from 5.0 to 9.5 $kg\ ha^{-1}\ yr^{-1}$). In the last phase, changes in N emissions are very limited and notably the N point sources amount to 2.6 $kg\ ha^{-1}\ yr^{-1}$ in 2015–2019, accounting for 17.1 % of total emissions, which is close to the 2000-2004 values.

The relative importance of the different P emission pathways (Fig. 7d) follows similar developments as the N emission pathways (Fig. 7c). We note that the contribution of the P point sources, and more specifically the treated component ($J^P_{ps,wwtpOut,all}$), tends to be lower than that of the N point sources, which. This can be explained by the higher efficiency on average of P treatment , as indicated by the ranges of the efficiency parameters reported (see efficiency of primary treatment in Table 1, and of secondary and tertiary treatment in Table 5). Unlike the N emissions, the dynamics of the P emission pathways are more governed by the very strong dynamics of the P gross emissions over the period 1950–2019 (Fig. 7b). P gross emissions more than triple between 1950–1954, with an average value of 1.32 $kg\ ha^{-1}\ yr^{-1}$, and 1974–1979, with an average value of 4.5 $kg\ ha^{-1}\ yr^{-1}$. The P gross emissions then sharply decrease to achieve about their early 1950s value on average over the period 1992–1996 (1.39 $kg\ ha^{-1}\ yr^{-1}$) and slightly increase again over the recent period reaching 1.66 $kg\ ha^{-1}\ yr^{-1}$ over 2014–2019. While N gross emissions depend only on population dynamics (Supplementary Fig. S8) and protein supply per capita (Fig. 3a), P emissions dynamics are driven by the marked development of the P detergents (Fig. 3b-d). P point sources amount to 0.70 $kg\ ha^{-1}\ yr^{-1}$ on average over 1950–1954 (53.1 % of total emissions), then increase to reach their maximum value over 1973–1978 equal to 2.06 $kg\ ha^{-1}\ yr^{-1}$ (46 % of total emissions). They then underwent undergo an almost ten-fold decrease to reach 0.22 $kg\ ha^{-1}\ yr^{-1}$ over 2000–2004 (13.5 % of total emissions), due to both a drop in detergent use and improvement in wastewater treatment. Similar to N point sources, P point sources are relatively constant over the recent period and amount to 0.19 $kg\ ha^{-1}\ yr^{-1}$ on average during 2015–2019 (11.6 % of total emissions). Since the N and P point sources

have different temporal dynamics, their N:P ratio strongly varies in time, namely between 3.5 and 17, and lies around 13.5 on average over the recent period 2015–2019 (see Supplementary Fig. S40).

At NUTS-1 (state) level, the N and P point sources follow in general temporal developments similar to that at NUTS-0 (Germany) level average, although differences are observed in their magnitude, for instance regarding the treated component shown in Fig. 5,6. The point sources for the recent period (average 2015–2019), and more specifically the treated component (as the untreated component is negligible for most states) varies in the range 0.4–29.3 $kg\ ha^{-1}\ yr^{-1}$ for N and 0.04–1.37 $kg\ ha^{-1}\ yr^{-1}$ for P. The highest values above 10 $kg\ ha^{-1}\ yr^{-1}$ for N and 0.5 $kg\ ha^{-1}\ yr^{-1}$ for P are reached in the city ~~(NUTS-1)~~ states of Hamburg, Bremen, and Berlin (Fig. 5-6.b,d,k) that also have the highest population density (Fig. 2). The lowest values below 1 $kg\ ha^{-1}\ yr^{-1}$ for N and 0.1 $kg\ ha^{-1}\ yr^{-1}$ for P are found in the states of Lower Saxony, Mecklenburg-Vorpommern and Saxony-Anhalt (Fig. 5-6c,m,o), that are also among the states with the lowest population density below 2 $people\ ha^{-1}$ for the recent period 2015–2019 (Fig. 2). The treated point sources are also low in Thuringia but the untreated point sources are large compared to the other states, namely 0.95 $kg\ ha^{-1}\ yr^{-1}$ for N and 0.1 $kg\ ha^{-1}\ yr^{-1}$ for P, resulting overall in a moderate amount of point sources (1.8 $kg\ ha^{-1}\ yr^{-1}$ for N and 0.2 $kg\ ha^{-1}\ yr^{-1}$ for P). Not only the magnitude of the pathways of the N and P gross emissions differs between NUTS-1 regions, but also the relative proportions of these pathways. Remarkably, the point sources contributions to the total emissions vary in rather wide ranges between NUTS-1 regions, for instance between 29.7 % and 63.1 % and between 19.0 % and 62.7 % in 1986–1990 for N and P, respectively, and between 9.3 % and 24.5 % and between 3.3 % and 20.8 % in 2015–2019 for N and P, respectively.

We also observe differences between the NUTS-1 regions of (former) West and East Germany (~~excluding Berlin), as~~ Supplementary Fig. S41–S45). Specifically, the contribution (in %) of the emissions that are not collected in the sewer system nor treated in WWTPs ($J^{Nutri}_{noSewerWwtp,all}$) and the untreated point sources ($J^{Nutri}_{ps,noTreat,all}$) are always larger on average over the NUTS-1 regions of East Germany. The contribution of the removal in WWTPs ($J^{Nutri}_{wwtpRem,all}$) is larger on average over the NUTS-1 regions of West Germany . ~~We refer to Supplementary Fig. S41–S61 for further details.~~ In Berlin, we observe different dynamics of the N and P emissions compared to the other ~~states, that are due to the specificity of the state of Berlin, in which~~ state. This is because sewage farms were an important pathway prior to 1982 (Fig. 4k) and ~~for which~~ a time-varying fraction of the population of Berlin is connected to WWTPs located in the neighbouring state of Brandenburg. For instance, we observe a sharp increase in the treated N and P emissions (Fig. 5-6k) in the 1980s, resulting from the discontinuation of the sewage farm operations.

Herein we ~~reported~~ report the emissions based on the mean estimates of the posterior ensemble. Variations in the magnitude of the N and P emissions occur across the different posterior realisations~~, and in particular~~. In particular, for some NUTS-1 regions, the uncertainty bounds of the different emission pathways ~~(corresponding to 100 posterior realisations)~~ show some overlap, which impacts to some degree their relative importance (Fig. S47–S48). Nevertheless, the relative importance of the different N and P emission pathways at NUTS-0 level is rather robust against uncertainty (Fig. S46).

## 5.3 Gridded point sources estimates and their uncertainty due to the spatial disaggregation scheme

Figure 8a presents the gridded map of the total N point sources normalised by the total grid cell area ($J^N_{ps,tot,all}$ ($kg\ ha^{-1}\ yr^{-1}$)) from the mean of the posterior realisations, that are averaged over the decade 1970–1979 and downscaled using population data only (as detailed in Sect. 2.2). ~~In this map, we observe the~~ The spatial distribution of $J^N_{ps,tot,all}$ ~~being~~ is strongly influenced by the distribution of the urban and rural population density ~~that can be visualised in~~ (Supplementary Fig. S5). The densely populated urban areas stand out by their high values of the point sources shown in red ~~to darker red~~ colours. More specifically, we distinguish the city states of Berlin, Bremen, and Hamburg (denoted as BE, HB and HH, respectively), as well as areas in the South to Middle West Germany that are within the European Megalopolis (so-called "Blue Banana" or "Liverpool-Milan Axis"). In contrast, Fig. 8d shows that as much as 42 % of the grid cells have a small value of the point sources lower than 0.5 $kg\ ha^{-1}\ yr^{-1}$ (dark blue colour), due in particular to the very limited generation of N point sources from the rural population that is largely not connected to the sewer system and WWTPs (details in Fig ~~S21~~ S23). An exception is the South to Middle West Germany where many cells display light blue colours, showing a higher contribution of the rural population to the N emissions.

The pattern evolves in time towards a more scattered distribution of the N point sources due to a progressively higher connection of the rural population to the sewer system and WWTPs (Fig. 8b,c,e,f). Notably, the percentage of grid cells with a small value of the point sources lower than 0.5 $kg\ ha^{-1}\ yr^{-1}$ (dark blue color) drops to 18 % in the 1990–1999 decade (Fig. 8e). This percentage is however increasing back after the 1990s, and it reaches 28 % in the 2010–2019 decade. This can be explained by the overall reduction in the N point sources after 1990 due to improvement in wastewater treatment ~~, as described in~~ (Sect. 5.2 and Fig. 7). In this regard, we also observe a reduction in the number of grid cells with large values of $J^N_{ps,tot,all}$ with time. Specifically, the percentage of grid cells with values of the load higher than 5 $kg\ ha^{-1}\ yr^{-1}$ (red colours) decreases from 32 % over 1970–1979 to 22.5 % over 1990–1999 and 10 % over 2010–2019 (Fig. d-f). The temporal changes are also influenced by the local population dynamics (~~reported in~~ Fig. S6). For example the decreases in N point sources between 1990–1999 and 2010–2019 due to improvement in the wastewater handling system are exacerbated by a reduction in population density in some urban areas in the NUTS-1 region of North Rhine-Wesphalia (denoted as NW in Fig. 8a).

The gridded maps of the total N point sources that are downscaled using WWTP data for the treated part (Fig 8g-i) report large differences in the spatial distributions compared to the maps that are downscaled based on population data only (Fig 8a-c). The uncertainty in the spatial pattern at grid level due to the choice of the disaggregation scheme tends to increases with time. This is in particular because the importance of the untreated part of the point sources, which is downscaled using population data, reduces with time (Sect. 5.2, Fig. 7). Using WWTP data for disaggregation results in the concentration of the point sources in as few grid cells where WWTPs are located (Fig 8j-l). These cells can have very high values of the load (dark red colours) up to 53,252, 36,320, and 15,270 $kg\ ha^{-1}\ yr^{-1}$ in 1970–1979, 1990–1999, and 2010–2019, respectively. In contrast, the number of grid cells with values of the load lower than 0.5 $kg\ ha^{-1}\ yr^{-1}$ (dark blue) is large, namely 59 %, 85 % and 92 % of total grid cells in 1970–1979, 1990–1999, and 2010–2019, respectively (Fig 8j-l). Consequently, the densely ~~population~~ populated urban areas are less visible from Fig 8g-i that appear to be more patchy than Fig 8a-c.

Similar observations can be drawn from the gridded maps of the total P point sources ~~that are reported for the two disaggregation schemes in~~ (Supplementary Fig. S62). Nonetheless, the temporal evolution of the P point sources differs from that of the N point sources because they are highly influenced by the dynamics of the P detergents (Sect. 5.2, Fig. 7). In this regard, the gridded maps of the P point sources for the decade 1970–1979 shows much higher values of the load compared to the maps for the following decades.
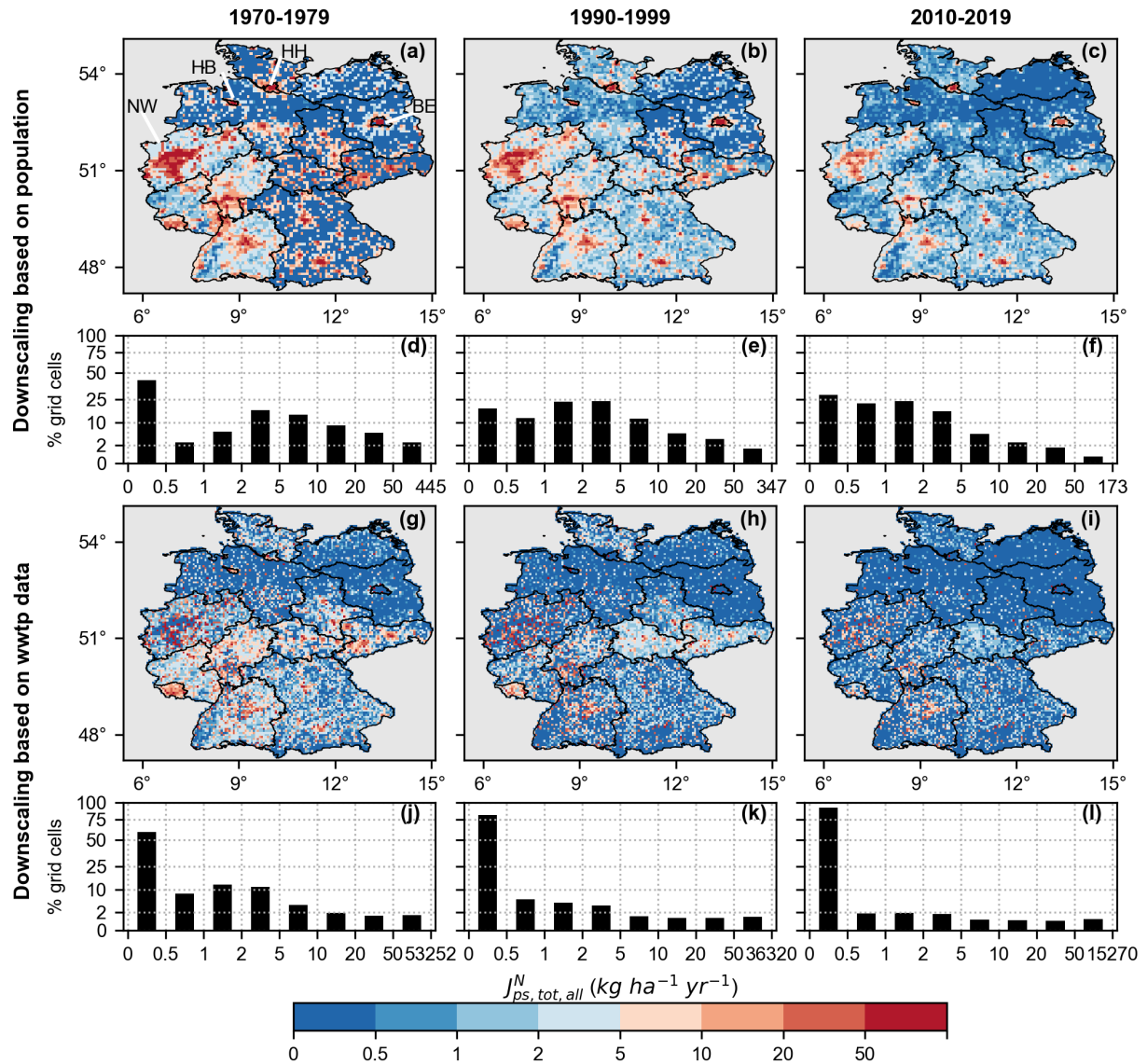
Figure 8. Gridded maps of total N point sources ($J^N_{ps,tot,all}$ in $kg\ ha^{-1}\ yr^{-1}$; panels a–c and g–i), including both the treated and untreated components, and distribution of the grid cell values of $J^N_{ps,tot,all}$ (panels d–f and j–l), obtained by downscaling the mean of the posterior realisations from NUTS-1 to grid level using population data (panels a–f) and WWTP data (panels g-l) for the treated component (panels g-l). The untreated components component is always downscaled based on population data. The figure provides the average values in our posterior ensemble over three different decades, that is, 1970–1979 (left panels), 1990–1999 (middle panels) and 2010–2019 (right panels). Panels (a–c) show the data at a resolution of 0.015625°, while panels (g–i) report aggregated values at a resolution of 0.0625° for depiction purpose (the maps at 0.015625° resolution are reported in Fig S63). Panels (d–f) and (j–l) all refer to the distribution of the grid cells values at 0.015625° resolution. In all panels, the values of $J^N_{ps,tot,all}$ were grouped into the same eight classes. The last class of values (depicted in dark red colour in the maps) includes all values higher than 50 $kg\ ha^{-1}\ yr^{-1}$. The figure also reports the boundaries of the 16 NUTS-1 units of Germany and indicates the location of the city states, that is, Berlin (BE), Bremen (HB), Hamburg (HH), and North Rhine-Westphalia (NW). P point sources are shown in Supplementary Fig. S62. Data source: the NUTS-1 map comes from the German Federal Agency for Cartography and Geodesy (BKG, 2020) ©GeoBasis-DE/BKG dl-de/by-2-0 license (https://www.govdata.de/dl-de/by-2-0). The map is shown in the WGS84 system in the figure.

5.4 Uncertainty in point source estimates at river basin level due to the spatial disaggregation scheme

~~We~~ Given the substantial uncertainties in the spatial pattern of the point sources at high spatial resolution (grid level, Sect 5.3), we examine the uncertainties in the point source estimates at river basin level, which is of interest to water quality studies, as further ~~discussion~~ discussed in Sect. 6.1. We aggregate our gridded point sources estimates over a large sample of basins from the HydroBASINS v1.c of the HydroSHEDS database (HydroSHEDS, 2014; Lehner and Grill, 2013). Basins are selected by intersecting all basins of all levels provided by the HydroSHEDS database over Germany, and removing duplicates. We also discard small basins that have an area less than $10\ km^2$, which is comparable to the original resolution of the HYDE population dataset (5'x5') that we use to disaggregate the data from NUTS-1 to grid level. This results in a total of 3778 basins (Fig. 9a), with an area between 10 and $263{,}580\ km^2$. Half of the basins have an area between 114 and $253\ km^2$ (25 % and 75 % percentile, respectively), while 5.4 % of the basins have an area higher than $1000\ km^2$ (~~the distribution of the basin areas is reported in Supplementary~~ see Fig. S65).

We first analyse the uncertainties in the total N point sources ~~normalised by the basin area~~ ($J_{ps,tot,all}^N$ ($kg\ ha^{-1}\ yr^{-1}$)) for six exemplary nested sub-basins of the Weser River Basin, which is Germany's largest national river basin. The sub-basin sizes vary between 210 and $44{,}949\ km^2$ and their extent is shown by the grey shaded areas in Fig. 9a-b. The temporal variations in the N point sources posterior ensembles are comparable to that observed at NUTS-0 and NUTS-1 level (Sect. 5.2, Fig. 7) for both downscaling ~~scheme for the treated point sources~~ schemes, that is the scheme based on population data (shown in green colours) and WWTP data (red colours) (Fig. 9.d-h). An exception is the smallest sub-basin (Fig. 9c) that has singular temporal dynamics. ~~In particular, the~~ The value of the N point sources ~~obtained using the downscaling approach~~ downscaled based on population data ~~(green colours)~~ is equal to 0 before 1970. This is because the population in the basin is exclusively rural. Before 1970, the basin did not have point source contributions because rural population is not connected to the sewer system nor to WWTPs.

Figure 9c-h also ~~report~~ reports the uncertainty due to the choice of the disaggregation scheme with purple dashed lines that represent the absolute differences between the two means of the posterior ensemble resulting from the two downscaling schemes (dark green and red solid lines), denoted as $\Delta\overline{J_{ps,tot,all}^N}$. For the basins of intermediate sizes (482–16,097 km^2; panels d-g), the absolute differences in the mean tend to first increase with time up to the 1970s-1980s —on average between 1.0 and 1.4 $kg\ ha^{-1}\ yr^{-1}$, and then decline to reach a value between 0.09 and 0.4 $kg\ ha^{-1}\ yr^{-1}$ on average over the decade 2010–2019. For the smallest (panel c) and largest (panel h) basins, $\Delta\overline{J_{ps,tot,all}^N}$ takes small values (lower than 0.3 $kg\ ha^{-1}\ yr^{-1}$). ~~The relative differences in the mean,~~ Notably, the absolute and relative (normalised by the average values between the two means~~and reported in~~, Supplementary Fig. ~~S67, show diverse temporal developments and in particular they increase in time for the basin of size 2178 km^2 (panel e) . Notably, the absolute and relative~~ S71) differences between the means tend to decrease with the basin size for the basins of intermediate to larger sizes (panels e-h).
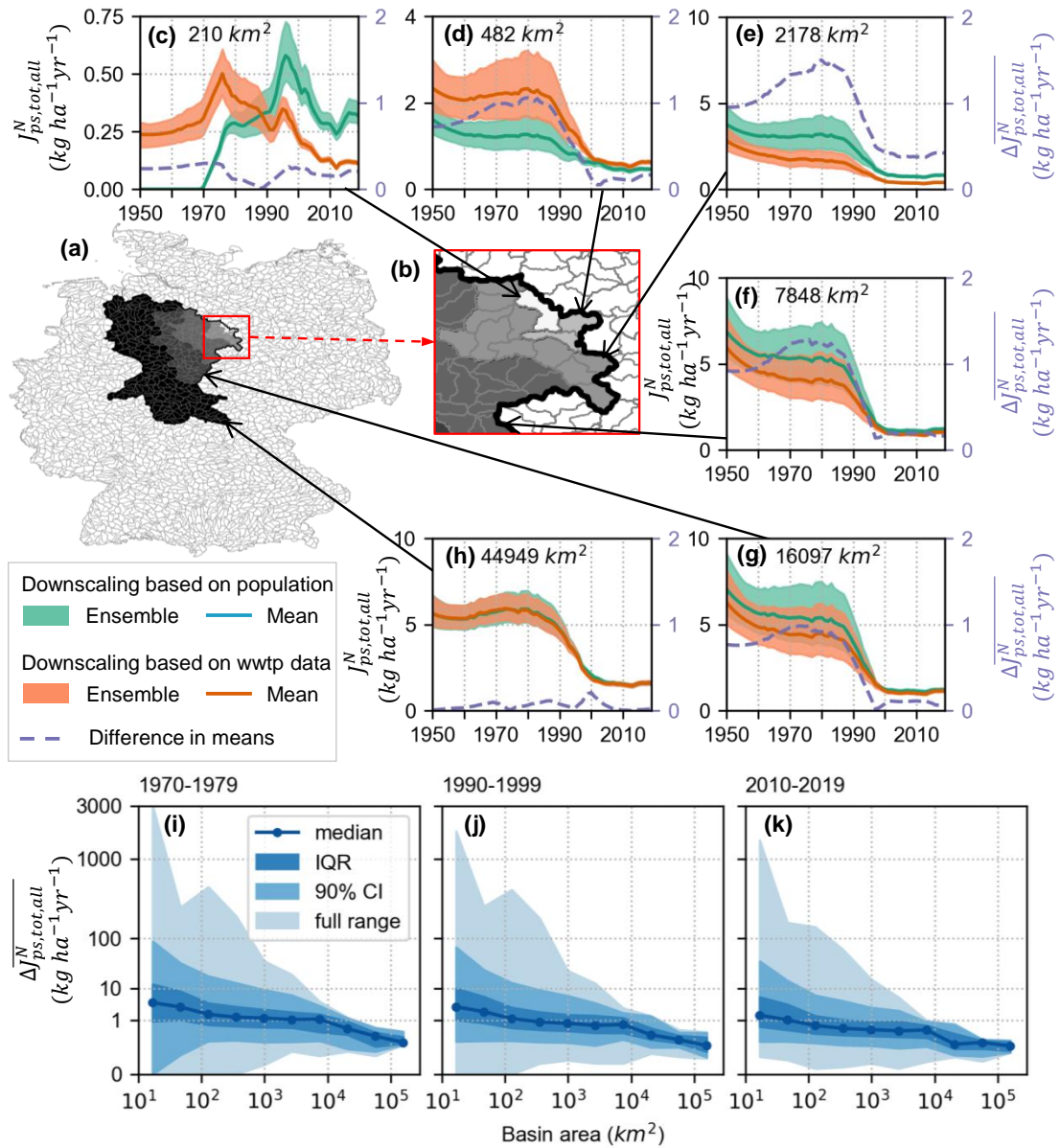
Figure 9. N point sources ($J_{ps,tot,all}^N$ *in kg* ha^{-1} yr^{-1}) and absolute differences in the mean of the posterior ensemble between the two downscaling approaches based on population and WWTP data ($\Delta \overline{J_{ps,tot,all}^N}$ *in kg* ha^{-1} yr^{-1}), aggregated at basins level for river basins of the HydroSHEDS database. Panels (a-b) report the map of the 3778 basins selected over Germany, including six exemplary nested sub-basins of the Weser River Basin reported in grey shaded areas. Panels (c-h) represent the N point sources for the six sub-basins of the Weser River Basin. The light green and red shaded areas delineate the minimum and maximum values over the 100 posterior realisations for the downscaling approach for the treated component based on population and WWTP data, respectively. The solid dark green and red lines represent the mean of the posterior realisations, and the absolute difference between these two means is shown as dashed purple lines (secondary y-axis). Panels (i-k) report the distribution (median, Interquartile range (IQR), 90 % confidence interval (CI) and full range) of the difference in the means ($\Delta \overline{J_{ps,tot,all}^N}$) against the basin area, that is grouped into ten classes. The classes centers are reported as dark blue circles along the median line. Distribution are shown for the 1970–1979 decade (panel i), 1990–1999 decade (panel j), and 2010–2019 decade (panel k). Data source: basin boundaries come from the HydroBASINS v1.c of the HydroSHEDS database (HydroSHEDS, 2014; Lehner and Grill, 2013) ©HydroSHEDS. The map is shown in the WGS84 system in the figure.

Figure 9i-k summarises the results regarding the absolute differences in the two means ~~of the posterior realisations of N point sources, obtained with the two downscaling schemes~~ ($\Delta \overline{J^N_{ps,tot,all}}$) for all 3778 selected HydroSHEDS basins. Results are grouped into ten classes of values according to the basin area, and for three time periods, namely 1970–1979 (panel i), 1990–1999 (panel j), and 2010–2019 ~~in~~ (panel k). The median value (dark blue solid line) depicts the overall trend in the absolute difference and shows that $\Delta \overline{J^N_{ps,tot,all}}$ tends to decrease with the basin area and with time. In particular, in the 1970–1979 decade (panel i), the median is decreasing from 3–4 $kg\ ha^{-1}\ yr^{-1}$ for basin areas lower than 100 km^2 (two left-most dark blue circles in panel i) to around 1.2 $kg\ ha^{-1}\ yr^{-1}$ for basin areas around 1000 km^2 (fifth dark blue circles from the left in panel i) and to 0.07–0.42 for basin areas higher than 10,000 km^2 (three right-most circles in panel i). In the 2010–2019 decade (panel k), the median is equal to 1–1.5, 0.4 and 0.04–0.07 $kg\ ha^{-1}\ yr^{-1}$ for basin areas lower than 100 km^2, around 1000 km^2, and higher than 10,000 km^2, respectively. The median of the relative differences (expressed in %~~), that are shown in~~; Supplementary Fig. S71~~,~~), also tends to decrease with the basin area. ~~Furthermore, relative differences exhibit important changes in time~~ ~~for~~ Relative differences are however more stable with time, apart from basins with an areas lower than 100 km^2 ~~only, for which the relative differences increase in~~ for which it tends to increase with time.

The width of the different envelops depicted in Fig 9i-k (interquartile range, 90 % confidence interval and full range) represent the variability in the absolute differences across basins given the basin size. All envelop widths tend to decrease with time and with the basin size, as the disaggregation of the point sources over smaller basins is influenced by the local specificity, while larger basins capture larger scale tendencies. For smaller basin sizes, the width of the 90 % confidence interval and full range take very high values up to 90 and 3038 $kg\ ha^{-1}\ yr^{-1}$ (left-most area class in panel i). In contrast, the 90 % confidence interval and full range are of limited magnitude for larger catchments and are always lower than 1.5 $kg\ ha^{-1}\ yr^{-1}$ for basin areas higher than 10,000 km^2. ~~The tendency of the variability across basins to decrease with the basin size is also verified~~ A similar tendency is found for the relative differences ~~between the means of the two posterior distributions of the N point sources~~ (Supplementary Fig. S71).

Further details on the absolute and relative differences for all decades in the period 1950–2019 are reported in Supplementary Fig. S66–S67. The uncertainty in the P point source estimates at river basin level due to the spatial disaggregation scheme show similar patterns as the N point sources ~~, as can be seen in~~ (Supplementary Fig. S68, ~~S69, S70 , S72~~ – S70 and S72).

6   Use of the point sources dataset and future perspectives

6.1   Value and potential use of the dataset

Our N and P point sources dataset builds upon a diverse set of information. In particular, we collect Germany specific detergent data from the German Federal Ministry of the Environment, Nature Conservation and Nuclear Safety (BMU) and the German Cosmetic, Toiletry, Perfumery and Detergent Association (IKW), that reflect legislation changes over the period 1975–2019. Instead, previous studies relies on proxy data (GDP) for the estimate of detergent use (Van Drecht et al., 2009; Morée et al., 2013) and on large-scale detergent datasets that have a limited ~~time~~ temporal extent (1985–2000) (Grizzetti et al., 2022; Vigiak et al., 2023). ~~We also~~

We incorporate in our modelling approach observational data of the treated part of the N and P point sources, that is WWTP outgoing load, 1) at NUTS-1 level, that are available over a time period starting between 1987 and 2008 depending on the NUTS-1 regions and ending in general in 2019 and that are used for parameter estimation and model evaluation; and 2) at WWTP level, that are representative of the period around the year 2016 and that are used for downscaling the treated point sources estimates. We ~~also~~ could thus verify the plausibility of our N and P point sources estimates at NUTS-1 level that capture the WWTP outgoing load observations and additional observational data of WWTP incoming load reasonably well (Sect. 5.1). The performance are lower with respect to WWTP incoming P emissions. This could be due not only to inaccuracies in our realisations, but also uncertainties in the observational data, that consist of measurements as well as possibly estimated values when observations are not available as reported in Sect. 3.5.2.

Importantly, our dataset goes beyond these WWTP observational data, by systematically providing long-term (1950–2019) and high resolution (0.015625° grid) point sources including both the part which is treated in WWTP, and the part which is not treated and directly discharged into surface waters from the sewer system. In this regard, the dataset accounts for the large temporal changes in the N and P point sources and their N:P ratio (Sect. 5.2)~~, in particular following~~. Temporal dynamics are strongly influenced by the 1991 Urban Waste Water Treatment Directive (EC, 1991), that led to an improvement in wastewater treatment and a large decrease in N and P point sources in the 1990s, and ~~following~~ the 1980 Phosphate Ordinance (PHochstMengV, 1980) and the subsequent self-commitment of the laundry detergent manufacturers to produce phosphate-free detergents, that resulted in a sharp reduction in detergent phosphate in the 1980s and early 1990s. It is also crucial to consider the point sources that are not treated in WWTP ~~, since~~ and for which no observational data are available. In fact, they contribute substantially to the total point sources before 1990 and over the entire 1950–2019 period for the state of Thuringia.

Such a long-term dataset can support water quality assessment studies~~as~~. This is demonstrated for example in Nguyen et al. (2022) and Sarrazin et al. (2022) that investigated the N legacies in the soil and groundwater for a large sample of river basins in Germany over the period 1950–2014 and 1960–2015, respectively, using a preliminary version of the N point sources dataset presented in this paper (based on ~~country-level~~ NUTS-0 level statistics). In particular, Sarrazin et al. (2022) found that N point sources can have a large impact on the magnitude and trend of the in-stream nitrate ~~loading~~ concentration in the Weser River Basin, since point sources ~~have an immediate impact on the in-stream concentration~~ immediatly reaches surface waters without prior degradation and transport through the soil and groundwater, as occur to the diffuse sources. Our dataset ~~also~~ complements existing datasets not only on ~~N and P point sources~~ point sources for Germany or at a larger-scale (Morée et al., 2013; UBA, 2020; Vigiak et al., 2020, 2023), but also on diffuse sources (Batool et al., 2022; Ebeling et al., 2022; Häußermann et al., 2020), that are also an important source for nutrient pollution. Our database also complement datasets focusing on other aspects of wastewater, such as the global datasets of Ehalt Macedo et al. (2022) on WWTPs location and characteristics, and of Jones et al. (2021) on wastewater volume and its pathways, both being limited to the year 2015.

Another important feature of our dataset is that it quantifies uncertainties due to 1) the parameter values (~~cofficients~~coefficients), including 100 point sources estimates that are most consistent with observations at NUTS-1 level, and 2) the downscaling scheme, providing two gridded point sources maps for each of the 100 realisations using population or WWTP spatial data. This results in a total of 200 gridded point sources realisations that capture the uncertainties that are due to the paucity of data

to constrain the parameters at NUTS-1 level, as well as the lack of spatially resolved data for spatial disaggregation to grid level.

The uncertainties tend to decrease with time. In particular, the parametric uncertainty is lower over the recent period compared to the earlier period where no observations are available to constrain the model parameters (Sect. 5.1). Regarding the downscaling, while the difference between the point sources estimates resulting from the two spatial disaggregation scheme reduces with time in absolute terms, it does not vary substantially in relative terms, apart from smaller aggregation scales (basins with area lower than 100 km^2) where it even increases with time depending on the local specificity (Sect. 5.4). Importantly, the disaggregation scheme based on WWTP data around the year 2016 produces more accurate results over the recent period than the other scheme. In earlier years, both approaches could be considered as being equiprobable, since they are both based on strong assumptions, namely that the WWTP setting did not change with time (scheme based on WWTP data), and that population density can be an adequate proxy for the spatial distribution of WWTP outgoing emissions (~~population data; as~~ scheme based on population data that was also adopted in previous studies such as Morée et al. (2013) and Vigiak et al. (2023)). Therefore, we encourage users of this dataset to devise a time-varying weighting scheme to take into account the varying accuracy in the spatial disaggregation approaches. Such weighting scheme could ensure that only the estimates downscaled using WWTP data are considered for the recent period, while the uncertainties arising from both downscaling schemes are considered for earlier years. We provide an illustration of such a scheme in Supplementary Fig. S73. Moreover, the uncertainties in our estimates decrease from grid level (Fig. 8) towards larger spatial aggregation units (e.g. river basins in Fig. 9). In particular, for earlier years, using the data at larger spatial scales of aggregation (above 100 km^2) is more reliable given the differences between the two downscaling schemes. Therefore, using the dataset directly at grid level (Sect. 5.3) may be prone to large uncertainties and our analysis at river basin level (Sect. 5.4) allows reflection on suitable spatial aggregation scales for water quality studies. In this respect, our gridded dataset provides the flexibility to aggregate the point sources estimates at any spatial level of interest. Our uncertainty characterisation goes beyond the study of Morée et al. (2013) that analysed the uncertainties in the model coefficients using spatially-uniform a priori ranges that were not constrained using observational data as in this study. Nonetheless, Morée et al. (2013) also found a tendency toward higher uncertainties in the past compared to the more recent period.

Overall, the point sources data for the past period have to be interpreted as providing temporal and spatial trends rather than precise estimates. A reason for this is the use of qualitative and proxy information not only to downscale the data to grid level, but also to derive the point sources estimates at NUTS-0 and NUTS-1 level in absence of quantitative data, in particular for the period before the 1990s for East Germany, before the 1980s for West Germany, and for the entire period 1950–2019 for industrial and commercial emissions~~(Table 2 provides an overview of the quantitative data used)~~. A comparison of our data with independent estimates provided by the German Environmental Agency for the period 1987–2016 (UBA, 2020) indeed shows ~~that~~ a good consistency in the temporal trends with increasing similarity between the two datasets with time (Sect. 5.1).

6.2   Directions for future improvement of the dataset

965   The point sources dataset presented in this paper accounts for N and P emissions from wastewater that are collected by the sewer system and/or treated in urban WWTPs, resulting from human excreta, detergent P use, and industrial indirect emissions. The incorporation of additional point sources constituents in future versions of this dataset would provide a more comprehensive database for N and P point source emissions. One of the main constituent to be addressed is the industrial direct emissions. These emissions are difficult to estimate due to a lack of data, and it may only be possible to recover a lower bound for these

970   emissions from the available data (such as the E-PRTR and IED records), as discussed in the introduction (Sect. 1). However, given their importance in magnitude, industrial direct emissions and their uncertainties would be worth further investigation in future studies. In this respect, the NUTS-0 level dataset of the German Environmental Agency (UBA, 2020) indicates that the ratio of industrial direct emissions to WWTP outgoing emissions over the period 1983–1987 is as much as 40 % for N and 12 % for P. This ratio decreases with time to reach 12 % for N and 4 % for P over the period 2012–2016, but, as mentioned earlier,

975   this is likely to be a lower bound.

Urban runoff that is collected in the sewer system (storm sewers or combined sewers) can also lead to large N and P loads being discharged to surface waters, because, on the one hand, it contains N and P ~~,~~ in particular due to atmospheric N and P deposition (Yang and Toor, 2018). On the other hand, when collected in combined sewer systems that also carry wastewater, urban runoff can cause combined sewer overflow (CSO), when the capacity of the effluent handling system is exceeded during

980   storm events (Fuchs et al., 2010; Pistocchi et al., 2019). This results in wastewater being ~~discharges~~ discharged directly to surface waters without treatment in WWTPs. Urban runoff and CSO have been identified as important sources of pollution to be addressed in the revision of the Urban Waste Water Treatment Directive, which is currently in preparation (Vikolainen, 2023). However, the magnitude of storm sewer N and P discharge and CSO are uncertain, since no observational data are currently available, and their estimation involves simplifying assumptions and uncertain coefficients (Fuchs et al., 2010; Pistocchi et al.,

985   2019).

Another potentially important contributor to point sources is the domestic emissions that are not connected to the sewer system nor to WWTPs via transport from cesspits~~, and it~~. While these emissions are overall of limited importance in the recent period, their magnitude is large in the earlier period (Sect. 5.2). It would be therefore critical to elucidate ~~the fate of these emissions~~ their fate in Germany and in other countries where these emissions can be substantial for the recent period

990   as well (Vigiak et al., 2020). These emissions are handled in particular in septic tanks or independent wastewater systems ~~and~~ (Vigiak et al., 2020, 2023). They could be either a diffuse source to soils or a point source to surface ~~water~~waters, as documented for example in MUGV (2010) for the recent period in Germany. ~~Previous studies made different simplifying assumptionsregarding this component, either considering that~~ Due to a lack of detailed information on these emissions, previous studies made simplifying assumptions. In Germany, Fuchs et al. (2010) consider that disconnected population is equipped with

995   septic tanks from which a part of the N and P is transported to WWTPs, while the other part is a diffuse source. In Europe, Grizzetti et al. (2022) and Vigiak et al. (2020) consider that it is entirely a diffuse source ~~(Grizzetti et al., 2022; Vigiak et al., 2020, 2023) or~~ and that N and P are reduced with the same efficiency as that of primary treatment (Vigiak et al., 2020) or possibly secondary

treatment (Grizzetti et al., 2022). Globally, Morée et al. (2013) and Van Puijenbroek et al. (2019) consider that the urine part is a point source~~(Fuchs et al., 2010)~~, while the feces part is a diffuse source. Other point sources constituents can also contribute to the total point sources, such as intensive livestock operations and fish farms. However, to our knowledge, these emissions were not considered in previous point sources assessment in Europe.

Moreover, our dataset can be further improved by explicitly accounting for N emissions embedded in food waste disposal in sewer. This can be considered as in the recent IPCC (2019) report using a coefficient that represents the ratio of protein disposed of via this route to total human N excreta and it is suggested to be 1.09 in Europe. The estimation of this coefficient is however highly uncertain due to a lack of reporting datasets. In Germany, food disposal down the drain could account for overall 14 % of the domestic food waste in Germany (Schmidt et al., 2019). However, this figure cannot be related to the amount of protein wasted, due to a lack of information regarding the specific food commodities involved. Here, we argue that, by considering a large uncertainty range in the fraction of protein supply wasted at the distribution and consumption level (f_{waste}^{pro}, see Table 1), we may implicitly account for the fact that part of the protein wasted may actually contribute to the point sources.

As previous point sources estimates (Morée et al., 2013; Van Drecht et al., 2009; UBA, 2020; Vigiak et al., 2020, 2023), our dataset does not distinguish between the different N and P species that constitute wastewater and that can have a different impact on the environment. For instance, nitrogen can be present as ammonia (NH_3), which is very toxic for aquatic ecosystem, or as ammonium (NH_4^+) or nitrate (NO_3^-), which are less toxic (Camargo and Alonso, 2006). In this respect, future studies could quantify the proportions of these different N species in point sources. For this, further data on population connection to advanced (tertiary) wastewater treatment could be collected, to account for not only treatment with targeted denitrification as in this study, but also treatment with targeted nitrification, which is the transformation from NH_3 to $N0_3^-$. ~~Such~~ Somz data are provided for instance by the Statistical Office of Germany (1975–2007; 2010–2022).

Overall, we propose that future water quality studies could perform sensitivity analysis to better understand the impact of different possible assumptions on the N and P pathways discussed in this section. For Germany, such investigation is facilitated as we provide all data that we produced, along with our model code (see code and data availability section).

7 Code and data availability

The dataset (version 1.1) is available from a Zenodo repository at https://doi.org/10.5281/zenodo.10500535 (Sarrazin et al., 2024). It includes (1) the input and calibration data at NUTS-1 level (CSV files), (2) the input data at grid level (netCDF files), (3) the 200 gridded (0.015625°resolution) N and P point sources estimates, resulting from the 100 posterior parameter sets and the two spatial disaggregation schemes (netCDF files), (4) the 100 NUTS-1 level estimates of N and P gross emissions and their different pathways from the 100 posterior parameter sets (CSV files), and (5) the 200 basin-level N and P point sources estimates for each of the 3778 basins selected from the HydroSHEDS database (CSV files). The code of the N and P point sources model is written in Python and is available from a Github repository (https://github.com/fannysarrazin/NP_point_sources_model, last access: 15 January 2024); it is assigned a DOI (https://doi.org/10.5281/zenodo.10501238; Sarrazin and Kumar, 2024).

8 Conclusions

The main features of N and P point sources dataset introduced in this paper is that it provides for Germany long-term (1950-2019) and high-resolution (0.015625°) estimates of not only the domestic and industrial point sources that are treated in urban WWTPs, but also the untreated domestic point sources that are collected in the sewer system, while considering the uncertainties in modelling choices (i.e. coefficients and downscaling approaches). Our data construction is the results of a substantial data collection from a multitude of different sources to extract state-level or at least country-specific data, wherever possible.

Overall, the magnitude of the N and P point sources strongly varies in time and space and we recommend that they should be careful examined in water quality studies. Despite the uncertainties involved in point sources estimation and the fact that not all point sources constituents are included in this version of the dataset (e.g. industry direct emissions), we believe that this dataset supports robust long-term water quality and nutrient assessment studies. In fact, we provide 200 realisations of the point sources, in an attempt to characterise their uncertainties, and a detailed discussion on how to interpret these estimates and to best use them, given their level of precision that varies in time and across spatial aggregation scales. In addition, we provide a transparent methodology, that builds on existing modelling concepts, to enable long-term reconstruction, while incorporating observational WWTP load data, to increase the accuracy and precision of the estimates. A similar approach could be adopted by other researchers to develop other national and regional datasets ~~, in an effort~~ where sub-national and observational point sources datasets are available. This would contribute to improve our large-scale understanding of nutrient point sources and their impact on the (aquatic) environment.

Author contributions. FS and RK conceptualised the study and designed the methodology. FS led the analyses, developed the code, and collected, processed, and cured the data. RK supervised the work. FS produced the original draft of the paper, including the figures and tables, that was reviewed and edited by all authors.

References

Arle, J., Blondzik, K., Claussen, U., Duffek, A., Grimm, S., Hilliges, F., Kirschbaum, B., Kirst, I., Koch, D., Koschorreck, J., Lepom, P., Leujak, W., Mohaupt, V., Naumann, S., Pirntke, U., Rechenberg, J., Schilling, P., Ullrich, A., Wellmitz, J., Werner, S., and Wolter, R.: Waters in Germany: Status and assessment German environment, German Environmental Agency (UBA), Dessau-Roßlau, Germany, https://www.umweltbundesamt.de/en/publikationen/waters-in-germany, last access: 03 October 2023, 2017.

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

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

Batool, M., Sarrazin, F. J., Attinger, S., Basu, N. B., Van Meter, K., and R., K.: Long-term annual soil nitrogen surplus across Europe (1850–2019), Scientific Data, 9, 612, https://doi.org/10.1038/s41597-022-01693-9, 2022.

Behrendt, H., Huber, P., Kornmilch, M., Opitz, D., Schmoll, O., Scholz, G., Uebe, R., Pagenkopf, W.-G., Bach, M., and Schweikart, U.: Nutrient emissions into River Basins of Germany, Texte|23/2000, German Environmental Agency (UBA), Berlin, Germany, https://www.umweltbundesamt.de/en/publikationen/nutrient-emissions-into-river-basins-of-germany, last access: 07 November 2021, 2000.

Berth, P., Berg, M., and Hachmann, K.: Mehrkomponentensysteme als Waschmittelbuilder, Tenside Detergents, 20, 276–282, https://doi.org/10.1515/tsd-1983-200610, 1983.

BKG (Bundesamt für Kartographie und Geodäsie): NUTS regions 1 : 250 000, 31.12.2020, GeoBasis-DE [data set], Leipzig, Germany, https://gdz.bkg.bund.de/index.php/default/nuts-gebiete-1-250-000-stand-31-12-nuts250-31-12.html, last access: 1 November 2022, 2020.

BMU (Bundesministers für Umwelt, Naturschutz und Reaktorsicherheit): Bericht der Bundesregierung an den Deutschen Bundestag uber Wirkungen des Wasch- und Reinigungsmittelgesetzes vom 19.12.1986, Drucksache 11/4315, Zugeleitet mit Schreiben des Bundesministers für Umwelt, Naturschutz und Reaktorsicherheit vom 5. April 1989 – WA I 3 – 521 140/1 – gemäß Beschluß des Deutschen Bundestages vom 13. November 1986 – Drucksache 10/6404, Dr. Hans Heger, Bonn, Germany, 1989.

Bouraoui, F., Grizzetti, B., and Aloe, A.: Long term nutrient loads entering European seas, Joint Research Centre JRC62873, Publications Office of the European Union, https://doi.org/10.2788/54513, 2011.

Büttner, O.: DE-WWTP - data collection of wastewater treatment plants of Germany (status 2015, metadata), HydroShare [data set], https://doi.org/10.4211/hs.712c1df62aca4ef29688242eeab7940c, 2020.

Byrnes, D. K., Van Meter, K. J., and Basu, N. B.: Long-Term Shifts in U.S. Nitrogen Sources and Sinks Revealed by the New TREND-Nitrogen Data Set (1930–2017), Global Biogeochemical Cycles, 34, e2020GB006 626, https://doi.org/10.1029/2020GB006626, 2020.

Camargo, J. A. and Alonso, A.: Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: A global assessment, Environment International, 32, 831–849, https://doi.org/10.1016/j.envint.2006.05.002, 2006.

Conley, D. J., Paerl, H. W., Howarth, R. W., Boesch, D. F., Seitzinger, S. P., Havens, K. E., Lancelot, C., and Likens, G. E.: Controlling Eutrophication: Nitrogen and Phosphorus, Science, 323, 1014–1015, https://doi.org/10.1126/science.1167755, 2009.

Diaz, R. J. and Rosenberg, R.: Spreading Dead Zones and Consequences for Marine Ecosystems, Science, 321, 926–929, https://doi.org/10.1126/science.1156401, 2008.

Dodds, W. K. and Smith, V. H.: Nitrogen, phosphorus, and eutrophication in streams, Inland Waters, 6, 155–164, https://doi.org/10.5268/IW-6.2.909, 2016.

Duarte, C. M., Agusti, S., Barbier, E., Britten, G. L., Castilla, J. C., Gattuso, J.-P., Fulweiler, R. W., Hughes, T. P., Knowlton, N., Love-lock, C. E., Lotz, H. K., Predragovic, M., Poloczanska, E., Roberts, C., and Worm, B.: Rebuilding marine life, Nature, 580, 39–51, https://doi.org/10.1038/s41586-020-2146-7, 2020.

Ebeling, P., Kumar, R., Lutz, S. R., Nguyen, T., Sarrazin, F., Weber, M., Büttner, O., Attinger, S., and Musolff, A.: QUADICA: water QUAlity, DIscharge and Catchment Attributes for large-sample studies in Germany, Earth System Science Data, 14, 3715–3741, https://doi.org/10.5194/essd-14-3715-2022, 2022.

EC: Council Directive 91/271/EEC of 21 May 1991 concerning urban waste water treatment, Official Journal of the European Communities, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31991L0271, last access: 5 February 2023, 1991.

EC: Regulation (EC) No 166/2006 of the European Parliament and of the Council of 18 January 2006 concerning the establishment of a European Pollutant Release and Transfer Register and amending Council Directives 91/689/EEC and 96/61/EC, Official Journal of the European Union, https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1578566025656&uri=CELEX:32006R0166, last access: 20 july 2024, 2006.

EC: Directive 2010/75/EU of the European Parliament and of the Council of 24 November 2010 on industrial emissions (integrated pol-lution prevention and control), Official Journal of the European Union, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32010L0075, last access: 2 October 2023, 2010.

EC: Regulation (EU) No 259/2012 of the European Parliament and of the Council of 14 March 2012 amending Regulation (EC) No 648/2004 as regards the use of phosphates and other phosphorus compounds in consumer laundry detergents and consumer automatic dishwasher detergents, Offical Journal of the European Union, https://eur-lex.europa.eu/eli/reg/2012/259/oj, last access: 11 May 2023, 2012.

EEA: Source apportionment of nitrogen and phosphorus inputs into the aquatic environment, European Environment Agency (EEA) report No 7/2005, Luxembourg, https://www.eea.europa.eu/ds_resolveuid/328c70e765d5978d6efec90d980c0a23, last access: 4 February 2023, 2005.

EEA: Industrial waste water treatment –pressures on Europe's environment, European Environment Agency (EEA) report No 23/2018, Pub-lications Office of the European Union, Luxembourg, https://www.eea.europa.eu/ds_resolveuid/9316d500ad82400bab4ad5383b118396, last access: 12 November 2023, 2019.

EEA: European Pollutant Release and Transfer Register (E-PRTR), European Environmental Agency (EEA) [data set], https://www.eea.europa.eu/ds_resolveuid/3578652f4e8e43bba4f0555a4b5933d0, last access: 20 july 2024, 2020.

EEA: Industrial Reporting under the Industrial Emissions Directive 2010/75/EU and European Pollutant Release and Transfer Register Reg-ulation (EC) No 166/2006, European Environmental Agency (EEA) [data set], https://www.eea.europa.eu/en/datahub/datahubitem-view/9405f714-8015-4b5b-a63c-280b82861b3d, last access: 5 February 2023, 2023a.

EEA: Waterbase - UWWTD: Urban Waste Water Treatment Directive – reported data, European Environmental Agency (EEA) [data set], https://www.eea.europa.eu/en/datahub/datahubitem-view/6244937d-1c2c-47f5-bdf1-33ca01ff1715, last access: 20 july 2024, 2023b.

EEA: European Environment Agency (EEA) Glossary, https://www.eea.europa.eu/help/glossary/eea-glossary, last access: 8 November 2023, 2023c.

EEA, Andersen, J., Harvey, T., and Reker, J.: Nutrient enrichment and eutrophication in Europe's seas: Moving towards a healthy ma-rine environment, European Environment Agency (EEA) report No 14/2019, Publications Office of the European Union, Luxembourg, https://doi.org/10.2800/092643, 2019.

Ehalt Macedo, H., Lehner, B., Nicell, J., Grill, G., Li, J., Limtong, A., and Shakya, R.: Distribution and characteristics of wastewater treatment plants within the global river network, Earth System Science Data, 14, 559–577, https://doi.org/10.5194/essd-14-559-2022, 2022.

Ehrhardt, S., Kumar, R., Fleckenstein, J. H., Attinger, S., and Musolff, A.: Trajectories of nitrate input and output in three nested catchments along a land use gradient, Hydrology and Earth System Sciences, 23, 3503–3524, https://doi.org/10.5194/hess-23-3503-2019, 2019.

Eurostat: Population connected to wastewater collection and treatment systems by NUTS 2 regions (env_wwcon_r2), 11.04.2016 update, European Commission [data set], https://db.nomics.world/Eurostat/env_wwcon_r2, last access: 22 June 2023, 2016.

1145   Eurostat: Population connected to wastewater treatment plants (env_ww_con), 23.03.2023 update, European Commission [data set], https://ec.europa.eu/eurostat/databrowser/view/ENV_WW_CON/default/table?lang=en, last access: 22 June 2023, 2023.

FAO: The state of food and agriculture 1951 - a survey of the world conditions and prospects, FAO, Washington, DC, USA, https://www.fao.org/3/ap639e/ap639e.pdf, last access: 13 November 2023, 1951.

FAO: Food Balances (-2013, old methodology and population) - 21.12.2021 update, FAO [data set], Rome, Italy, http://www.fao.org/faostat/

1150   en/#data/FBSH, last access: 16 May 2023, 2021.

FAO: Food Balances (2010-) - 14.02.2022 update, FAO [data set], Rome, Italy, http://www.fao.org/faostat/en/#data/FBS, last access: 16 May 2023, 2022.

FAO and SIK: Global food losses and food waste - Extent, causes and prevention, Study conducted for the International Congress SAVE FOOD! at Interpack2011 Düsseldorf, Germany, FAO, Rome, Italy, https://www.fao.org/3/i2697e/i2697e.pdf, last access: 26 October 2023,

1155   2011.

Floyd, P., Zarogiannis, P., and Fox, K.: Non-surfactant organic ingredients and zeolite-based detergents – Final report prepare for the European Commission, Risk & Policy Analysts Limited (RPA), Swindon, UK, https://fdocuments.in/document/europa-enterprise-chemicals-non-surfactant-organic-non-surfactant-organic.html, last access: 07 December 2021, 2006.

Foroutan-Rad, M.: Phosphatsubstitute für Wasch- und Reinigungsmittel - okologische und toxikologische Betrachtung, Umwelt, 3, 218–224,

1160   1981.

Fuchs, S., Scherer, U., Wander, R., Behrendt, H., Venohr, M., Opitz, D., Hillenbrand, T., Marscheider-Weidemann, F., and Götz, T.: Calculation of Emissions into Rivers in Germany using the MONERIS Model, Texte|46/2010, German Environmental Agency (UBA), Berlin, Germany, https://www.umweltbundesamt.de/en/publikationen/calculation-of-emissions-into-rivers-in-germany, last access: 26 January 2023, 2010.

1165   Fuchs, S., Weber, T., Wander, R., Toshovski, S., Kittlaus, S., Reid, L., Bach, M., Klement, L., Hillenbrand, T., and Tettenborn, F.: Effizienz von Maßnahmen zur Reduktion von Stoffeinträgen - Neue Ansätze und aktualisierte Ergebnisse von MoRE-DE, Texte|05/2017, German Environmental Agency (UBA), Berlin, Germany, https://www.umweltbundesamt.de/publikationen/effizienz-von-massnahmen-zur-reduktion-von, last access: 26 January 2023, 2017.

Fuchs, S., Brecht, K., Gebel, M., Bürger, S., Uhlig, M., and Halbfaß, S.: Phosphoreinträge in die Gewässer bundesweit modellieren - Neue

1170   Ansätze und aktualisierte Ergebnisse von MoRE-DE, Texte | 142/2022, German Environmental Agency (UBA), Berlin, Germany, https://www.umweltbundesamt.de/publikationen/phosphoreintraege-in-die-gewaesser-bundesweit, last access: 26 January 2023, 2022.

Glennie, E. B., Littlejohn, C., Gendebien, A., Hayes, A., Palfrey, R., Sivil, D., and Wright, K.: Phosphates and alternative detergent builders – Final report prepared for the EU Environment Directorate, UC 4011, Water Research Centre Limited, Swindon, UK, https://www.oieau.fr/eaudoc/system/files/documents/36/183877/183877_doc.pdf, last access: 20 July 2024, 2002.

1175   Grimvall, A., Stålnacke, P., and Tonderski, A.: Time scales of nutrient losses from land to sea — A European perspective, Ecological Engineering, 14, 363–371, https://doi.org/10.1016/S0925-8574(99)00061-0, 2000.

Grizzetti, B., Vigiak, O., Udias, A., Bisselink, B., Pistocchi, A., Bouraoui, F., Malagó, A., Aloe, A., Zanni, M., Weiss, F., Hristov, J., Wilson, J., Pisoni, E., De Meij, A., De Roo, A., Macias, M., and Stips, A.: A European assessment of freshwater availability and nutrient pollution :

historical analysis and scenarios developed in the project Blue2.2, Joint Research Centre JRC130025, Publications Office of the European Union, Luxembourg, https://doi.org/10.2760/924432, 2022.

Groß, R., Leisewitz, A., and Moch, K.: Untersuchung der Einsatzmengen von schwer abbaubaren organischen Inhaltsstoffen in Wasch- und Reinigungsmitteln im Vergleich zum Einsatz dieser Stoffe in anderen Branchen im Hinblick auf den Nutzen einer Substitution, prepared for the German Environment Agency, Tech. Rep. UBA-FB 3709 65 430, öko-Institut and öko-Recherche, https://www.bmuv.de/fileadmin/Daten_BMU/Pools/Forschungsdatenbank/fkz_3709_65_430_wasch_und_reinigungsmittel_bf.pdf, last access 13 November 2023, 2012.

Gustafsson, B. G., Schenk, F., Blenckner, T., Eilola, K., Meier, H. E. M., Müller-Karulis, B., Neumann, T., Ruoho-Airola, T., Savchu, O. P., and Zorita, E.: Reconstructing the Development of Baltic Sea Eutrophication 1850–2006, AMBIO volume, 14, 534—548, https://doi.org/10.1007/s13280-012-0318-x, 2012.

Happel, O., Armbruster, D., Brauch, H.-J., Rott, E., and Minke, R.: Phosphonate in Wasch- und Reinigungsmitteln und deren Verbleib in der Umwelt – Entwicklung von Analyseverfahren und deren praktische Anwendung bei Proben von Oberflächenwasser, Abwasser und Sediment, Texte 69/2021, German Environment Agency, Berlin, Germany, https://www.umweltbundesamt.de/publikationen/phosphonate-in-wasch-reinigungsmitteln-deren, last access 13 November 2023, 2021.

Häußermann, U., Klement, L., and Breuer, L.: Nitrogen soil surface budgets for districts in Germany 1995 to 2017, Environmental Sciences Europe, 32, 109, https://doi.org/10.1186/s12302-020-00382-x, 2020.

Herbert, D. and Fourqurean, J.: Ecosystem Structure and Function Still Altered Two Decades After Short-Term Fertilization of a Seagrass Meadow, Ecosystems, 11, 688–700, https://doi.org/10.1007/s10021-008-9151-2, 2008.

HydroSHEDS: HydroBASINS v1.c, HydroSHEDs [data set], https://www.hydrosheds.org/products/hydrobasins, last access: 23 October 2023, 2014.

IKW (Industrieverband Körperpflege- und Waschmittel e.V.): Nachhaltigkeitbericht 2005/2006, https://www.ikw.org/fileadmin/IKW_Dateien/downloads/Haushaltspflege/HP_Nachhaltigkeitsbr-2006-i.pdf, last access: 13 November 2023, 2005.

IKW (Industrieverband Körperpflege- und Waschmittel e.V.): Nachhaltigkeitbericht 2009/2010, https://www.ikw.org/fileadmin/IKW_Dateien/downloads/Haushaltspflege/HP_Nachhaltigkeitsbericht09-10.pdf, last access: 13 November 2023, 2011.

IKW (Industrieverband Körperpflege- und Waschmittel e.V.): Nachhaltigkeitbericht 2013/2014, https://www.ikw.org/fileadmin/IKW_Dateien/downloads/Haushaltspflege/HP_Nachhaltigkeitsbericht2013-2014.pdf, last access: 13 November 2023, 2015.

IKW (Industrieverband Körperpflege- und Waschmittel e.V.): Nachhaltigkeitbericht 2015/2016, https://www.ikw.org/fileadmin/IKW_Dateien/downloads/Haushaltspflege/HP_Nachhaltigkeitsbericht__15_16.pdf, last access: 13 November 2023, 2017.

IKW (Industrieverband Körperpflege- und Waschmittel e.V.): Nachhaltigkeitbericht 2019, https://www.ikw.org/fileadmin/IKW_Dateien/downloads/Haushaltspflege/2019_Nachhaltigkeitsbericht_interaktiv.pdf, last access: 13 November 2023, 2019.

IPCC: Chapter 6 - Wastewater Treatment and Discharge, in: 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Volume 5 Waste, edited by Buendia, E., Tanabe, K., Kranjc, A., Baasansuren, J., Fukuda, M., S., N., Osako, A., Pyrozhenko, Y., Shermanau, P., and Federici, S., pp. 6.1–6.72, IPCC, Switzerland, https://www.ipcc-nggip.iges.or.jp/public/2019rf/vol5.html, last access: 25 June 2023, 2019.

Jager, A. D. and Vogt, J.: Rivers and Catchments of Europe - Catchment Characterisation Model (CCM), European Commission, Joint Research Centre [data set], http://data.europa.eu/89h/fe1878e8-7541-4c66-8453-afdae7469221, last access: 02 October 2023, 2007.

Jaworska, J., Van Genderen-Takken, H., Hanstveit, A., van de Plassche, E., and Feijtel, T.: Environmental risk assessment of phosphonates, used in domestic laundry and cleaning agents in the Netherlands, Chemosphere, 47, 655–665, https://doi.org/10.1016/S0045-6535(01)00328-9, 2002.

Jenkinson, D. S.: The Rothamsted long-term experiments: Are they still of use?, Agronomy Journal, 83, 2–10, https://doi.org/10.2134/agronj1991.00021962008300010008x, 1991.

Jones, E. R., van Vliet, M. T. H., Qadir, M., and Bierkens, M. F. P.: Country-level and gridded estimates of wastewater production, collection, treatment and reuse, Earth System Science Data, 13, 237–254, https://doi.org/10.5194/essd-13-237-2021, 2021.

Jones, H. P., Jones, P. C., Barbier, E. B., Blackburn, R. C., Rey Benayas, J. M., Holl, K. D., McCrackin, M., Meli, P., Montoya, D., and Mateos, D. M.: Restoration and repair of Earth's damaged ecosystems, Proceedings of the Royal Society B: Biological Sciences, 285, 20172 577, https://doi.org/10.1098/rspb.2017.2577, 2018.

Kemp, W. M., Testa, J. M., Conley, D. J., Gilbert, D., and Hagy, J. D.: Temporal responses of coastal hypoxia to nutrient loading and physical controls, Biogeosciences, 6, 2985–3008, https://doi.org/10.5194/bg-6-2985-2009, 2009.

Klein Goldewijk, K., Beusen, A., Doelman, J., and Stehfest, E.: Anthropogenic land use estimates for the Holocene – HYDE 3.2, Earth System Science Data, 9, 927–953, https://doi.org/10.5194/essd-9-927-2017, 2017.

Klein Goldewijk, K., Beusen, A., Doelman, J., and Stehfest, E.: History Database of the Global Environment 3.2, Data publication platform of Utrecht University [data set], Utrecht, the Netherlands, https://doi.org/10.24416/UU01-MO2FF3, 2022.

Kloepfer, M. and Kröger, H.: Das Umweltrecht in der deutschen Einigung - Zum Umweltrecht im Einigungsvertrag und zum Umweltrahmengesetz, Schriften zum Umweltrecht (SUR), Band 22, Duncker & Humblot, Berlin, Germany, https://www.duncker-humblot.de/buch/das-umweltrecht-in-der-deutschen-einigung-9783428472390/?page_id=1, last access: 06 January 2022, 1991.

Kuliński, K., Rehder, G., Asmala, E., Bartosova, A., Carstensen, J., Gustafsson, B., Hall, P. O. J., Humborg, C., Jilbert, T., Jürgens, K., Meier, H. E. M., Müller-Karulis, B., Naumann, M., Olesen, J. E., Savchuk, O., Schramm, A., Slomp, C. P., Sofiev, M., Sobek, A., Szymczycha, B., and Undeman, E.: Biogeochemical functioning of the Baltic Sea, Earth System Dynamics, 13, 633–685, https://doi.org/10.5194/esd-13-633-2022, 2022.

Le Moal, M., Gascuel-Odoux, C., Ménesguen, A., Souchon, Y., Étrillard, C., Levain, A., Moatar, F., Pannard, A., Souchu, P., Lefebvre, A., and Pinay, G.: Eutrophication: A new wine in an old bottle?, Science of The Total Environment, 651, 1–11, https://doi.org/10.1016/j.scitotenv.2018.09.139, 2019.

Lee, M., Shevliakova, E., Malyshev, S., Milly, P. C. D., and Jaffé, P. R.: Climate variability and extremes, interacting with nitrogen storage, amplify eutrophication risk, Geophysical Research Letters, 43, 7520–7528, https://doi.org/10.1002/2016GL069254, 2016.

Lehner, B. and Grill, G.: Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems, Hydrological Processes, 27, 2171–2186, https://doi.org/10.1002/hyp.9740, 2013.

LEinfG: Ländereinführungsgesetz vom 22. Juli 1990 (GBl. DDR 1990 I S. 955), https://www.gesetze-im-internet.de/leinfg/DDNR009550990.html, last access: 10 August 2023, 1990.

Lemley, D. A. and Adams, J. B.: Eutrophication, in: Encyclopedia of Ecology (Second Edition), edited by Fath, B., pp. 86–90, Elsevier, Oxford, UK, second edition edn., https://doi.org/10.1016/B978-0-12-409548-9.10957-1, 2019.

Lin, L., Clair, S. S., Gamble, G. D., Crowther, C. A., Dixon, L., Bloomfield, F. H., and Harding, J. E.: Nitrate contamination in drinking water and adverse reproductive and birth outcomes: a systematic review and meta-analysis, Scientific Reports, 13, 563, https://doi.org/10.1038/s41598-022-27345-x, 2023.

Lottermoser, B. G.: Effect of long-term irrigation with sewage effluent on the metal content of soils, Berlin, Germany, Environmental Geochemistry and Health, 34, 67–76, https://doi.org/10.1007/s10653-011-9391-5, 2012.

LU-RP and MKUEM-RP (Landesamt für Umwelt, und Ministerium für Klimaschutz, Umwelt, Energie und Mobilität Rheinland-Pfalz): Stand der Abwasserbeseitigung in Rheinland-Pfalz - Lagebericht 2020, Ministerium für Klimaschutz, Umwelt, Energie und

1255   Mobilität Rheinland-Pfalz, Mainz, https://wasser.rlp-umwelt.de/servlet/is/1120/Lagebericht_2020.pdf?command=downloadContent&filename=Lagebericht_2020.pdf, last access: 30 June 2023, 2021.

Macias Moy, D., Stips, A., Grizzetti, B., Aloe, A., Bisselink, B., De Meij, A., De Roo, A., Dutiel, O., Ferreira, N., Garcia Gorriz, E., Gonzalez-Fernandez, D., Hristov, J., Miladinova-Marinova, S., Pärn, O., Piroddi, C., Pisoni, E., Pistocchi, A., Polimene, L., Serpetti, N., Thoma, C., Udias Moinelo, A., Vigiak, O., Weiss, F., Wilson, J., and Zanni, M.: Water/marine Zero Pollution Outlook, Joint Research
1260   Centre JRC 131197, Publications Office of the European Union, Luxembourg, https://doi.org/10.2760/681817, 2022.

McCrackin, M. L., Jones, H. P., Jones, P. C., and Moreno-Mateos, D.: Recovery of lakes and coastal marine ecosystems from eutrophication: A global meta-analysis, Limnology and Oceanography, 62, 507–518, https://doi.org/10.1002/lno.10441, 2017.

Mehlhart, G., Bulach, W., Moch, K., and Blepp, M.: Relevanz der gewerblichen Textil- und Geschirrreinigung am Eintrag von Phosphat und anderen Phosphorverbindungen (P) in das Abwasser, 98/2021, Umwelbundesamt, Dessau-Roßlau, Germany, https://www.
1265   umweltbundesamt.de/publikationen/relevanz-der-gewerblichen-textil-geschirrreinigung, last access: 06 January 2022, 2021.

Millenium Ecosystem Assessment: Ecosystems and human well-being: Synthesis, Island Press, Washington, DC., http://www.millenniumassessment.org/, last access: 8 November 2023, 2005.

Mittelstet, A. and Storm, D.: Quantifying Legacy Phosphorus Using a Mass Balance Approach and Uncertainty Analysis, JAWRA Journal of the American Water Resources Association, 52, 1297–1310, https://doi.org/10.1111/1752-1688.12453, 2016.

1270   Morée, A. L., Beusen, A. H. W., Bouwman, A. F., and Willems, W. J.: Exploring global nitrogen and phosphorus flows in urban wastes during the twentieth century, Global Biogeochemical Cycles, 27, 836–846, https://doi.org/10.1002/gbc.20072, 2013.

Moreno-Mateos, D., Barbier, E. B., Jones, P. C., Jones, H. P., Aronson, J., López-López, J. A., McCrackin, M. L., Meli, P., Montoya, D., and Rey Benayas, J. M.: Anthropogenic ecosystem disturbance and the recovery debt, Nature Communications, 8, 14 163, https://doi.org/10.1038/ncomms14163, 2017.

1275   Moreno-Mateos, D., Alberdi, A., Morriën, E., van der Putten, W. H., Rodríguez-Uña, A., and Montoya, D.: The long-term restoration of ecosystem complexity, Nature Ecology and Evolution, 4, 676–685, https://doi.org/10.1038/s41559-020-1154-1, 2020.

MUGV (Ministerium für Umwelt, Gesundheit und Verbraucherschutz Brandenburg): Wegweiser für den Einsatz von Kleinkläranlagen und Sammelgruben, Potsdam, https://mluk.brandenburg.de/sixcms/media.php/9/Wegweiser-Kleinklaeranlage-Sammelgruben.pdf, last access: 19 May 2024, 2010.

1280   MUNLV-NW (Ministerium für Umwelt und Naturschutz, Landwirtschaft und Verbraucherschutz des Landes Nordrhein-Westfalen): Entwicklung und Stand der Abwasserbeseitigung in Nordrhein-Westfalen. 19 Auflage, Düsseldorf, https://www.lanuv.nrw.de/fileadmin/lanuv/wasser/abwasser/lagebericht/00_EStAb2020_Gesamtversion.pdf, last access: 29 June 2023, 2020.

Nguyen, T. V., Sarrazin, F. J., Ebeling, P., Musolff, A., Fleckenstein, J. H., and Kumar, R.: Toward Understanding of Long-Term Nitrogen Transport and Retention Dynamics Across German Catchments, Geophysical Research Letters, 49, e2022GL100 278,
1285   https://doi.org/10.1029/2022GL100278, 2022.

Noleppa, S. and Cartsburg, M.: Das grosse Wegschmeissen- Vom Acker bis zum Verbraucher: Ausmaß und Umwelteffekte der Lebensmittelverschwendung in Deutschland, WWF Deutschland, https://www.wwf.de/fileadmin/fm-wwf/Publikationen-PDF/WWF_Studie_Das_grosse_Wegschmeissen.pdf, last access: 26 October 2023, 2015.

Nork, M. E.: Umweltschutz in unternehmerischen Entscheidungen - Eine theoretische und empirische Analyse, Deutscher Universitäts-
1290   Verlag, Wiesbaden, Germany, https://doi.org/10.1007/978-3-322-86359-1, 1992.

O'Connell, D. W., Ansems, N., Kukkadapu, R. K., Jaisi, D., Orihel, D. M., Cade-Menun, B. J., Hu, Y., Wiklund, J., Hall, R. I., Chessell, H., Behrends, T., and Van Cappellen, P.: Changes in Sedimentary Phosphorus Burial Following Artificial Eutrophication

of Lake 227, Experimental Lakes Area, Ontario, Canada, Journal of Geophysical Research: Biogeosciences, 125, e2020JG005713, https://doi.org/10.1029/2020JG005713, 2020.

1295 OECD: Diffuse Pollution, Degraded Waters: Emerging Policy Solutions, OECD Studies on Water, OECD Publishing, Paris, https://doi.org/10.1787/9789264269064-en, 2017.

Parkhurst, T., Standish, R. J., and Prober, S. M.: P is for persistence: Soil phosphorus remains elevated for more than a decade after old field restoration, Ecological Applications, 32, e2547, https://doi.org/10.1002/eap.2547, 2022.

Pavinato, P. S., Cherubin, M. R., Soltangheisi, A., Rocha, G. C., Chadwick, D. R., and Jones, D. L.: Revealing soil legacy phosphorus to

1300 promote sustainable agriculture in Brazil, Scientific Reports, 10, 15615, https://doi.org/10.1038/s41598-020-72302-1, 2020.

Penuelas, J., Poulter, B., Sardans, J., Ciais, P., van der Velde, M., Bopp, L., Boucher, O., Godderis, Y., Hinsinger, P., Llusia, J., Nardin, E., Vicca, S., Obersteiner, M., and Janssens, I. A.: Human-induced nitrogen–phosphorus imbalances alter natural and managed ecosystems across the globe, Nature Communications, 4, 2934, https://doi.org/10.1038/ncomms3934, 2013.

Penuelas, J., Janssens, I. A., Ciais, P., Obersteiner, M., and Sardans, J.: Anthropogenic global shifts in biospheric N and P concentrations and

1305 ratios and their impacts on biodiversity, ecosystem productivity, food security, and human health, Global Change Biology, 26, 1962–1985, https://doi.org/10.1111/gcb.14981, 2020.

PHochstMengV: Phosphathöchstmengenverordnung vom 4. Juni 1980 (BGBl. I S. 664), http://www.gesetze-im-internet.de/ph_chstmengv/, last access: 07 December 2021, 1980.

Pianosi, F., Sarrazin, F., and Wagener, T.: A Matlab toolbox for Global Sensitivity Analysis, Environmental Modelling & Software, 70,

1310 80–85, https://doi.org/10.1016/j.envsoft.2015.04.009, 2015.

Pistocchi, A., Dorati, C., Grizzetti, B., Moinelo, A. U., Vigiak, O., and Zanni, M.: Water quality in Europe: effects of the Urban Wastewater Treatment Directive, Joint Research Centre JRC115607, Publications Office of the European Union, Luxembourg, https://doi.org/10.2760/303163, 2019.

Puckett, L. J., Tesoriero, A. J., and Dubrovsky, N. M.: Nitrogen contamination of surficial aquifers — A growing legacy, Environmental

1315 Science & Technology, 45, 839–844, https://doi.org/10.1021/es1038358, 2011.

Rudolph, K.-U. and Block, T.: The German water section - policies and experiences, Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BUNR) and German Environmental Agency (UBA), https://www.umweltbundesamt.de/sites/default/files/medien/publikation/long/2752.pdf, last access: 06 February 2023, 2001.

Sarrazin, F. J. and Kumar, R.: fannysarrazin/NP_point_sources_model: v1.0 (v1.0), Zenodo [code],

1320 https://doi.org/10.5281/zenodo.10501238, 2024.

Sarrazin, F. J., Kumar, R., Basu, N. B., Musolff, A., Weber, M., Van Meter, K. J., and Attinger, S.: Characterizing Catchment-Scale Nitrogen Legacies and Constraining Their Uncertainties, Water Resources Research, 58, e2021WR031587, https://doi.org/10.1029/2021WR031587, 2022.

Sarrazin, F. J., Attinger, S., and Kumar, R.: Gridded dataset of nitrogen and phosphorus point sources from wastewater in Germany (1950-

1325 2019) (v1.1), Zenodo [data set], https://doi.org/10.5281/zenodo.10500535, 2024.

Schmidt, T., Schneider, F., and Claupein, E.: Food waste in private households in Germany – Analysis of findings of a representative survey conducted by GfK SE in 2016/2017 (Thünen Working Paper 92a), https://www.thuenen.de/media/publikationen/thuenen-workingpaper/ThuenenWorkingPaper_92a.pdf, last access: 17 December 2021, 2019.

Sebilo, M., Mayer, B., Nicolardot, B., Pinay, G., and Mariotti, A.: Long-term fate of nitrate fertilizer in agricultural soils, Proceedings of the

1330 National Academy of Sciences, 110, 18185–18189, https://doi.org/10.1073/pnas.1305372110, 2013.

Seeger, H.: The history of German waste water treatment, European Water Management, 2, 1999.

Selman, M., Sugg, Z., and Greenhalgh, S.: Eutrophication and Hypoxia in Coastal Areas A Global Assessment of the State of Knowledge, WRI Policy Note, Water Quality: Eutrophication and Hypoxia no 1, World Resources Institute, Washington, DC, https://www.wri.org/research/eutrophication-and-hypoxia-coastal-areas, last access: 24 January 2023, 2008.

Sharpley, A., Jarvie, H. P., Buda, A., May, L., Spears, B., and Kleinman, P.: Phosphorus Legacy: Overcoming the Effects of Past Management Practices to Mitigate Future Water Quality Impairment, Journal of Environmental Quality, 42, 1308–1326, https://doi.org/10.2134/jeq2013.03.0098, 2013.

Smith, V. H.: Eutrophication of freshwater and coastal marine ecosystems a global problem, Environmental Science and Pollution Research, 10, 126–139, https://doi.org/10.1065/espr2002.12.142, 2003.

SO-BE (Statistisches Landesamt Berlin): Statistisches Jahrbuch 2001, Kulturbuch-Verlag, Berlin, Germany, https://www.statistischebibliothek.de/mir/receive/BBAusgabe_mods_00000271, last access: 18 May 2023, 2001.

SO-DE (Statistisches Bundesamt): 12411-0010- Population: Länder, reference date, GENESIS–Database of the Federal Statistical Office of Germany [data set], Wiesbaden, Germany, https://www-genesis.destatis.de/, last access: 16 May 2023, 2022a.

SO-DE (Statistisches Bundesamt): 12411-0015-Population: Administrative districts, reference date, GENESIS–Database of the Federal Statistical Office of Germany [data set], Wiesbaden, Germany, https://www-genesis.destatis.de/, last access: 16 May 2023, 2022b.

SO-DE (Statistisches Bundesamt): Fachserie : 15, Wirtschaftsrechnungen - Einkommens- und Verbrauchsstichprobe. Heft 1, Ausstattung privater Haushalte mit ausgewählten Gebrauchsgütern und Versicherungen, https://www.statistischebibliothek.de/mir/receive/DESerie_mods_00000154, last access: 16 May 2023, Fachserie 15 Heft 1.

SO-EDE (Staatlichen Zentralverwaltung für Statistik DDR): Statistisches Jahrbuch der Deutschen Demokratischen Republik, Staatsverl. d. Deutschen Demokratischen Republik, Berlin, Germany, https://www.statistischebibliothek.de/mir/receive/DESerie_mods_00007446, last access: 10 August 2023, 1955–1990.

SO-WDE (Statistisches Bundesamt): Statistisches Jahrbuch für die Bundesrepublik Deutschland 1970, W. Kohlhammer GmbH, Stuttgart and Mainz, Germany, https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00130491, last access: 10 August 2023, 1970.

Statistical Office of Germany (Statistisches Bundesamt): Fachserie. 19, Umwelt. Reihe 2, Wasserversorgung und Abwasserbeseitigung. 1, öffentliche Wasserversorgung und Abwasserbeseitigung 1975 (Fachserie / 19 / 2 / 1), Wiesbaden, https://www.statistischebibliothek.de/mir/receive/DESerie_mods_00000203, last access: 03 Octobre 2023, 1975–2007.

Statistical Office of Germany (Statistisches Bundesamt): Fachserie. 19, Umwelt. 2, Wasserversorgung und Abwasserbeseitigung. 1, öffentliche Wasserversorgung und öffentliche Abwasserentsorgung. 2, öffentliche Abwasserbehandlung und -entsorgung 2019 (Fachserie / 19 / 2 / 1 / 2), Wiesbaden, https://www.statistischebibliothek.de/mir/receive/DESerie_mods_00000929, last access: 03 Octobre 2023, 2010–2022.

Stuart, M., Chilton, P., Kinniburgh, D., and Cooper, D.: Screening for long-term trends in groundwater nitrate monitoring data, Quarterly Journal of Engineering Geology and Hydrogeology, 40, 361—376, https://doi.org/10.1144/1470-9236/07-040, 2007.

UBA: Einträge von Nähr- und Schadstoffen in die Oberflächengewässer, German Environment Agency (UBA) [data set], https://www.umweltbundesamt.de/bild/stickstoff-phosphoreintraege-aus-punktquellen, last access: 26 January 2023, 2020.

Van Drecht, G., Bouwman, A. F., Harrison, J., and Knoop, J. M.: Global nitrogen and phosphate in urban wastewater for the period 1970 to 2050, Global Biogeochemical Cycles, 23, GB0A03, https://doi.org/10.1029/2009GB003458, 2009.

Van Meter, K. J., Basu, N. B., Veenstra, J. J., and Burras, C. L.: The nitrogen legacy: Emerging evidence of nitrogen accumulation in anthropogenic landscapes, Environmental Research Letters, 11, 035 014, https://doi.org/10.1088/1748-9326/11/3/035014, 2016.

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

Van Meter, K. J., McLeod, M. M., Liu, J., Tenkouano, G. T., Hall, R. I., Van Cappellen, P., and Basu, N. B.: Beyond the Mass Balance: Watershed Phosphorus Legacies and the Evolution of the Current Water Quality Policy Challenge, Water Resources Research, 57, e2020WR029 316, https://doi.org/10.1029/2020WR029316, 2021.

Van Puijenbroek, P., Beusen, A., and Bouwman, A.: Global nitrogen and phosphorus in urban waste water based on the Shared Socio-economic pathways, Journal of Environmental Management, 231, 446–456, https://doi.org/10.1016/j.jenvman.2018.10.048, 2019.

Vero, S., Basu, N., Van Meter, K., Richards, K. G., Mellander, P.-E., Healy, M. G., and Fenton, O.: Review: the environmental status and implications of the nitrate time lag in Europe and North America, Hydrogeology Journal, 26, 7–22, https://doi.org/10.1007/s10040-017-1650-9, 2018.

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

Vigiak, O., Aloe, B. G. A., Zanni, M., Bouraoui, F., and Pistocchi, A.: Domestic waste emissions to European freshwaters 1990 - 2016 (v. 1.0), European Commission, Joint Research Centre [data set], http://data.europa.eu/89h/104da345-4ba1-444b-8df4-97bb57950a8e, last access: 02 October 2023, 2023.

Vikolainen, V.: Initial Appraisal of a European Commission Impact Assessment - Revising the Urban Wastewater Treatment Directive, PE 740.240, EPRS | European Parliamentary Research Service, Luxembourg, https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/740240/EPRS_BRI(2023)740240_EN.pdf, last access: 03 October 2023, 2023.

Wang, L., Stuart, M. E., Bloomfield, J. P., Butcher, A. S., Gooddy, D. C., McKenzie, A. A., Lewis, M. A., and Williams, A. T.: Prediction of the arrival of peak nitrate concentrations at the water table at the regional scale in Great Britain, Hydrological Processes, 26, 226–239, https://doi.org/10.1002/hyp.8164, 2012.

WHO: Nitrate and nitrite in drinking-water - Background document for development of WHO Guidelines for Drinking-water Quality, WHO/FWC/WSH/16.52, World Health Organization (WHO) Press, Geneva , Switzerland, 2016.

Yang, S., Büttner, O., Jawitz, J. W., Kumar, R., Rao, P. S. C., and Borchardt, D.: Spatial Organization of Human Population and Wastewater Treatment Plants in Urbanized River Basins, Water Resources Research, 55, 6138–6152, https://doi.org/10.1029/2018WR024614, 2019a.

Yang, S., Büttner, O., Kumar, R., Jäger, C., Jawitz, J. W., Rao, P., and Borchardt, D.: Spatial patterns of water quality impairments from point source nutrient loads in Germany's largest national River Basin (Weser River), Science of The Total Environment, 697, 134 145, https://doi.org/10.1016/j.scitotenv.2019.134145, 2019b.

Yang, Y.-Y. and Toor, G. S.: Stormwater runoff driven phosphorus transport in an urban residential catchment: Implications for protecting water quality in urban watersheds, Scientific Reports, 8, 11 681, https://doi.org/10.1038/s41598-018-29857-x, 2018.

ZEODET: Zeolites for detergents as nature intended, European Chemical Industry Council (Cefic), Bruxelles, Belgium, https://www.euzepa.eu/images/3.ZEODETbrochure.pdf, last access: 06 January 2022, 2000.

Zhang, L., Chen, J., and Chu, G.: Legacy phosphorus in calcareous soil under 33 years of P fertilizer application: Implications for efficient P management in agriculture, Soil Use and Management, 38, 1380–1393, https://doi.org/10.1111/sum.12792, 2022.