Interreg
Danube Region

Co-funded by
the European Union

Tethys

Output 2.1

Testing and demonstration of HS emissions model as an operative fit-for-purpose tool for transnational modelling-based risk and scenarios assessment under new challenges and pressures

Project co-funded by European Union funds Danube Region Program and National Funds of the participating countries

O2.1: Testing and demonstration of HS emissions model as an operative fit-for-purpose tool for transnational modelling-based risk and scenarios assessment under new challenges and pressures

PROJECT TITLE: Coordinated Danube Action for the titanic endeavor of tackling hazardous substances water pollution under changing pressures, challenges and targets

ACRONYM: Tethys

DATE OF PREPARATION: 31.12.2025

RESPONSIBLE FOR THE OUTPUT:
Budapest University of Technology and Economics, Hungary (BME)

AUTHORS

Name co-author	Project partner
Zsolt Jolankai	Budapest University of Technology and Economics, HU
Katalin Maria Dudas	Budapest University of Technology and Economics, HU
Máté K. Kardos	Budapest University of Technology and Economics, HU
Vivien Potó	Budapest University of Technology and Economics, HU
Emese Szandányi	Budapest University of Technology and Economics, HU
Tímea Lajkó	Budapest University of Technology and Economics, HU
Adrienne Clement	Budapest University of Technology and Economics, HU

CONTRIBUTING PARTNERS

Name co-author	Project partner
Steffen Kittlaus	TU Wien, AT
Meiqi Liu	TU Wien, AT
Ottavia Zoboli	TU Wien, AT
Matthias Zessner	TU Wien, AT
Dimitar Mihalkov	Bulgarian Water Association, BG
Radoslav Tonev	Bulgarian Water Association, BG
Silviya Petkova	Bulgarian Water Association, BG
Marko Nikolic	Center for Eco-Toxicological Research Podgorica, ME
Anja Babic	Center for Eco-Toxicological Research Podgorica, ME
Danijela Sukovic	Center for Eco-Toxicological Research Podgorica, ME
Đorđa Medić	Croatian Waters “Hvratska Vode”, HR
Jasmina Antolić	Croatian Waters “Hvratska Vode”, HR
Darko Barbalić	Croatian Waters “Hvratska Vode”, HR
Luka Vukmanić	Croatian Waters “Hvratska Vode”, HR
Marianne Bertine Broer	Environment Agency Austria, AT
Oliver Gabriel	Environment Agency Austria, AT
Adam Kovacs	ICPDR, AT
Zoran Major	ICPDR, AT
Prvoslav Marjanović	Jaroslav Černi Water Institute, RS
Marko Marjanović	Jaroslav Černi Water Institute, RS

Name co-author	Project partner
David Mitrinović	Jaroslav Černi Water Institute, RS
David Kocman	Jozef Stefan Institute, SI
Radmila Milačič Ščančar	Jozef Stefan Institute, SI
Thomas Rosmann	National Administration "Romanian Waters", RO
Ioana Nedelea	National Administration "Romanian Waters", RO
Mihai Enciu	National Administration "Romanian Waters", RO
Alexandru Bandea	National Administration "Romanian Waters", RO
Andreea Dăescu	National Administration "Romanian Waters", RO
Olha Ukhan	Ukrainian Hydrometeorological Institute, UA
Yuliia Luzovitska	Ukrainian Hydrometeorological Institute, UA
Denis Klebanov	Ukrainian Hydrometeorological Institute, UA
Dajana Kučić Grgić	University of Zagreb, HR
Matija Cvetnić	University of Zagreb, HR
Michal Kirchner	Water Research Institute, SK
Michal Kunštek	Water Research Institute, SK
Miroslav Kandera	Water Research Institute, SK
Jelena Vićanović	Vode Srpske, BA

ABSTRACT

The development of the MoRE model for the Danube River Basin (DRB) aims to improve the understanding and management of heavy metal, pharmaceutical, and PFAS emissions in accordance with European legislative frameworks, in particular the Water Framework Directive (WFD). Building on the HS emission model developed within the DTP Danube Hazard m3c project, the model was jointly further developed, upgraded with new data and algorithms, and validated to support a broader range and higher complexity of emission scenarios. This represents a key step in establishing a fit-for-purpose tool capable of addressing emerging challenges and advanced risk assessment requirements.

Output 2.1 presents the upgraded calculations and the results of the model development.

Chapter 2 describes the enhanced calculation pathways, including erosion, surface runoff, tile drainage, urban emissions, and industrial sources. Both direct emissions (e.g. wastewater discharges) and indirect emissions (e.g. atmospheric deposition and waste storage) are considered. These pathways capture substance fate and transport processes across relevant environmental compartments, including hotspots, while integrating improved hydrological data to enhance spatial resolution and support climate change scenario analysis.

All project partners actively contributed by providing input and validation data. **Chapter 3** summarises the datasets used, including hydrological and land use data, urban population and sewerage infrastructure information, wastewater treatment technologies, road networks, and the location of industrial sites, mining areas, and potential hotspots.

The upgraded model introduces new parameters and methodologies to quantify unknown emissions along each pathway. **Chapter 4** presents the resulting variants and outcomes, highlighting the importance of accounting for both known and potential emission sources, as many identified sources (e.g. industrial facilities) may release pollutants in the absence of reported emission data.

Chapter 5 outlines four additional upgrades and extensions: (1) R-based model mirroring to increase flexibility and enable uncertainty analysis of selected parameters; (2) enhanced retention calculations incorporating photodegradation, biotransformation, and hydrolysis; (3) development of a stochastic framework to address uncertainty arising from multiple sources within the modelling process; and (4) the mathematical and technical implementation of scenario analyses.

Model validation, described in **Chapter 6**, was designed to make use of existing databases and to ensure robust performance under different conditions.

Note: Model results are presented in Output 2.3 "Fully operative transnational HS emissions model as a fit-for-purpose tool for risk assessment and evaluation of scenarios for policy support under new complex challenges and pressures."

CONTENT

FIGURES

TABLES

1 AIMS AND SCOPE OF HAZARDOUS SUBSTANCES EMISSION MODELING IN THE DANUBE RIVER BASIN

1.1 EMISSION MODELLING IN THE DANUBE RIVER BASIN

Emission modelling in the Danube River Basin (DRB) has a long history, guided by the European legislative frameworks (WFD). The first Danube scale modelling study was delivered in the Danubs project[1] investigating nutrient emission patterns in the DRB, followed by a series of modelling studies coordinated by the International Commission for the Protection of the Danube River (ICPDR). For the latter, modelling has been delivered by the MONERIS model, a lumped parameter static, empirical emission model, prepared by the IGB Berlin[2].

A comparative modelling study has been prepared by a joint effort of ICPDR, JRC and IGB Berlin in 2015[3], in which three models were compared (MONERIS, SWAT, GREEN) to see similarities and differences between the modelling approaches. The study showed that *"The model results in general agreed well with each other and with the observed flow data. Similarly, the models showed good agreement in estimating total nitrogen loads. Differences among models estimates were more noticeable in Morava, Sava, Tisa, Jiu, Olt, Arges-Vedea, Buzau-Ialomita, SiretPrut-Buzau"* .

Hazardous substance emissions in this scale was first modelled in the Danube Hazard m^{3}c project with a water quality model named the Danube River Hazardous Substance Model. This model is based on a river water quality model developed by Deltares in the late 90'ies and developed further in the following years[4] . This model is a source-based model, focusing on sources and riverine processes and less on actual pathways. The model has been applied for the Danube in a spatial resolution determined by the E-Hype hydrologic model. The model results showed a good agreement for some chemicals and poor for others. The main difference between these chemicals was the knowledge about their sources and the relative contribution of certain pathways that were less known at the time of the model application.

Based on the above-described modelling studies, the current study focuses on the implementation of a more defined database on sources and pathways of selected elements (6 heavy metals and As) and compounds (PFOA, PFOS, Diclophenac and Carbamazepine). In case of heavy metals, the main emission sources are clearly the various industries, or which European wide databases are available, with a growing number of reported data, including emissions, discharges, concentrations. A throurough analysis of the emission data has been prepared with the aim to create more accurate emissions for each industry segment. In case of pharmaceuticals, a refinement of municipal point source emissions are kept in the focus based on the analysis of emission data, while the retention of these chemicals is also

[1] Kroiss, H., Zessner, M., & Lampert, C. (2006). daNUbs: Lessons learned for nutrient management in the Danube Basin and its relation to Black Sea euthrophication. *Chemistry and Ecology*, *22*(5), 347–357. https://doi.org/10.1080/02757540600917518

[2] Gericke, A. (IGB-B., & Venohr, M. (IGB-B. (2021). Nutrient Emissions and Loads in the Danube River Basin - Current situation and scenarios for the 3rd Danube River Basin Management Plan – Final report.

[3] Malago, A. (JRC), Venohr, M. (IGB), Gericke, A. (IGB-B., Vigiak, O. (JRC), Bouraoui, F. (JRC), Grizetti, B. (JRC), & Kovacs, A. (ICPDR). (2015). Modelling nutrient pollution in the Danube River Basin: a comparative study of SWAT, MONERIS and GREEN models. https://dx.doi.org/10.2788/156278

[4] van Gils, J. (2015). The Danube Water Quality Model and Its Application in the Danube River Basin BT - The Danube River Basin (I. Liska (ed.); pp. 61–83). Springer Berlin Heidelberg. https://doi.org/10.1007/698_2015_335

integrated in the model based on advanced knowledge about their fate and behaviour in the aquatic environment[5].

For this study, the MoRE emission model[6] is used, which is a modelling framework, able to implement empiric models, such as the MONERIS model[7], which provides the methodological base of the HS emission model currently used. The model was developed and tested in several instances, the developer Institute, Karlsruhe Institute of Technology applied it for the country scale nutrient emission analysis of Germany[8], PFAS compounds emissions were successfully regionalized in Austria[9], and in the Upper Danube Basin[10]. The model was applied for seven pilot catchments in the DRB within the framework of the Danube Hazard m^{3}c project[11].

1.2 LEGAL BACKGROUND

1.2.1 Water Framework Directive (WFD)

The MoRE model has been further developed and implemented in the Tethys project in a way that enables to meet the requirements of the Water Framework Directive (WFD 60/2000/EC) and technical guidance document 28 (TGD 28)[12] in a harmonized form for the whole Danube River Basin.

The **TGD28 aims to help Member States establish the emission inventories** and to reduce the burden by focusing on substances that are relevant at the River Basin District level. TGD28 summarizes well – in its introduction - why it is necessary to create an emissions inventory:

"According to Article 5 of the Directive 2008/105/EC on Environmental Quality Standards in the Field of Water Policy (the EQS Directive), Member States are obliged to establish an inventory of emissions, discharges and losses of all Priority Substances and pollutants listed in Part A of Annex I to this Directive.

The inventories should give information on the relevance of the substance at the spatial scale of the river basin district, and on the loads discharged to the aquatic environment, thus supporting Member

---

[5] Honti, M., Bischoff, F., Moser, A., Stamm, C., Baranya, S., & Fenner, K. (2018). Relating Degradation of Pharmaceutical Active Ingredients in a Stream Network to Degradation in Water-Sediment Simulation Tests. Water Resources Research, 54(11), 9207–9223. https://doi.org/https://doi.org/10.1029/2018WR023592

[6] Fuchs, S., Kaiser, M., Kiemle, L., Kittlaus, S., Rothvoß, S., Toshovski, S., Wagner, A., Wander, R., Weber, T., & Ziegler, S. (2017). Modeling of Regionalized Emissions (MoRE) into Water Bodies: An Open-Source River Basin Management System. In Water (Vol. 9, Issue 4). https://doi.org/10.3390/w9040239

[7] Venohr, M., Hirt, U., Hofmann, J., Opitz, D., Gericke, A., Wetzig, A., Natho, S., Neumann, F., Hürdler, J., Matranga, M., Mahnkopf, J., Gadegast, M., & Behrendt, H. (2011). Modelling of Nutrient Emissions in River Systems – MONERIS – Methods and Background. Int. Rev. Hydrobiol., 96(5), 435–483. https://doi.org/https://doi.org/10.1002/iroh.201111331

[8] Morling, K., & Fuchs, S. (2024). Zusammenführung der bundesweiten Modellierung von Wasserhaushalt (LARSIM-ME) und Stoffeinträgen (MoRE).

[9] Kittlaus, S., Clara, M., van Gils, J., Gabriel, O., Broer, M. B., Hochedlinger, G., Trautvetter, H., Hepp, G., Krampe, J., Zessner, M., & Zoboli, O. (2022). Coupling a pathway-oriented approach with tailor-made monitoring as key to well-performing regionalized modelling of PFAS emissions and river concentrations. Science of The Total Environment, 849, 157764. https://doi.org/10.1016/j.scitotenv.2022.157764

[10] Liu, M., Saracevic, E., Oudega, T. J., Obeid, A. A. A., Nagy-Kovács, Z., László, B., Kittlaus, S., Zoboli, O., Krampe, J., Derx, J., & Zessner, M. (2025). Investigating the extent of PFAS contamination in the Upper Danube Basin across environmental compartments. Environmental Sciences Europe, 37(1), 99. https://doi.org/10.1186/s12302-025-01141-6

[11] Gabriel, O., & Broer, M. B. (2023). Report on improved system understanding as basis for adapted transnational emission modelling at DRB scale - O T2.2 - Output Danube Hazard m3c project. https://dtp.interreg-danube.eu/uploads/media/approved_project_output/0001/56/52b44806c4c7637a97f77af4f9476a043bbcee60.pdf

[12] Common Implementation Strategy for the Water Framework Directive (2000/60/EC) Guidance Document No. 28, Technical Guidance on the Preparation of an Inventory of Emissions, Discharges and Losses of Priority and Priority Hazardous Substances, ISBN : 978-92-79-23823-9, doi : 10.2779/2764, © European Communities, 2012

States in subsequent river basin management and WFD implementation. For the public, the inventories should give greater transparency with regard to existing problems and on the need for measures to address them. Additionally these inventories will be used by the Commission for compliance checking with the environmental objectives of the WFD (WFD, Article 4) on reduction of discharges, emissions and losses for Priority Substances and cessation or phase out of discharges, emissions and losses for Priority Hazardous Substances (PHS). These inventories will be an important input for the Commission report according to Art. 7(1) of the EQS Directive on the possible need to amend existing acts, and the need for additional specific Community-wide measures such as emission controls.

Furthermore, the preamble of the EQS Directive (Recital 20) foresees the need to have an appropriate tool for quantifying losses of substances occurring naturally, or produced through natural processes, in which case complete cessation or phase out from all potential sources is impossible.

These inventories are to be compiled for every RBD or the national part of international RBDs and to provide not only yearly inputs but also to include, as appropriate, concentrations in sediment and biota (e.g. helping to substantiate the relevance of a substance for the RBD)."

1.2.2 Industrial Emission Directive (IED)

A large share of Europe's pollution stems from industrial production processes, including the emission of air pollutants, waste water discharge and waste generation.

The new Industrial and Livestock Rearing Emissions Directive 2024/1785 (IED 2.0) [13] is the main EU instrument to reduce these emissions into air, water and land, and to prevent waste generation from large industrial installations and intensive livestock farms (pig and poultry). It amends Directive 2010/75/EU. The latest rules will help promote innovation in new and emerging technologies and foster material efficiency and decarbonisation by encouraging greener practices. This will help guide large European industries towards meeting the EU's zero pollution ambition by 2050.

Reporting on industrial emissions has proven to be an important tool for setting and measuring the effectiveness of environmental policies. Various pieces of EU law contain requirements for countries to monitor and report key parameters relevant for verifying the implementation progress of these legal obligations and the pressures on the environment generated by pollutants. In the field of industrial emissions, the following main pieces of legislation contain reporting obligations with a high level of interrelation: the **Industrial Emission Directive (IED)**, **the Large Combustion Plants (LCP) Directive** and **Waste Incineration (WI) Directive** (both superseded via integration into the IED) and the **European Pollutant Release and Transfer Register (E-PRTR) Regulation**. But reporting on industrial emissions is also directly or indirectly linked to other pieces of EU law, in the water, waste, chemicals or climate domains, for example the **EU Seveso Directive** on the control of major-accident hazards involving dangerous substances or the EU scheme for **greenhouse gas emission allowance trading** (EU-ETS).

Proper implementation of EU industrial emissions legislation is important for EU environmental law and policy. Industrial Emissions Directive (IED) is the main EU instrument regulating pollutant emissions from industrial installations. **The IED aims to achieve a high level of protection of human health and the environment taken as a whole by reducing harmful industrial emissions across the EU, in particular through the application of Best Available Techniques (BAT).** The IED covers around 50 000 installations in the EU. These are required to operate in accordance with permits issued by Member States' competent authorities. The permits contain conditions set down in accordance with the principles and provisions of the IED.

[13] https://environment.ec.europa.eu/topics/industrial-emissions-and-safety/industrial-and-livestock-rearing-emissions-directive-ied-20_en

BAT conclusions are the reference for setting or revising the permit conditions of IED installations. BAT is described for each sector in reference documents, the so-called BREFs[14] . The BAT conclusions are a document containing the parts of a BREF with conclusions on BAT and their associated environmental performance levels (BAT-AEPLs), including **the associated emission levels (BAT-AELs).**

Emission levels associated with BAT

Emission levels associated with the best available techniques (**BAT-AELs**) for emissions to water given in these BAT conclusions refer to values of concentrations (mass of emitted substances per volume of water), expressed in µg/l or mg/l.

Unless otherwise stated, the BAT-AELs refer to flow-weighted yearly averages of 24-hour flow-proportional composite samples, taken with the minimum frequency set for the relevant parameter and under normal operating conditions. Time-proportional sampling can be used provided that sufficient flow stability is demonstrated.

The flow-weighted yearly average concentration of the parameter (*cw*) is calculated using the following equation:

$$c_w = \sum_{i=1}^{n} c_i \cdot q_i \Big/ \sum_{i=1}^{n} q_i$$

Where

n = number of measurements;

c_i = average concentration of the parameter during ith measurement;

q_i = average flow rate during ith measurement.

1.2.3 Urban Wastewater Treatment Directive (UWWTD)

The adoption of the **revised Urban Wastewater Treatment Directive (UWWTD)** has introduced additional requirements for municipal wastewater collection and treatment that must be incorporated into the model development process. Directive (EU) 2024/3019[15] constitutes a comprehensive revision of the former 1991 Urban Wastewater Treatment Directive, introducing significantly expanded obligations for EU Member States. The revised framework addresses emerging environmental pressures, public health concerns and climate policy objectives while reinforcing the principle that the polluter should pay.

A major structural change is the extension of the directive's scope. **The minimum threshold for regulated agglomerations is reduced from 2,000 to 1,000 population equivalent (p.e.),** substantially increasing the number of settlements subject to wastewater collection and treatment obligations. Member States must ensure that all agglomerations at or above this threshold are equipped with appropriate collection systems and treatment facilities by 2035, with limited possibilities for derogations under specific conditions.

The revised directive introduces stricter treatment requirements with clearly defined deadlines: (i) Secondary treatment (biodegradable organic matter removal) becomes mandatory for all agglomerations ≥ 1,000 p.e. by 2035; (ii) Tertiary treatment (nutrient removal, nitrogen and phosphorus) is required reinforcing EU efforts to combat eutrophication and protect sensitive waters; and (iii) A new and transformative requirement is the **introduction of quaternary treatment** for large wastewater

[14] Joint Research Centre (JRC), European Bureau for Research on Industrial Transformation and Emissions

[15] Directive (EU) 2024/3019 of the European Parliament and of the Council of 27 November 2024 concerning urban wastewater treatment (recast)

treatment plants by 2045, **targeting the removal of micropollutants, including pharmaceutical residues and chemicals from personal care products**. This obligation implies the deployment of advanced treatment technologies and significant upgrades of existing infrastructure.

By 2045, large wastewater treatment plants (≥ 150,000 p.e.) must remove micropollutants. Furthermore, areas where micropollutant concentrations pose a risk to human health or the environment (e.g., low-dilution waters, drinking water catchments, bathing waters, aquaculture) must be identified. In these sensitive areas, quaternary treatment must also be applied by 2045 (with a risk-based approach determining where necessary).

The directive goes beyond treatment plant performance to mandate a more system-wide planning approach. Member States must establish Integrated Urban Wastewater Management Plans that cover sewer networks, treatment capacity, stormwater management and combined sewer overflows. These plans must also address climate change adaptation, encouraging the adoption of nature-based and green–blue infrastructure solutions to reduce pollution and enhance resilience.

Overall, Directive (EU) 2024/3019 represents a transition from a compliance-focused wastewater policy to a comprehensive, forward-looking framework that incorporates environmental protection, public health, climate neutrality, and the principles of a circular economy. This entails substantial new responsibilities in terms of planning, technology, monitoring and financing, with long-term implications for the development and governance of urban water infrastructure.

1.3 REQUIREMENT FOR AN IMPROVED EMISSION MODEL TO OVERCOME DATA GAPS

The MoRE model provides a framework for quantifying emissions. However, its applicability depends significantly on the quality of the available data sources. Therefore, a conceptual revision of the model was required, as well as the collection of new input data and the upgrading of calculation approaches. This work was supported by a gap analysis based on the Danube Hazard m^3c inventory database.

The intention was to upgrade the model so that it can account for regional differences within the Danube River Basin. This primarily depends on data availability. To this end, an intensive data collection exercise was conducted, which included acquiring new data by investigating the Member States' monitoring results published in the 3rd River Basin Management Plan (mandatory measurements of priority substances according to Directive 39/2013/EU, as well as measurements of substances on the watch list). Furthermore, the BME initiated bilateral meetings with all PPs to gather information on the potential data supply.

The model combines a source-based approach with the quantification of emission pathways. The former requires the development of specific emission factors (e.g. per capita emissions). Therefore, in addition to collecting concentration data (e.g. HS concentration in domestic and industrial wastewater, soil and groundwater), volumetric information (e.g. municipal and industrial wastewater discharges, and the consumption and production of pharmaceuticals and industrial chemicals) had to be collected as well.

Some shortcomings of the MoRE model needed to be addressed. For example, retention in river systems was not considered for hazardous substances. To address this issue, the model has been updated with a new concept. This new algorithm will also be able to perform a sensitivity analysis of the parameters. The revised model approach can perform the scenario assessments required to assess the new policy challenges and objectives at the European level. As part of the analysis, changes in input data and model parameters were made and their impact assessed (e.g. upgrading of WWTP technologies, implementation of measures to reduce diffuse urban and rural pollution, etc - in line with new regulations and policies (recast of the UWWTD, IED, CAP and the EU Zero Pollution Action Plan).

2 OVERVIEW OF THE MODEL DEVELOPMENT PROCESS

2.1 GENERAL DESCRIPTION OF THE MORE MODEL APPROACH

The MoRE model[16] is a semi-empirical emission model, which operates on the mesoscale (tenth to hundreds of square kilometers) and on annual time steps (in this model application period 2015-2020). It is a further development derived from the MONERIS emissions model[17] mainly developed for nutrients and differs in particular by a modified technical model realization.

The MoRE model was initially available only in a German version for an extended number of organic and inorganic micro pollutants. Additionally, a very basic application for nutrients was available in English. The latter was taken and built up into a fully functional English version for a wide range of substances and with a wide range of calculation approaches in the Interreg Danube Hazard m^{3}c project. In the current application, the model was further translated and updated to fulfil some additional requirements of the current work.

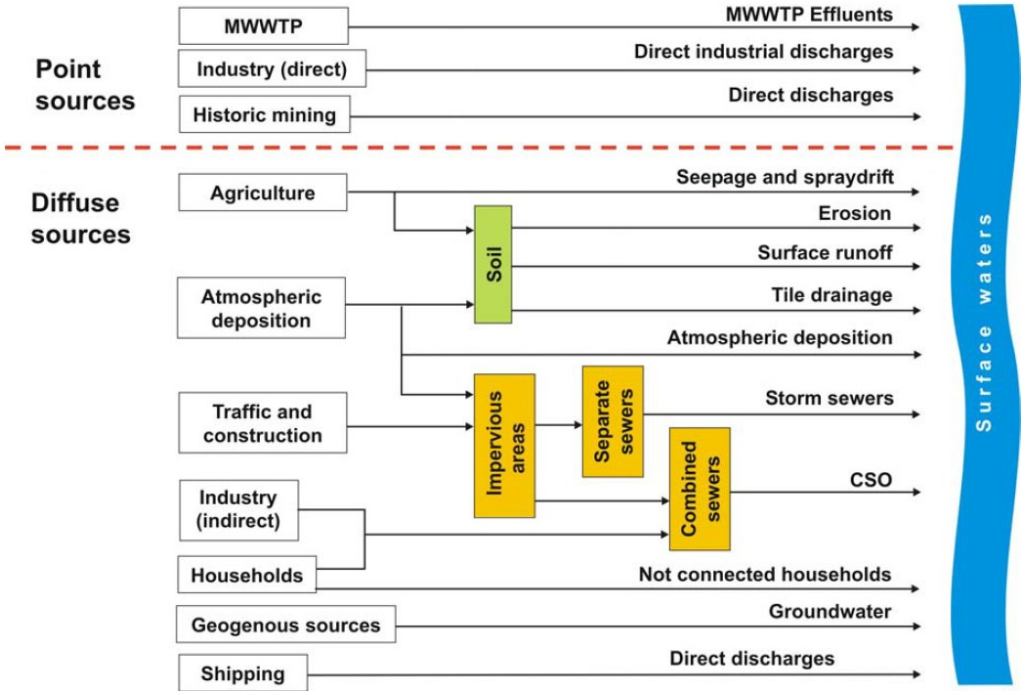

Figure 1 - Substance sources and emission pathways for water pollution (Fuchs et al., 2010)

In the current version some modification of the pathways was undertaken, e.g. the historic mining was included as an industrial pathway, which is detailed to individual industry segments.

[16] Fuchs, S., Kaiser, M., Kiemle, L., Kittlaus, S., Rothvoß, S., Toshovski, S., Wagner, A., Wander, R., Weber, T., & Ziegler, S. (2017). Modeling of Regionalized Emissions (MoRE) into Water Bodies: An Open-Source River Basin Management System. In *Water* (Vol. 9, Issue 4). https://doi.org/10.3390/w9040239

[17] Behrendt, H., Kornmilch, M., Opitz, D., Schmoll, O., & Scholz, G. (2002). Estimation of the nutrient inputs into river systems – experiences from German rivers. *Reg. Environ. Change*, *3*(1), 107–117. https://doi.org/10.1007/s10113-002-0042-3

2.2 PATHWAYS FOR HEAVY METALS

2.2.1 Industries

Industry is one of the most important pathway for heavy metals and PFAS, as considerable emissions are related to these activities either directly via point sources or indirectly via diffuse pathways.

Point source pathways

Within this group we can identify three types of discharge:

1. Direct wastewater discharge through an industrial wastewater treatment facility that receives industrial discharge from one/many industrial facility/facilities.
2. Industrial facilities often discharge wastewater into the public sewer system, which then flows through the urban WWTP (UWWTP) into the receiving watercourse.
   From a modelling perspective, it would be useful to handle the industrial wastewater fraction of UWWTPs as a separate 'industrial' pathway, as this would allow better estimation for the necessary pollution control measures. However, this would require an internationally homogeneous database (fraction of industrial wastewater and its industrial sources), which is not currently available.
3. Direct industrial discharge to a receiving water body without wastewater treatment facility.

While the geographical location of direct point sources emissions is relatively easy to determine, determining the quantity of pollutants emitted is a complex task. The loads from industrial sites are only reported and collected for larger sites and for a few regulated contaminants. Our approach assumes that all sites – small, medium and large – are relevant emitters that need to be quantified in some way. The methods for quantification are detailed in the relevant sections of Chapter 4.

Parameters introduced for this pathway include concentration, load, discharge variables, e.g.: for the cadmium emission from power generation sites (TPS) the variable ID_TPS_CONC_HM_CD is defined following the naming convention of the MoRE model, in which the ID refers to the industrial site, the TPS refers to the Thermal Power Station type, HM refers to the substance group heavy metal, while CD refers to cadmium. If concentrations are available, then discharge [m^3/year] is also needed: ID_TPS_ps_Q. If emitted loads are reported, the ID_TPS_E_HM_CD parameter is used, if not then emission factors may be used ID_TPS_EF_HM_CD , where EF refers to the emission factor notation.

The detailed inventory of industry types, which potentially emit metals are described in Chapter 4.2.

Diffuse pathways

In these pathways we can also differentiate between different cases:

- Through the emission via the atmosphere
  This type of emission is modelled with the following principle: Aerial emission is linked to the atmospheric deposition pathway. We use two approaches:
  a. If aerial emission values are known, either by direct reported data, by generated emission factors, the emitted loads are converted to atmospheric deposition values by simple assumptions of the impact area and by assuming, that all the emitted substance load is deposited within that area. This is described under atmospheric deposition section.
  b. If the emission amount is unknown, then the background concentration of heavy metals is multiplied by an emission multiplication factor $F_{AD,HM}$ (currently not used in the model).

- Emission from industrial wastes: A typical example are the tailing facilities of mines, where high stock of heavy metal pollutants is stored, and potentially released through leaching, or accidental pollution (e.g. *Bia Mare*, *Băile Borșa* accidents in Romania, or the Red sludge acci-

dent in *Ajka,* Hungary). While the former can be estimated based on concentrations of leachates from these sites (See legacy hot spots, section 2.2.9), the latter, even though this can be an important source, due to lack of data, we do not account for this in the model.

2.2.2 Municipal waste water treatment plants

This pathway describes direct emissions from waste water treatment plants to receiving water bodies. The method includes the estimation of the annual metal loads from each individual wastewater treatment plants in the first step.

The calculation concept is the following:

WWTP_ps_E_HM = WWTP_ps_Q · WWTP_CONC_HM
where
WWTP_ps_E is the point source load from the individual treatment plants to the receiving water body.
WWTP_ps_Q is the discharge of the individual treatment plants.
WWTP_CONC_ is the concentration of the given metal in the treatment plant effluent.

The overall equation taking into account treatment stages is as follows:

WWTP_ps_E_HM =
if(WWTP_treatment_type=0, WWTP_ps_Q*WWTP_CONC_NOTREAT_HM/1000000,
if(WWTP_treatment_type = 1, WWTP_ps_Q*WWTP_CONC_PRIM_HM/1000000,
if(WWTP_treatment_type = 2, WWTP_ps_Q*WWTP_CONC_SEC_HM/1000000,
f(WWTP_treatment_type = 3, WWTP_ps_Q*WWTP_CONC_TERT_HM/1000000,
WWTP_ps_Q*WWTP_CONC_QUART_HM/1000000))))

The concentration values for the primary, secondary, tertiary and quaternary treatment have been derived from available effluent concentration data (see Section 4.1.1).

In the second step the emissions are aggregated for the analytical units by summing up all the treatment plants discharging to the main receiving rivers within the AU, including the small plants that are not handled as point source units but as estimated total loads:

WWTP_E_HM is the heavy metal emission for a given analytical unit, $WWTP\_ps\_E\_HM_i$ is the point source emissions from the individual treatment plants (i) in the AU and WWTP_small_E_HM is the load from small treatment plants, which is calculated by the following method:

WWTP_small_E_HM = WWTP_s_CONC_HM * WWTP_small_Q / 1000 / 1000

Where WWTP_s_CONC_HM is the heavy metal concentration of the effluents from small WWTP-s. These values (Figure 2) were also obtained from observation data and were introduced in the Danube Hazard m^3c project.

WWTP_s_CONC_HM_AS	1		3	µg/L	18.11.2022	waste_water
WWTP_s_CONC_HM_CD	1		1.55	µg/L	30.09.2022	DHm3c all data
WWTP_s_CONC_HM_CR	1		0.00221	µg/L	30.09.2022	DHm3c all data
WWTP_s_CONC_HM_CU	1		2.81	µg/L	30.09.2022	DHm3c all data
WWTP_s_CONC_HM_NI	1		4.51	µg/L	30.09.2022	DHm3c all data
WWTP_s_CONC_HM_PB	1		6.81	µg/L	30.09.2022	DHm3c all data
WWTP_s_CONC_HM_ZN	1		1.43	µg/L	30.09.2022	DHm3c all data

Figure 2 -- Excerpt of data base used for the model, showing constant concentration values for small WWTPs

2.2.3 Urban pathways

In the Moneris and MoRE approach, the urban pathways are the most complex pathways with plenty of empirical assumptions of discharge amounts per runoff pathways or concentration values related to them. The urban pathways include the following:

- Emission through **Combined Sewer Overflows (CSOs)** from inhabitants, commercial areas and impervious areas.
- Emission from surface runoff through **separated storm sewer systems** („**_SS_**“ notation).
- Emission through **inhabitants not connected to sewer systems** („**_NSS_**“ notation), which is a pathway that emits to groundwater. Currently this emission is not fed into the rivers as no information is available on retention of the modelled substances in soils and groundwater.
- Emission from runoff from impervious areas that are not connected to sewer systems („_NSS_“ notation).
- Emission through **sewer network not connected to treatment plants** (only connected to sewers systems, „**_OSS_**“ notation). Part of this is transported into the groundwater, which is not reaching the river system in the current modelling setup.

The detailed description of these pathways is not introduced here as these equations are not developed further within this model application. An overview of the pathway flowchart can be seen in the Figure 3. Further details can be found in O2.3 Annex I.

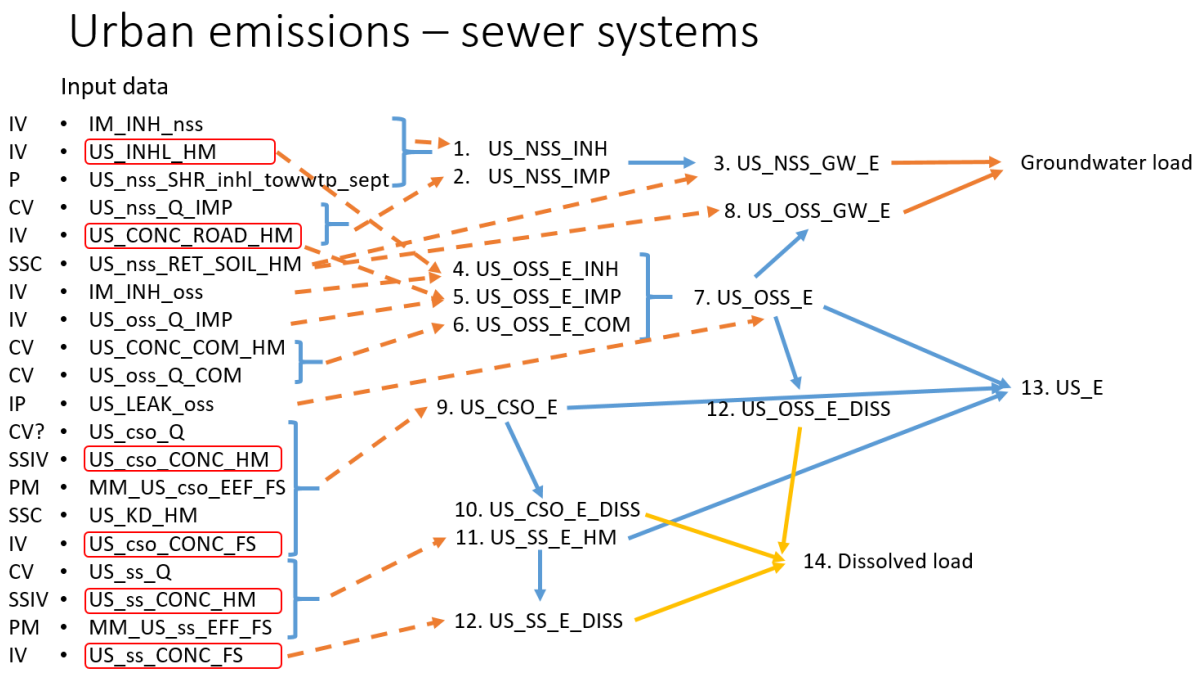

Figure 3 - A schematic overview of the urban pathways in the MoRE model application for heavy metals.

2.2.4 Open roads (outside urban areas)

Open (or country) roads emit primarily heavy metals due to the intensive road traffic. This emission has been quantified for the seven pilots in the Danube Hazard m^3c project. In the current application, the open road pathway is used similarly, using the following relationship:

$$OR\_E\_HM = (OR\_SFL\_HM - MM\_AD\_RATE\_dep\_red\_HM / 100 \cdot BI\_RATE\_dep\_HM) \cdot IM\_A\_OR\_qsr / 10$$

where

OR_SFL_HM is a specific emission of HM on open road surfaces. Values were derived from the Austrian Stobimo project[18] (g/ha/a),

MM_AD_RATE_dep_re – mitigation measure applied for atmospheric deposition,

BI_RATE_dep_HM – atmospheric deposition of HM in general (applied in the AU everywhere),

IM_A_OR_qsr is the area of non-urban roads discharging into water bodies.

A key parameter in the estimation is the IM_A_OR_qsr, which is hard to estimate. Discharge from roads, especially motorways, are reduced by retention ponds in many countries. However, information is hardly available. For the sake of the simplicity of the modelling, it is assumed that all road surfaces are connected; however, not all roads are estimated in the approach (see Section 3.3.2).

2.2.5 Atmospheric deposition

Generally, the source of metals in the atmosphere is related to industrial activities, traffic, agriculture, and some other sources. The emitted particles and dissolved phase elements are then transported by atmospheric processes to different parts of the basin or outside of it.

There are two approaches to deal with the description of this pathway:

1. Use results of a complex atmospheric transport model. The EMEP program[19] has a transport model (EMEP MSC-E) built up for the whole of Europe. This model handle metals and persistent organic pollutants apart from nutrients and some other substances. The output from this model is the most possible accurate input for the MoRE model; its availability is however under negotiation at the moment.
2. The second option covers a very simple estimation of the deposition from "background" emissions and hot spot emissions. In this approach by "background" we mean the deposition of metals in remote places (outside the impact area of hot spots) caused by long-range transport.

Description of the 2nd method:

Atmospheric deposition (AD) is given by the sum of two sources:

1. There is a background value, which is determined from the extrapolation of monitoring results carried out for seven pilot sites within the Danube Hazard m^{3}c project.
   In this pathway the calculation method is the following:

AD_E_HM = BI_RATE_dep_HM * (1- MM_AD_RATE_dep_red_HM / 100) * IM_A_WS / 10
where
AD_E_HM is the metal emission via atmospheric deposition,
BI_RATE_dep_HM is the metal deposition rate (periodical AU variable),
MM_AD_RATE_dep_red_HM is the mitigation measure, which defines the level of reduction for the deposition,
IM_A_WS is the water surface.

2. There is hot spot emission in the vicinity of known aerial sources, such as oil refinery emissions, or chemical industry sites, where these substances are used. The value of the emitted load comes

[18] Amann, A., Clara, M., Gabriel, O., Hochedlinger, G., Humer, M., Humer, F., Kittlaus, S., Kulcsar, S., Scheffknecht, C., Trautvetter, H., Zessner, M., & Zoboli, O. (2019). *STOBIMO Spurenstoffe - Stoffbilanzmodellierung für Spurenstoffe auf Einzugsgebietsebene. BMNT. Final report.*

[19] EMEP (2023). EMEP Monitoring Network. European Monitoring and Evaluation Programme under the UNECE Convention on Long-range Transboundary Air Pollution. Norwegian Meteorological Institute, Oslo.

from industrial site emissions (Figure 4). The **impact area** of the emissions is determined by the following assumptions:

a. Chimney emissions have a concentric emission area.
b. The impact area has 50 km wide diameter, which is an order of magnitude estimation based on (Galloway et al., 2020[20]; Hangen et al., 2010[21]; Moghadasi et al., 2023[22]).
c. Emission within the emission area is homogeneous.
d. Total emission from AD per AU is estimated by the deposition rate and the water surface area (see below).
e. Long range transport is considered by the background values.

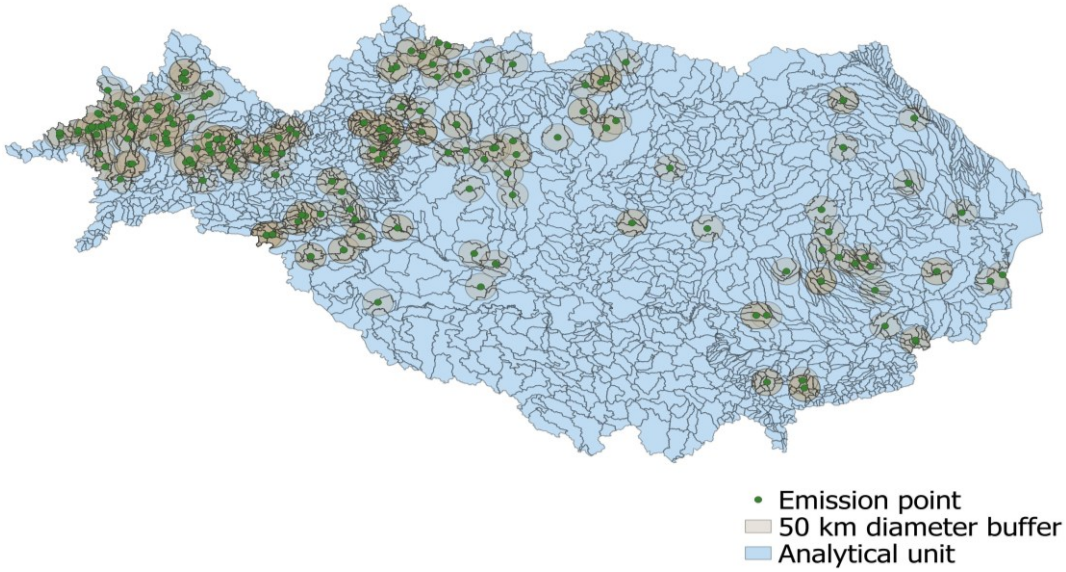

Figure 4 - Industrial site emissions with their 50 km diameter impact area

AD on water surfaces from industrial emitters:

In the first step aerial emitters are **aggregated for the analytical unit, but this is done outside the model, as a pre-processing step.** For this, first, the distribution ratios between the AUs are estimated using a uniform circular emission zone around each chimneys with a diameter of 50 km (rough estimate of the zone, where short distance deposition occurs). These circular zones are intersected with the AUs and overlap ratios are calculated. Loads per AU are calculated based on these ratios, loads are then aggregated for each Aus.

$$ID\_AD\_HM = \sum_{i=1}^{n} ID\_AD\_ps\_HM_i$$

where $ID\_AD\_ps\_HM_i$ are the individual industrial point source emitters of HM.

[20] Galloway, J. E., Moreno, A. V. P., Lindstrom, A. B., Strynar, M. J., Newton, S., May, A. A., & Weavers, L. K. (2020). Evidence of Air Dispersion: HFPO–DA and PFOA in Ohio and West Virginia Surface Water and Soil near a Fluoropolymer Production Facility. Environmental Science & Technology, 54(12), 7175–7184. https://doi.org/10.1021/acs.est.9b07384

[21] Hangen, E., Klemm, A., Kronawitter, H., & Schubert, A. (2010). Perfluorooctanoate (PFO) in Forest Soils near a Fluoropolymer Manufacturing Facility. Water, Air, & Soil Pollution, 212(1), 491–499. https://doi.org/10.1007/s11270-010-0365-5

[22] Moghadasi, R., Mumberg, T., & Wanner, P. (2023). Spatial Prediction of Concentrations of Per- and Polyfluoroalkyl Substances (PFAS) in European Soils. Environmental Science & Technology Letters, 10(11), 1125–1129. https://doi.org/10.1021/acs.estlett.3c00633

In the second step **the deposition load on water surfaces** is calculated by the proportion of the water surface area to the whole catchment area in the model:

ID_AD_E_HM = ID_AD_HM · BI_A_WAT/BI_A

where

ID_AD_E_HM is the total HM deposition on water surfaces, ID_AD_HM is the total atmospheric deposition of heavy metals from industrial areal emitters for the given AU, BI_A_WAT is the water surface area of the AU and BI_A is the total area of the AU.

2.2.6 Erosion

Erosion is an important pathway of metal transports due to the presence of all the metals substances in soils. Their presence comes from multiple sources, including **atmospheric deposition as background** and also **as hot spot emission** from chimneys or local mining activities. Another important source is the **agriculture**, where some fertilizers and pesticide compounds contain some of these elements. In this model we account for higher emissions from aerial hot spot emissions through this pathway also.

The concept of the erosion estimation is simple in the project, the emission is calculated by multiplying metal concentration of the soils with annual soil loss. This is calculated for five distinct land uses: **agricultural lands, grassland, natural vegetation, glaciers and mountainous areas**. The first three are given by the estimates from soil metal content maps, the latter two are calculated from rock metal concentration values (adapted from the Danube Hazard m^3c model).

Soil concentration values are defined as described in Chapter 4.3.

Soil loss is a key variable for heavy metal emissions especially in areas with relatively high metal concentrations (e.g. some regions of the western-Balkans). However, soil loss estimation is not necessary to estimate within the project as there are open data sources with this information: For most regions the **European Soil Data Centre contains erosion information**, while for the remaining areas **global soil estimates are available**. This information has been obtained from the Moneris dataset. Please note, that the **annual variation of soil loss** values can be significant, which is taken into account in the model by a simple approach of the ratio of **actual summer precipitation compared to the long-term average summer precipitation**. Actual annual soil erosion estimates would be preferable for the entire basin, but at this analysis would not fit within the framework of this project.

Soil loss for glaciers (two variants: 25 t/(ha · year) and 35 t/(ha · year) and mountains (two variants: 10 t/(ha · year) and 0.2 t/(ha · year) are estimated by constant values. The former are defined in the approach described in Zoboli et al. (2018) and Zessner et al. (2011)[23]. The latter are described in Zoboli et al. (2018) and in the Moneris handbook (Venohr et al., 2011a)[24].

Erosion for agricultural lands is calculated by the following formula (Original Moneris model):

ER_agrl_E_HM = [ER_agrl_SL_AL · ER_agrl_CONT_SOIL_top_AL_HM · (1 - MM_ER_EFF_AL_SED /100) + ER_agrl_SL_PST · ER_agrl_CONT_SOIL_top_PST_HM]· ER_SDR · 0.01 · ER_ENR_AL /1000

where:

- ER_agrl_SL_AL is the soil loss from arable land,

[23] Zessner, M., Kovacs, A., Schilling, C., Hochedlinger, G., Gabriel, O., Natho, S., Thaler, S., & Windhofer, G. (2011). Enhancement of the MONERIS Model for Application in Alpine Catchments in Austria. *International Review of Hydrobiology*, *96*(5), 541–560. https://doi.org/https://doi.org/10.1002/iroh.201111278

[24] Venohr, M., Hirt, U., Hofmann, J., Opitz, D., Gericke, A., Wetzig, A., Natho, S., Neumann, F., Hürdler, J., Matranga, M., Mahnkopf, J., Gadegast, M., & Behrendt, H. (2011). Modelling of Nutrient Emissions in River Systems – MONERIS – Methods and Background. *Int. Rev. Hydrobiol.*, *96*(5), 435–483. https://doi.org/https://doi.org/10.1002/iroh.201111331

- ER_agrl_CONT_SOIL_top_AL_HM is the HM content of the soil in the arable lands in the given analytical unit,
- MM_ER_EFF_AL_SED is the efficiency of the erosion reduction measure introduced in the model (if applicable),
- ER_agrl_SL_PST is the soil loss from pastures,
- ER_agrl_CONT_SOIL_top_PST_HM is the HM content of the soil in pastures in the given analytical unit,
- ER_SDR is the sediment delivery ratio, which is calculated according to the original method defined in the Moneris model[7],
- ER_ENR_AL (-) is the enrichment ratio, which is calculated according to the original method defined in the Moneris model[7]

ER_ENR_HM =

if(ER_E_spec_AGRL_SED < 1, then **ER_FCT_a_ENR_HM**;

if(ER_FCT_a_ENR_HM * ER_E_spec_AGRL_SED ^ ER_EXP_ENR_HM < 1, then 1;

otherwise ER_FCT_a_ENR_HM · ER_E_spec_AGRL_SED ^ ER_EXP_ENR_HM))

The values for HM enrichment factors are defined in the MoRE model by (Fuchs et al., 2002)[25].

Erosion for glaciers and mountainous areas are calculated with the following formula:

ER_glc_E_SED = ER_glc_E_spec_lt_SED · 100 · IM_A_GLC

HM emission via glacier erosion is calculated:

ER_glc_E_HM = ER_glc_E_SED · ER_CONT_ROCK_HM / 1000

Where ER_glc_E_SED is the sediment input from glaciers and ER_CONT_ROCK_HM is the HM content of the rock.

Erosion from natural areas is also calculated in the model as for agricultural soils. The method applied is the same as the one used in the DHm3c project, no changes have been implemented.

ER_nat_E_SED * ER_nat_CONT_SOIL_HM / 1000

In previous applications constant values have also been used to estimate sediment loss from natural areas (ER_nat_E_SED, but in this case, we accounted for it using Soil loss information from regional estimates. The constant option is also in the model as a calculation variant.

If (ER_CODE_method_nat > 10, ER_NAT_E_SED_max, ER_nat_SL_AU * 100 * ER_nat_SDR / 100 * IM_A_NAT * ER_FCT_corr_FCT_r_RUSLE)

where ER_nat_SL_AU is the AU specific natural soil loss rate from natural areas, ER_nat_SDR is the sediment delivery ratio for natural lands, and ER_FCT_corr_FCT_r_RUSLE is the correction factor of USLE R factor for the given year.

ER_nat_SDR is calculated by empirical formula based on slope (BI_SLP):

ER_nat_SDR = ER_FCT_SDR_nat * exp (BI_SLP * ER_EXP_SDR_nat) * 100

[25] Fuchs, S., Scherer, U., Hillenbrand, T., Marscheider-Weidemann, F., Behrendt, H., & Opitz, D. (2002). *Schwermetalleinträge in die Oberflächengewässer Deutschlands. Unter Mitarbeit von H. Herata*. https://www.umweltbundesamt.de/sites/default/files/medien/publikation/long/2225.pdf

ER_FCT_SDR_nat and ER_EXP_SDR_nat are the multiplier factor and exponent of the empirical formula.

corr_FCT_r_RUSLE is calculated by the division of the actual R factor and the long term R factor:

corr_FCT_r_RUSLE = ER_FCT_r_RUSLE_alp / ER_FCT_r_lt_RUSLE_lt

where

ER_FCT_r_lt_RUSLE_lt = if (ER_PREC_s_lt < 0, 1, (ER_PREC_s_lt * ER_FCT_a_FCT_r_RUSLE_alp) ^ ER_EXP_a_FCT_r_RUSLE_alp - ER_FCT_b_FCT_r_RUSLE_alp)

and

ER_FCT_r_RUSLE_alp = if (IM_PREC_s < 0, 1, (IM_PREC_s * ER_FCT_a_FCT_r_RUSLE_alp) ^ ER_EXP_a_FCT_r_RUSLE_alp - ER_FCT_b_FCT_r_RUSLE_alp)

In which ER_PREC_s_lt is the long-term average summer precipitation and IM_PREC_s is the annually calculated summer precipitation. Regression factors ER_FCT_a_FCT_r_RUSLE_alp, ER_EXP_a_FCT_r_RUSLE_alp and ER_FCT_b_FCT_r_RUSLE_alp are constant for the entire basin.

2.2.7 Surface runoff

Surface runoff, taken into account as wash-off from soils and vegetated surfaces also contain metals in small concentrations, especially in agricultural areas. In the earlier model approaches, the concentration value was constant based on monitored values (Danube Hazard pilot model, Stobimo project data[26]). As surface runoff measurements are very limited, built in model constants are used as reference values and soil metal concentrations are used to scale the variation of the surface runoff. The actual concentrations are calculated from soil concentrations and partition coefficients available from sediments.

The following assumptions are made:

- Even small surface runoff events deliver soils and transport metals.
- Annual delivered soil/suspended matter is proportional to the transported metal in dissolved form of total surface runoff.
- Known Kd_{sed} (from earlier model versions) values for metals are representative of the runoff/erosion events.

Based on these assumptions the surface runoff concentrations of metals are calculated with the following formulas:

1. ER_agrl_CONT_SOIL_top_AL_ENR_HM = ER_agrl_E_HM / SR_VEG_Q
   (calculating the Metal concentration in the sediment from the total erosion driven Metal load that already takes into account sediment delivery and enrichment, unit: ug/kg HM)
2. SR_CONC_CALC_HM = ER_agrl_CONT_SOIL_top_AL_ENR_HM / SR_KD_sed_HM * UF
   (SR_CONC_CALC_HM is the calculated concentration in the ater face of surface runoff, SR_KD_sed_HM is the Kd coefficient of th sediment, UF is th eunit conversion factor)

Where

ER_agrl_CONT_SOIL_top_AL_ENR_HM is the concentration

[26] Amann, A., Clara, M., Gabriel, O., Hochedlinger, G., Humer, M., Humer, F., Kittlaus, S., Kulcsar, S., Scheffknecht, C., Trautvetter, H., Zessner, M., & Zoboli, O. (2019). *STOBIMO Spurenstoffe - Stoffbilanzmodellierung für Spurenstoffe auf Einzugsgebietsebene. BMNT. Final report.*

2.2.8 Landfill sites

Landfill sites store a large variety of materials, products that contain Mmetal compounds such as electronic products, batteries, etc.. Leachate of these sites cause the contamination of groundwater that is transferred to surface water. These type of sites is treated as legacy pollution deposit in groundwater. Concentration values are available for leachate water from a few sites in Europe (Greece: Nika et al., 2020, AT: Promisces project report 2024).

Retention in groundwater is not taken into account. **Emissions from landfills** are calculated using a concentration · discharge approach:

$$LF\_E\_HM = LF\_CONC\_HM \cdot LF\_ps\_Q$$

where

LF_CONC_HM is the **concentration** of the given heavy **metal in the leachate**,

LF_ps_Q is the discharge of the given landfill sites.

The *ps* notation shows, that it is considered as a point source in the model.

$$LF\_ps\_Q = BI\_A\_LF \cdot (BI\_PREC\_LF - AD\_EVAPO\_LF)$$

where

BI_PREC_LF is the **annual precipitation** at the landfill sites,

AD_EVAPO_LF is the **annual evapotranspiration** at the landfill sites.

These latter values have been calculated from annual precipitation (Figure 5) and evapotranspiration (Figure 5) grid time series that are the output of the Danube Water Balance model. It thin model, daily discharges have been calibrated for the entire Danube Basin with satisfactory results, therefore both precipitation and evapotranspiration can be considered as a reliable input.

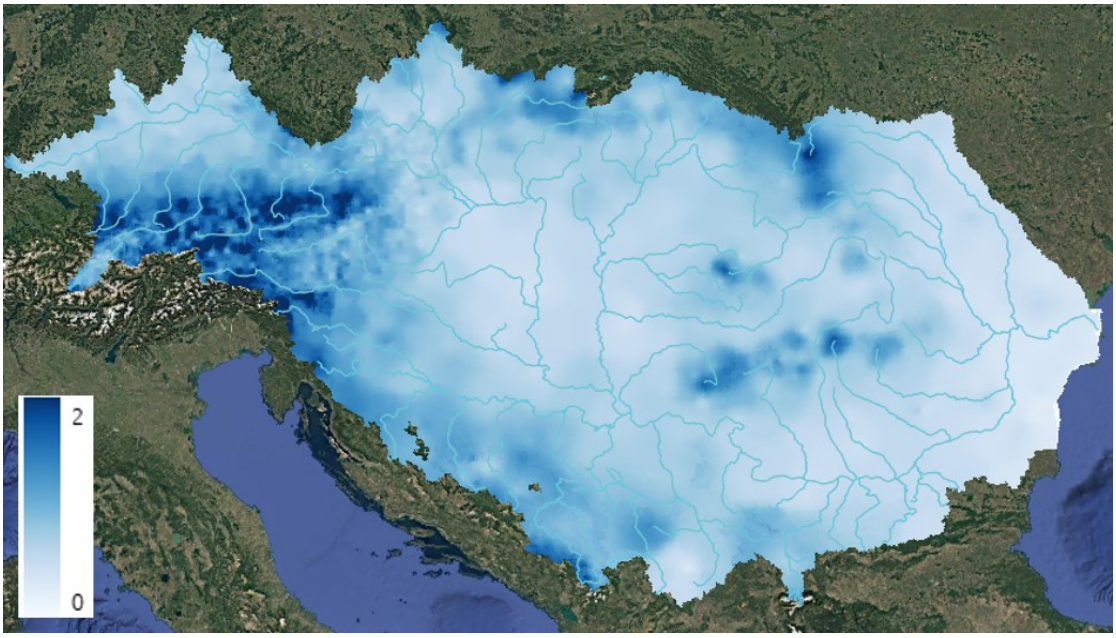

Figure 5 - Annual total precipitation as used in the CWatM model of the Danube Water Balance application

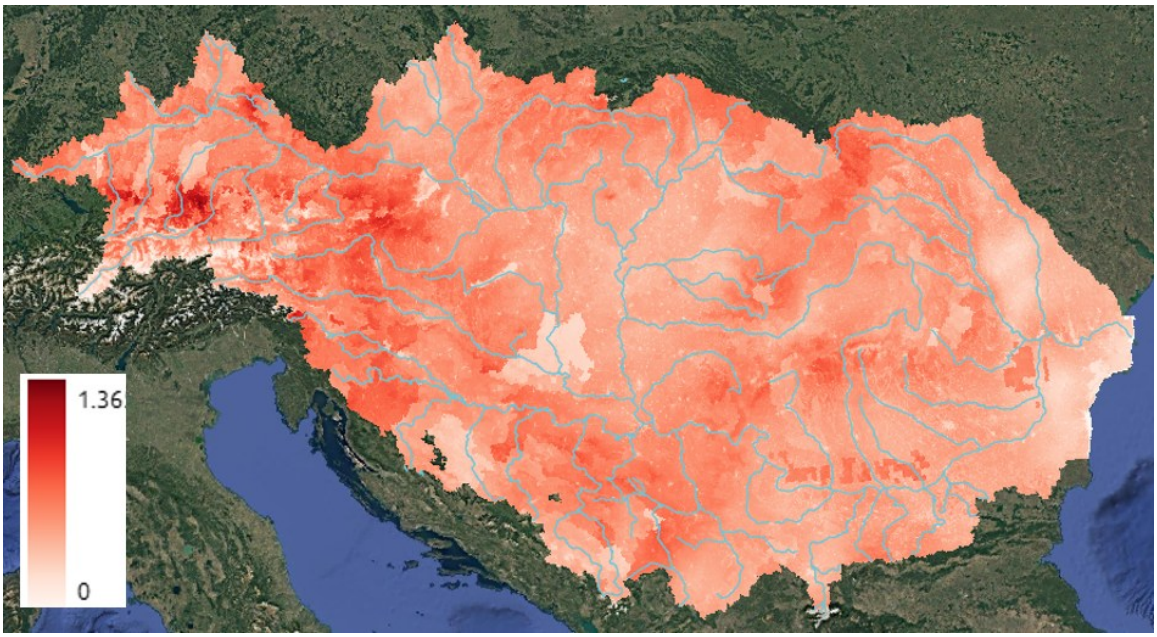

Figure 6 - Total evapotranspiration as calculated by the CWatM model of the Danube Water Balance application

2.2.9 Other Legacy hot spots

This pathway is created to account for sites, where known groundwater contamination is present. There have been many sites all over Europe, where the soil and groundwater have been contaminated with metals (Metallokemia and Borsodchem (Hg) from Hungary, etc.). Most of these sites have been rehabilitated, but there are still sites with existing contamination. Data of contaminated sites can be loaded in the model from national databases. **Emissions from legacy hot spots** are calculated by using a concentration · discharge approach:

$$LHSG\_E\_HM = LHSG\_CONC\_HM \cdot LHSG\_ps\_Q$$

where

LSHG_E_HM is the emission though groundwater from general (unspecified) legacy hot spot

LHSG _CONC_HM is the concentration of the given heavy metal in the contaminated groundwater,

LHSG_ps_Q = BI_A_LHSG · (BI_PREC_ LHSG - AD_EVAPO_ LHSG)

where

BI_PREC_ LHSG is the **annual precipitation** at the legacy hot spots areas,

AD_EVAPO_ LHSG is the **annual evapotranspiration** at the legacy hot spots areas.

2.2.10 Groundwater

The groundwater pathway can be differentiated between several hot spots and background load. Background load is associated with metal concentration in groundwater due to leaching of metal content in soils. With the term background we thus do not refer to natural background but to large scale diffuse sources of metal emissions. In this section only the groundwater background is described. Hot spot type emissions are described in separate sections (Industrial, Landfill sites, aerodomes). Emissions through the groundwater is calculated by the following formula:

GW_E_HM = GW_Q · GW_CONC_HM · (100-MM_DI_HM)/100
where

GW_Q = (GW_Q_spec · GW_A_rech),
where GW_Q_spec is calculated based on water balance methods applied. If baseflow is coming from the Danube water balance model, the GW_Q_scpec directly calculated form the modelled baseflow, if not, then it is calculated by subtracting all the water balance components from the net total runoff for the given AU.

2.2.11 Tile drainage

In former model applications, the tile drainages were included for nutrients but not for heavy metals or other compounds. The reason for this is the poor knowledge of concentration ranges in this pathway, as only a few monitoring study examined this in detail. In our approach we state that each pathway should be included in the model and try to give the right order of magnitude for the emissions.

There are two options to estimate tile drain flow concentrations:

1. As soil metal concentrations are fairly well known (the order of magnitude, at least), at least on the large scale, the tile drain concentration ranges can be regionalized. The best available method to do this is to account for the adsorption capacity of the soils through the organic matter content, which is one of the main factors that influence the adsorption potentials. For the emission estimation, therefore the Kd,OC values for the metals, and the soil organic fractions have to be known. The former is known by literature values, the latter can be obtained from soil databases. The detailed method is described in 4.2.2.7.
2. Groundwater concentrations are also known (with high uncertainty with regard to the spatial distribution) across the basin, this data can be used to estimate tile drain flow by using a conversion factor that translates concentration values between the two matrices. This latter is used currently in the model, and calculated as follows:

TD_E_HM = (TD_Q * 86400 * 365 * GW_CONC_HM) / (1000 * 1000)*TD_FCT_GW_CONC_RATE_HM
where
TD_E_HM is the Metal emission through tile drainage (t/y)
TD_Q is the discharge in the AU vie tile drainage (m^3/s)
GW_CONC_HM is heavy metal concentration in groundwater (ug/l)

TD_FCT_GW_CONC_RATE_HM (-) is the conversion factor between groundwater concentration and tile drain flow (please refer to section 4.7 for the derivation of these values).

2.3 PATHWAYS FOR PHARMACEUTICALS

Pathways for pharmaceutical compounds are much simpler than for the other two substance groups as there are **only three types of pathways** that are relevant based on the current knowledge.

- Waste water emissions (through municipal WWTP or through direct discharge).
- Urban systems: 1) CSOs and 2) separated storm water runoff.
- Industrial discharge from pharmaceutical production plants (can be significant due to high concentrations).
- Landfills

2.3.1 Municipal waste water treatment plants

2.3.2 Two approaches can be used for the emission estimation via this pathway:

1. Using population data and population specific emissions of the given pharmaceutical compounds.
2. Using waste water discharge data and pharmaceutical concentrations of the WWTP effluents.
   WWTP_ps_E_PHAR = WWTP_ps_Q*WWTP_CONC_PHAR

where WWTP_ps_E_PHAR is the pharmaceutical load in the treatment plant effluent for the given plant.
WWTP_ps_Q is the discharge of the individual treatment plants.
WWTP_CONC_PHAR is the concentration of the given pharmaceutical in the treatment plant effluent.

In the second step the emissions are aggregated for the analytical units by summing up all the emissions from treatment plants discharging to the main receiving rivers within the AU:

$$WWTP\_E\_PHAR = \sum_{i=1}^{n} WWTP\_ps\_E\_PHAR_i$$

where WWTP_E_PHAR is the Metal emission for a given analytical unit

$WWTP\_ps\_E\_PHAR_i$ –is the point source emissions from the individual treatment plants (i) in the AU

Currently, the second approach (discharge * concentration) is applied in the model.

2.3.3 Urban pathways

The pathways within this include the following:

- emission through Combined Sewer Overflows (CSOs) from inhabitants, commercial areas and impervious areas,;
- emission through separated sewer systems (stormwater runoff);
- emission through inhabitants not connected to sewer systems;
- emission through impervious areas not connected to sewer systems;
  emission through sewer network not connected to treatment plants (only connected to sewers systems);
  For pharmaceuticals this pathway is important, as many larger cities have a high share of collected and untreated waste water.
- emission through groundwater originated by the loss from sewer systems (leaking sewers). This pathway is less relevant for those chemicals that adsorb on soil particles and more relevant for those that goes through porous media without significant retention.

The detailed description of these pathways is not introduced here as these equations are not developed further within this model application.

2.3.4 Industrial discharge

As for the other substances, industrial sites are handled as point source emitters. Emission is either defined by discharge · concentration method, or if discharge is missing, by estimated annual loads. For pharmaceutical compounds, only direct dischargers are considered.

2.3.5 Landfills

Landfill sites store disposed medicines that mainly contain the compounds included in the modelling study. Leachate of these sites may cause the contamination of groundwater that is transferred to surface water. These type of sites are treated as legacy pollution deposit in groundwater. There have been a few studies examining groundwater concentrations of selected pharmaceuticals. Buszka et al.[27]

[27] Buszka, P. M., Yeskis, D. J., Kolpin, D. W., Furlong, E. T., Zaugg, S. D., & Meyer, M. T. (2009). Waste-Indicator and Pharmaceutical Compounds in Landfill-Leachate-Affected Ground Water near Elkhart, Indiana, 2000–2002. *Bulletin of Environmental Contamination and Toxicology*, *82*(6), 653–659. https://doi.org/10.1007/s00128-009-9702-z

found above LOQ concentration for pain killers (Ibuprofen). Other studies reported pharmaceutical concentrations in leachate effluents[28,29]. Retention in groundwater is not considered. Emissions are calculated using a concentration x discharge approach:

LF_E_PHAR = LF_CONC_PHAR · LF_ps_Q
where
LF_CONC_PHAR is the concentration of PHAR in the leachate,
LF_ps_Q is the discharge of the given landfill sites. It is assumed that the discharge to the stream equal the recharge amount to the groundwater.
The *ps* notation shows, that it is considered as a point source in the model.
LF_ps_Q = BI_A_LF · (BI_PREC_LF - AD_EVAPO_LF)
where
BI_PREC_LF is the annual precipitation at the landfill sites,

AD_EVAPO_LF is the annual evapotranspiration at the landfill sites (see details at the description of landfill sites for PFAS substances.)

2.4 PATHWAYS FOR PFAS COMPOUNDS

2.4.1 Industries

There are a very wide range of activities that emit PFAS compounds into the environment, either by direct emission to water, air, soil, or by indirect emission through groundwater (leeching sewers, landfill sites etc.). In our current work we have selected those industries that emit PFOA and/or PFOS compounds. This has been taken from an extensive overview study (Glüge et al., 2020). Below we list those activities that we accounted for (Table 1).

Table 1 - Review of industrial sectors that use and emit PFOS and/or PFOA substance

Category	BREF	Activity code (IED)	Description of activity	Chemical involved (type of use)
Oil refineries	REF	1.2	Refining of mineral oil and gas	PFOA (Chemical driven oil production: Increase the effective permeability of the formation, Foaming agent for fracturing subterranean formations, Heavy crude oil well polymer blocking remover)
Production of organic chemicals	OFC, LVOC, CWW	4.1.(a)	Chemical installations for the production on an industrial scale of basic organic chemicals	PFOS, PFOA (Lubricants, sealants, adhesives, brake fluids, etc.)
Surfactants		4.1. (k)k	Production of organic chemicals: surface-active agents and surfactants	PFOS, PFOA (surface-active agents and surfactants)

[28] Rezaei Adaryani, A., & Keen, O. (2022). Occurrence of pharmaceuticals and plasticizers in leachate from municipal landfills of different age. Waste *Management*, *141*, 1–7. https://doi.org/https://doi.org/10.1016/j.wasman.2022.01.023

[29] Yu, X., Lyu, S., Zhao, W., Guo, C., Xu, J., & Sui, Q. (2024). A picture of pharmaceutical pollution in landfill leachates: Occurrence, regional differences and influencing factors. *Waste Management*, *184*, 20–27. https://doi.org/https://doi.org/10.1016/j.wasman.2024.05.019

Category	BREF	Activity code (IED)	Description of activity	Chemical involved (type of use)
Polymer production	POL	4.1.(h)	Production of organic chemicals: plastic materials (polymers, synthetic fibres and cellulose-based fibres)	PFOS (rubber, antistat in rubber and plastic),
Electroplating of metal	STM &G	2.3.(c)	Processing of ferrous metals: application of protective fused metal coats	PFOS (Chrome plating)
Surface treatment of metals and plastics		2.6	Surface treatment of metals or plastic materials using an electrolytic or chemical process where the volume of the treatment vats exceeds 30 m^3	PFOS (Cr plating)
Paper mills	PP	6.1. (b)	Production in industrial installations of paper or card board	PFOS, PFOA (Wood based building materials, particle boards, in adhesive resins)
Textile processing units	TXT	6.2	Pre-treatment (operations such as washing, bleaching, mercerisation) or dyeing of textile fibres or textiles	PFOS, PFOA
Pesticide production facilities	OFC	4.4	Production of plant protection products or of biocides	PFOA (general use)
Surface treatments sites	STS	6.7	Surface treatment of substances, objects or products using organic solvents, in particular for dressing, printing, coating, degreasing, waterproofing, sizing, painting, cleaning or impregnating, with an organic solvent consumption capacity of more than 1	PFOA (paints, coatings)
Waste incineration	WI	5.2	**Disposal or recovery** of waste in waste (co-) **incineration**	PFOS, PFOA
Disposal or recovery of hazardous waste	WT	5.1	Disposal or recovery of hazardous waste	PFOS, PFOA

In the current model version all industrial units that are in the model will be taken into account as a point source emission and the method of emission calculation is following the concentration times discharge approach.

ID_ps_E_PFAS = ID_ps_Q * ID_CONC_PFAS/1000/1000*(100-MM_PS_PFAS)/100

Where ID_ps_E_PFAS is the industrial point source emission (kg/year), ID_ps_Q is the point source discharge (m^3/s), ID_CONC_PFAS is the median sectoral PFAS concentration of the industrial units and MM_PS_PFAS is the reduction measure for industrial emissions.

Concentration data for emissions are taken from the database of the Promisces project[30], where three variants for all the most important PFAS substances has been derived from an Austrian database of PFAS emissions. It is important to note that there is very limited data available for this pathway, and further data on industrial emissions should be gathered. There would be a need to diversify concentrations to the above listed industrial sectors.

Industrial discharge is taken from the Tethys database, which hase been constructed from national data for most DRB cointries including Slovakia, Serbia, Slovenia, Hungary and the DHm3c database.

To estimate total loads from direct discharges, these PS loads are aggregated.

ID_E_PFAS = more_psaggrau(ID_ps_E_PFAS)

Whew ID_E_PFAS is the emission of PFAS from industrial units within an AU, the more_psaggrau is the aggregation mechanism of the MoRE model that summes up all point source units for each AUs.

2.4.2 Municipal waste water treatment plants

Waste water treatment plant discharges are handled as point source emissions in the model. This means that metadata has to be supplied to each plant that is implemented in the model, including data on commissioning date, discharge point location, sector (municipal or industrial), etc. Exact location is not used in the current model version, the point source loads are aggregated to the AUs by the built-in algorithm of the MoRE model. Loads are calculated using two approaches: 1. Concentration * discharge and 2. Specific emission factors of inhabitants * number of inhabitants.

This pathway describes direct emissions from waste water treatment plants to receiving water bodies. The method includes the estimation of the annual PFAS loads from each individual waste water treatment plant in the first step:

The conceptual equation is

$$WWTP\_ps\_E\_PFAS = WWTP\_ps\_Q * WWTP\_CONC\_PFAS$$

Where

WWTP_ps_E is the point source load from the individual treatment plants to the receiving water body.

WWTP_ps_Q is the discharge of the individual treatment plants.

WWTP_CONC_PFAS_ is the concentration of the given PFAS in the treatment plant effluent.

The overall equation taking into account treatment stages is as follows:

WWTP_ps_E_PFAS = if(WWTP_treatment_type=0, WWTP_ps_Q*WWTP_CONC_NOTREAT_PFAS/1000000, if(WWTP_treatment_type = 1, WWTP_ps_Q*WWTP_CONC_PRIM_PFAS/1000000, if(WWTP_treatment_type = 2, WWTP_ps_Q*WWTP_CONC_SEC_PFAS/1000000, if(WWTP_treatment_type = 3, WWTP_ps_Q*WWTP_CONC_TERT_PFAS/1000000, WWTP_ps_Q*WWTP_CONC_QUART_PFAS/1000000))))

The concentration values for the primary, secondary, tertiary and quaternary treatment have been derived from available effluent concentration data (see section 4.1).

There is an alternative way of calculation waste water loads, based on treatment plant capacities:

[30] Liu, M. (2024). Per- and Polyfluoroalkyl Substances (PFAS) Concentrations in the Upper Danube Catchment: Integrated Dataset from H2020 Project PROMISCES - Case Study 2 (2.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14994113

WWTP_ps_E_PFAS = if(WWTP_ps_PE<2000,WWTP_ps_Q * WWTP_CONC_WWTP_0to2kEW_PFAS/1000000,if(WWTP_ps_PE<5000,WWTP_ps_Q * WWTP_CONC_WWTP_2to5kEW_PFAS/1000000,if(WWTP_ps_PE<10000,WWTP_ps_Q * WWTP_CONC_WWTP_5to10kEW_PFAS/1000000,if(WWTP_ps_PE<100000,WWTP_ps_Q * WWTP_CONC_WWTP_10to100kEW_PFAS/1000000,WWTP_ps_Q * WWTP_CONC_WWTP_100kEW-plus_PFAS/1000000))

Where WWTP_lowerPE_to_upperPE refers to the population equivalent of the given treatment plant. Effluent concentration of these categories have been derived as described in section 4.1.

In the second step the emissions are aggregated for the analytical units by summing up all the treatment plants discharging to the main receiving rivers within the AU:

$$WWTP\_E\_HM = \sum_{i=1}^{n} WWTP\_ps\_E\_PFAS_i$$

Where

WWTP_E_PFAS is the PFAS emission for a given analytical unit

$WWTP\_ps\_E\_PFAS_i$ is the point source emissions from all the treatment plants (i) in the AU

Small treatment plants are estimated in the model based on an expert judgement of total flows for the given AU based on population number and the connected population of the actual treatment plants in the AU.

WWTP_small_E_PFAS = WWTP_s_CONC_PFAS * WWTP_small_Q / 1000 / 1000

Where

WWTP_s_CONC_PFAS is the PFAS concentration of the effluents from small WWTP-s. This is also obtained from observation data. For PFOA this is a constant value of 0.03 µg/l for the DRB and for PFOS this is 0.05 µg/l for the entire DRB.

2.4.3 Urban pathways

There are four distinct pathways related to Urban runoff. The methodology used in the current model version is based on the method used in the Danube Hazard m^3c and PROMISCES projects. For the description of all variables, refer to *Annex III.*

Emission through **CSOs**

Urban wash-off that is going down in combined systems. Defined as follows:

US_cso_E_PFAS = US_cso_Q · US_cso_CONC_PFAS · (100-MM_PS_PFAS) / 100

Please not that flows are calculated from three different pathways: impervious areas, inhabitant specific runoff and from commercial areas.

Emission through separated sewer systems (stormwater runoff)

Urban wash-off that is going down the storm sewer network. Defined as follows:

US_E_SS = US_SS_COC_PFAS*US_SS_Q

Apart from the emission loads from runoff, fine sediment (FN) emission is also estimated, in order to differentiate between particulate and dissolved runoff.

US_ss_E_FS = US_ss_CONC_FS · (1- (MM_US_ss_EFF_FS /100)) · US_ss_Q

Where the runoff in this pathway is calculated as:

US_ss_Q = US_ss_A_IMP · IM_PREC_yr · US_RC

Emission from areas and population that is **only connected to sewer systems**:

In many cities of the Danube Basin the collected sewage is not treated, but directly discharged to the rivers, this is therefore a separate pathway:

US_oss_E_IMP_PFAS = US_oss_Q_IMP ·US_ss_CONC_PFAS · (100-MM_PS_PFAS) / 100

US_oss_E_INH_PFAS = US_INHL_PFAS · IM_INH_oss · (100-MM_PS_PFAS) / 100

US_oss_E_PFAS = US_oss_E_IMP_PFAS + US_oss_E_INH_PFAS

US_oss_GW_E_PFAS = (US_oss_E_IMP_PFAS + US_oss_E_INH_PFAS) · (1 - US_LEAK_oss)

Emission from areas and population that is **not connected to sewer systems**:

In many cities and especially in rural areas/villages of the Danube Basin the generated sewage is not collected and either not treated, or treated by delivering to WWTP by road transport. If it is not treated, then it infiltrates to the soil and groundwater, if treated then it either treated locally or in decentralized plants or in most cases delivered to the WWTP nearby. This pathway contains three sub-pathways, runoff from impervious areas (1), inhabitant specific runoff (2) and through groundwater(3) (Figure 3).

US_nss_E_INH_HM = if (BI_CODE_coun<=99, 0, IM_INH_nss / 1000 · US_INHL_PFAS (1 - US_nss_SHR_inhl_towwtp_sept / 100))

US_nss_E_IMP = US_nss_Q_IMP · US_CONC_ROAD_PFAS

US_nss_GW_E_PFAS = US_nss_E_INH_PFAS · (1 - US_nss_RET_SOIL_PFAS/100) · (100-MM_PS_PFAS) / 100

2.4.4 Atmospheric deposition

Atmospheric deposition is given by the sum of two sources:

1. There is a background value, which is determined from the extrapolation of monitoring results carried out for seven pilot sites within the Danube Hazard project. The description of the process and the final values can be found in chapter 4.
2. There are hot spot emissions in the vicinity of known aerial sources, such as fluorochemical plants, oil refinery emissions, or waste incineration plants. The value of the emitted load comes from the industrial site emissions. The impact area of the emissions are determined by the following assumptions:
    a. chimney emissions has a concentric emission area,
    b. the impact area has 50 km wide diameter[31,32,33] ,
    c. emission within the emission area is homogeneous,
    d. AU emission is estimated by the ratio of the impacted area divided by the AU area, and
    e. long range transport is neglected.

[31] Galloway, J. E., Moreno, A. V. P., Lindstrom, A. B., Strynar, M. J., Newton, S., May, A. A., & Weavers, L. K. (2020). Evidence of Air Dispersion: HFPO–DA and PFOA in Ohio and West Virginia Surface Water and Soil near a Fluoropolymer Production Facility. *Environmental Science & Technology*, *54*(12), 7175–7184. https://doi.org/10.1021/acs.est.9b07384

[32] Hangen, E., Klemm, A., Kronawitter, H., & Schubert, A. (2010). Perfluorooctanoate (PFO) in Forest Soils near a Fluoropolymer Manufacturing Facility. *Water, Air, & Soil Pollution*, *212*(1), 491–499. https://doi.org/10.1007/s11270-010-0365-5

[33] Moghadasi, R., Mumberg, T., & Wanner, P. (2023). Spatial Prediction of Concentrations of Per- and Polyfluoroalkyl Substances (PFAS) in European Soils. *Environmental Science & Technology Letters*, *10*(11), 1125–1129. https://doi.org/10.1021/acs.estlett.3c00633

2.4.5 Erosion

Erosion is not a dominant pathway of PFAS emissions, but as PFAS has been found in all soils around the DRB during the pilot monitoring campaign of the Danube Hazard m^{3}c project (Kardos et al., 2024) it may be a significant pathway, therefore we account for it in the model. Soil PFAS comes from multiple sources, including atmospheric deposition as background diffuse contamination and also via local emissions from fertilizer use or pesticide use (Glüge et al., 2020). In this model we account for the soil contamination using a linear regression of soil PFAS concentration and GDP as described in (Kardos et al., 2024)[34].

The calculation of erosion is the same as for heavy metals, except that erosion from glaciers or mountainous areas are neglected, due to lack of data.

Erosion for agricultural lands is calculated by the following formula (Original Moneris):

ER_agrl_E_HM = [ER_agrl_SL_AL · ER_agrl_CONT_SOIL_top_AL_PFAS · (1 - MM_ER_EFF_AL_SED /100) + ER_agrl_SL_PST · ER_agrl_CONT_SOIL_top_PST_PFAS]· ER_SDR · 0.01 · ER_ENR_AL /1000

Where:

ER_agrl_SL_AL is the soil loss from arable land,

ER_agrl_CONT_SOIL_top_AL_PFAS is the PFAS content of the soil in the arable lands in the given analytical unit,
MM_ER_EFF_AL_SED is the efficiency of the erosion reduction measure introduced in the model (if applicable),
ER_agrl_SL_PST is the soil loss from pastures,

ER_agrl_CONT_SOIL_top_PST_PFAS is the PFAS content of the soil in pastures in the given analytical unit,
ER_SDR is the sediment delivery ratio, which is calculated according to the original method defined in the Moneris model[7],
ER_ENR_AL (-) is the enrichment ratio, which is calculated according to the original method defined in the Moneris model[7]:

ER_ENR_HM = if (ER_E_spec_AGRL_SED < 1, then **ER_FCT_a_ENR_HM**;
if (ER_FCT_a_ENR_HM * ER_E_spec_AGRL_SED ^ ER_EXP_ENR_HM < 1, then 1;
otherwise ER_FCT_a_ENR_HM · ER_E_spec_AGRL_SED ^ ER_EXP_ENR_HM)).

2.4.6 Surface runoff

Surface runoff is taken into account as wash off from soils and natural surfaces also contain PFAS in small concentrations. The concentration value is taken from earlier model applications (EU Promisces report[22]). It is quite evident that there is a strong spatial heterogeneity of these concentrations, but there is no direct evidence of this in the literature. Variation in runoff, however, can be linked to the variation in soil concentrations. A direct linear relationship is used to scale the concentration variability. The relative difference of soil PFAS concentrations of a given AU to a reference AU is taken into account. The reference AU is the AU, where surface runoff measurements were delivered.

[34] Kardos, M. K., Clement, A., Jolánkai, Z., & Zessner, M. (2024). Development and testing of an efficient micropollutant monitoring strategy across a large watershed Lessons learned from a Danube Basin-wide project. *Science of the Total Environment*, *948*(174760). https://doi.org/10.1016/j.scitotenv.2024.174760

Transport of most PFAS substances is primarily happening in the dissolved water phase but particulate transport is also present. The Adsorption of PFAS compounds is measurable on various surfaces, including soil organic matter[35] and inorganic particles[36]. K_d and K_{oc} values are also available and described in Table 25.

2.4.7 Landfill sites

Landfill sites store a large variety of materials, products that contain PFAS compounds as these compounds are widely used in clothes and other textile wares, surface treatment of plastic and metal products, sanitary products etc. Leachate of these sites cause the contamination of groundwater that is transferred to surface water. This type of sites are treated as legacy pollution deposit in groundwater. Concentration values are available for leachate water from several sites in Europe (Greece[37], Upper Danube[38]).

Retention in groundwater is not considered. Emissions are calculated using a concentration x discharge approach:

$$LF\_E\_PFAS = LF\_CONC\_PFAS \cdot LF\_ps\_Q$$

where

LF_CONC_PFAS is the concentration of PFAS in the leachate,

LF_ps_Q is the discharge of the given landfill sites.

The *ps* notation shows, that it is considered as a point source in the model.

$$LF\_ps\_Q = BI\_A\_LF \cdot (BI\_PREC\_LF - AD\_EVAPO\_LF)$$

where

BI_PREC_LF is the annual precipitation at the landfill sites,

AD_EVAPO_LF is the annual evapotranspiration at the landfill sites.

Values for these two variables have been calculated from annual precipitation (Figure 4) and evapotranspiration (Figure 5) grid time series that are the output of the Danube Water Balance model. In this model, daily discharges have been calibrated for the entire Danube Basin with satisfactory results, therefore both precipitation and evapotranspiration can be considered as a reliable input.

2.4.8 Aerodromes and firefighting centers

Aerodromes and firefighting training centers use aqueous film forming foams (AFFF) to extinguish hydrocarbon fires. Other hot spots are those areas where large fires have been extinguished with these materials. These foams contain a long list of PFAS compounds, including PFOA and PFOS. At these sites, the infiltrated water washes the PFAS compounds into the groundwater, where it accumulates. Concentrations in groundwater are available in many sites across Europe (Forever Map database, Le

[35] Higgins, C. P., & Luthy, R. G. (2006). Sorption of Perfluorinated Surfactants on Sediments. *Environ. Sci. Technol.*, *40*(23), 7251–7256. https://doi.org/10.1021/es061000n

[36] Zhang, R., Yan, W., & Jing, C. (2015). Experimental and molecular dynamic simulation study of perfluorooctane sulfonate adsorption on soil and sediment components. *Journal of Environmental Sciences*, *29*, 131–138. https://doi.org/https://doi.org/10.1016/j.jes.2014.11.001

[37] Nika, M. C., Ntaiou, K., Elytis, K., Thomaidi, V. S., Gatidou, G., Kalantzi, O. I., Thomaidis, N. S., & Stasinakis, A. S. (2020). Wide-scope target analysis of emerging contaminants in landfill leachates and risk assessment using Risk Quotient methodology. *Journal of Hazardous Materials*, *394*, 122493. https://doi.org/https://doi.org/10.1016/j.jhazmat.2020.122493

[38] Liu, M., Saracevic, E., Oudega, T. J., Obeid, A. A. A., Nagy-Kovács, Z., László, B., Kittlaus, S., Zoboli, O., Krampe, J., Derx, J., & Zessner, M. (2025). Investigating the extent of PFAS contamination in the Upper Danube Basin across environmental compartments. *Environmental Sciences Europe*, *37*(1), 99. https://doi.org/10.1186/s12302-025-01141-6

Monde), but not many measurements are available in the Danube River Basin. To account for these hot spots, average concentration is derived from available data (Table 2).

AA_E_PFAS = AA_CONC_PFAS · AA_ps_Q

Discharge is calculated based on annual water balance difference:
AA_ps_Q = BI_A_AA · (BI_PREC_AA - AD_EVAPO_AA)
where
BI_PREC_AA is the **annual precipitation** at the aerodromes and firefighting training centers,
AD_EVAPO_AA is the **annual evapotranspiration** at the aerodromes and firefighting training centers.

Table 2 – Concentration values used in the model for AFFF contamination sites

variable	Value (ng/l)	variant	name of input data set
AA_CONC_PFAS_PFOA	25.39	1	PROMISCES monitoring and IWR inventory 2024
AA_CONC_PFAS_PFOA	6.46	2	PROMISCES monitoring and IWR inventory 2024
AA_CONC_PFAS_PFOA	67.7	3	PROMISCES monitoring and IWR inventory 2024
AA_CONC_PFAS_PFOS	261.57	1	PROMISCES monitoring and IWR inventory 2024
AA_CONC_PFAS_PFOS	19.48	2	PROMISCES monitoring and IWR inventory 2024
AA_CONC_PFAS_PFOS	1307.86	3	PROMISCES monitoring and IWR inventory 2024

2.4.9 Other legacy hot spots

This pathway is created to account for sites, where known elevated groundwater contamination is present. One well-known site is the Gendorf site, a PFAS manufacturing site, which does emit PFAS through direct discharge and aerial discharge. According to research studies at other sites[39] and also according to the Forever Map database[40], such sites may have sever groundwater contamination due to its past constant areal discharge. Other Eurpoean sites, where this is supported by are Miteni, and Cerazzo, Sarego, IT, Zwijndrecht, BL, where groundwater PFAS levels are in the range of 1000 to 40000 ng/l for the sum of PFOS and PFOA. There is no direct evidence of the presence of such high concentrations at sites in the Danube RB, but for the handling of likely contamination sites this pathway is included in the model.

LHSG _E_PFAS = LHSG _CONC_PFAS * LHSG _ps_Q

where

LHSG_ps_Q = BI_A_LHSG * (BI_PREC_ LHSG - AD_EVAPO_ LHSG)

where

BI_PREC_ LHSG is the **annual precipitation** at the legacy hot spots areas,

[39] Pétré, M.-A., Genereux, D. P., Koropeckyj-Cox, L., Knappe, D. R. U., Duboscq, S., Gilmore, T. E., & Hopkins, Z. R. (2021). Per- and Polyfluoroalkyl Substance (PFAS) Transport from Groundwater to Streams near a PFAS Manufacturing Facility in North Carolina, USA. Environmental Science & Technology, 55(9), 5848–5856. https://doi.org/10.1021/acs.est.0c07978

[40] Forever Pollution Project (2023). The Forever Pollution Map: Mapping PFAS contamination in Europe. Available at: https://foreverpollution.eu/ (accessed: 10 05 2025).

AD_EVAPO_ LHSG is the **annual evapotranspiration** at the legacy hot spots areas.

2.4.10 Groundwater

The groundwater pathway can be differentiated between several hot spots and diffuse loads. Diffuse load is associated with PFAS concentration in groundwater due to leaching of PFAS compounds from soils. The distribution of PFAS concentrations in the Danube basin is related in the model to soil concentrations, except in those areas where groundwater concentrations monitoring data is available. In this section, only the groundwater emissions from diffuse source are described. Hot spot type emissions are described in separate sections (Industrial, Landfill sites, aerodromes).

$$GW\_E\_PFAS = GW\_Q \cdot GW\_CONC\_PFAS \cdot (100\text{-}MM\_DI\_PFAS)$$

where

$$GW\_Q = (GW\_Q\_spec \cdot GW\_A\_rech)$$

Where GW_Q_spec is calculated based on water balance methods applied. If baseflow is coming from the Danube water balance model, the GW_Q_scpec directly calculated form the modelled baseflow, if not, then it is calculated by subtracting all the water balance components from the net total runoff for the given AU.

2.4.11 Tile drainage

In former applications, the tile drainage was included for nutrients but not for Metals or other compounds. The reason for this is the poor knowledge of concentration ranges in this pathway; only a few studies examined this in detail. In our approach, we state that each pathway should be included in the model and try to give the right order of magnitude for the emissions.

The starting point of the analysis is that soil concentrations of PFOA and PFOS can be estimated across the DRB, which can serve as a source for tile drainage load. Concentrations can be estimated using distribution coefficients between water and soil organic carbon (K_{oc}). K_{oc} in soils can be attained from a few studies (section 4.4, Table 17), while the organic fraction content of soils is also available (See Section 4.7.2).

Based on concentrations, the annual loads can be calculated by multiplying the estimated annual discharge (see Annex II for the algorithms of water balance, or O2.3 Annex I for the description of the algorithm):

$$TD\_E\_PFAS = (TD\_Q \cdot TD\_CONC\_PFAS)$$

where,

TD_E_PFAS is the PFAS emission through tile drainage (t/y)

TD_Q is the discharge in the AU vie tile drainage (m^3/s), that is estimated by an empirical approach.

TD_CONC_PFAS is PFAS concentration in tile drain flow (ug/l)

3 PROCESSING OF AVAILABLE INPUT DATA

3.1 ANALYTICAL UNITS (AU) IN THE MODEL

The delineation of the AUs has been done in the latest model application of the Moneris model, delivered by the IGB Berlin in 2021 for the RBMP of the Danube River Basin. The delineation has been reviewed and corrected in a few instances.

- Two units has been merged in Hungary, to remove a very small flatland AU.
- One Bulgarian catchment has been divided to two units, as these very in reality drain to two different AUs.

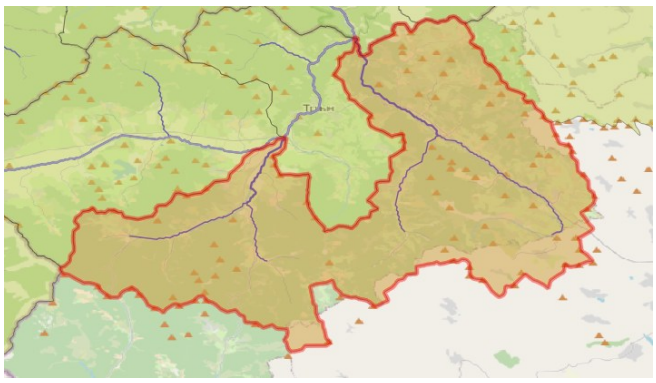

Figure 7 – Old delineation of a catchment in Bulgaria

- Three small AUs has been removed as those were not part of the Danube River Basin, but the Vistula Basin.
- The Croatian Karst basins has been reviewed by Croatian experts and were found to be correct.

Finalized AU map of the DRB:

The Danube River Basin contains 1727 units in total. The average area of the units is 468.6 km^2, while the median value is 322 km^2. There are quite a few very small AUs due to administrative boarders: 20 AUs below 10 km^2, AND 114 AUs below 50 km^2. The maximal AU size is almost 4600 km^2. The reason for this large heterogeneity in the AU sizes stems from the earlier modelling projects in the DRB. Some countries adapted higher resolution (AT, BG and MO).

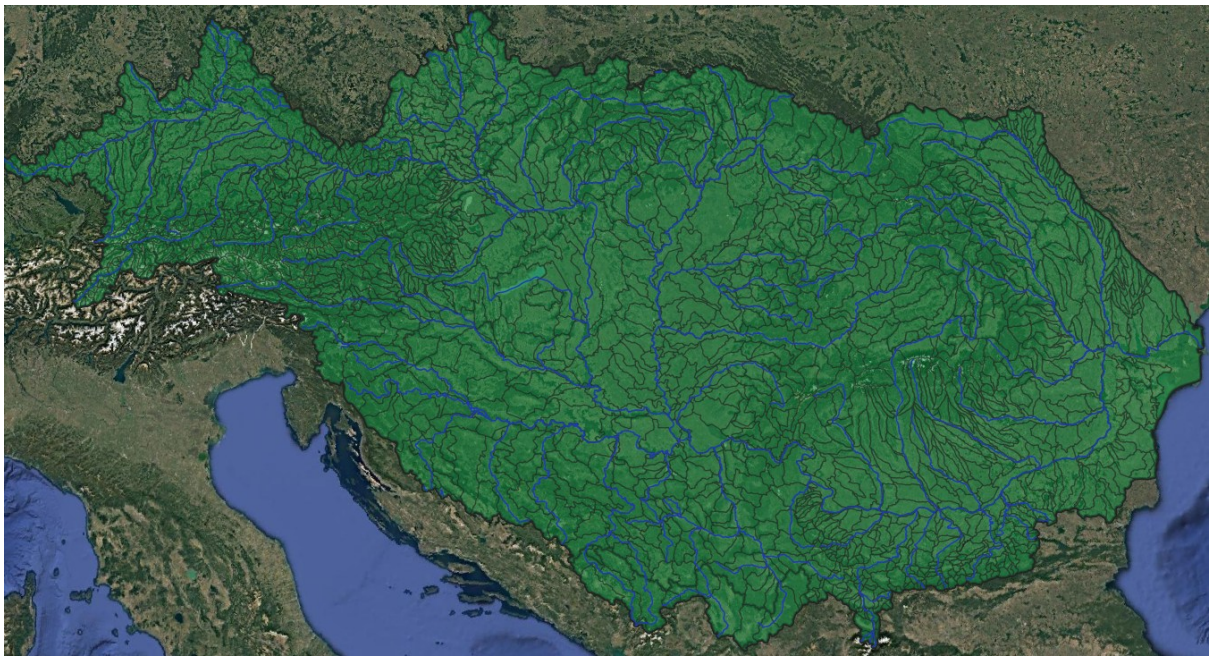

Figure 8 – Analytical units in the DRB for the MoRE model in the Tethys project

3.2 BASIC MODEL INPUT VARIABLES

The model has different types of variables and there are three major categories: constants, spatial data, and temporal data.

There is also a different approach of classifying the input data:

A. Data that describes the environment, such as topography, land use, and soil types.
B. Data that is related to the substances we model: concentrations, or stocks in environmental compartments, emission rates through pathways, etc.

The B datatype is described in chapter 4. In the current chapter, we only focus on type A of input data.

3.2.1 Moneris model datasets

Most of the generic input data, such as landuse, topography, etc. is taken from the database of the Moneris model, applied by IGB Berlin for the third DRBMP. The following table describes the datasets coming from this data source:

Table 2 - Model input variables taken from the Moneris nutrient model for the DRBMP in 2021

Analytical unit data	BI_A	area of analytical unit
	BI_ELEVA	mean elevation of the analytical unit
	BI_INH	number of inhabitants
	BI_PREC_yr_lt	mean long term average yearly sum of precipitation
	BI_SLP	mean slope in the Analytical unit
	BI_A_AL_slope_0_1	area of arable land with slope <1%
	BI_A_AL_slope_1_2	arable of arable land with slope 1-2%
	BI_A_AL_slope_2_4	arable of arable land with slope 2-4%
	BI_A_AL_slope_4_8	arable of arable land with slope 4-8%
	BI_A_AL_slope_8	area of arable land with slope > 8%
	BI_A_ALP	area alp (alpine pasture)

	BI_A_FOR	area of forest
	BI_A_GLC	area glaciers
	BI_A_O	area of open spaces
	BI_A_OPM	area of open mining
	BI_A_PST	area of grassland
	BI_A_URB	settlements
	BI_A_WL	area of wetlands
	BI_A_WS_sat	area of water surfaces from satellite data
	BI_A_WS_lakes_mr	area of lakes along main rivers
	BI_A_WS_mr	area of main rivers
	BI_A_WS_trib	area of tributaries
	RM_T_H2O	water temperature
Urban systems	US_SHR_a_cs_tss	proportion of the area drained by the combined system to the total area drained by the sewage system.
	US_cso_VOL_SOT	storage volume of stormwater overflow tanks in combined sewer systems
	IM_INH_con	number of inhabitants that are connected to sewer systems
	IM_INH_conWWTP	number of inhabitants that are connected to sewer systems and waste water treatment plants
	IM_INH_nss	number of inhabitants that are not connected to sewer systems
	IM_INH_oss	number of inhabitants that are only connected to sewer systems
	IM_INH_DCTP_GW	inhabitants connected to DCTP and discharging by groundwater recharge (soil aquifer treatment, SAT)
Erosion	ER_PREC_s_lt	summer precipitation, longterm average
	ER_agrl_SL_spec_lt_PST	soil loss from pastures, specific, longterm average
	ER_agrl_SL_pot_spec_AL_slp_0_1	potential soil loss from arable land, slope < 1%, specific
	ER_agrl_SL_pot_spec_AL_slp_1_2	potential soil loss from arable land, slope 1-2%, specific
	ER_agrl_SL_pot_spec_AL_slp_2_4	potential soil loss from arable land, slope 2-4%, specific
	ER_agrl_SL_pot_spec_AL_slp_4_8	potential soil loss from arable land, slope 4-8%, specific
	ER_agrl_SL_pot_spec_AL_slp_8	potential soil loss from arable land, slope > 8%, specific
	ER_nat_E_spec_lt_SED	Sediment input from naturally covered areas, specific, long-term average.
Atmo. Dep	AD_EVAPO_lt	evapotranspiration, longterm average
Tile drainage	TD_SHR_a_td_agrl	Share of drained area in the agricultural area

3.2.2 Open roads

Non-urban road surfaces are an important source of heavy metal emissions in heavily populated regions of the Basin. This spatial data had to be derived for the whole basin.

As an initial step, different datasets have been tested to see their applicability for the task. The testing process focused on specific regions of Hungary to ensure a localized and context-specific analysis. For this, different data sources were examined: the CORINE Land Cover dataset[41] , the OpenStreetMap

[41] https://land.copernicus.eu/en/products/corine-land-cover/clc2018

(OSM) dataset[42] , and the Hungarian NÖSZTÉP (Hungary Ecosystem Basemap) database[43] , which is the most detailed land cover dataset in Hungary. The primary disadvantage of relying on national databases is the need to integrate multiple data sources for multi-country research, which often leads to inconsistencies in data formats, resolution, and methodologies. Despite this limitation, if a national database is found to best meet the specific requirements of the study—offering superior accuracy, detail, or relevance—it can still be a valuable choice, provided that the inconsistencies can be managed through harmonization techniques or supplementary analyses. The objective of the analysis was to identify and extract area data for roads with the highest traffic volumes, as this information is critical for the model. Each dataset was examined for its ability to provide accurate and consistent road network data.

The CORINE Land Cover raster dataset contained the relevant data in the category 1.2.2 Road and rail networks and associated land. This category does not contain subcategories, so for example, railway lines, that are not currently relevant for us, are included in the data source, but when it was loaded into the QGIS software, it can be seen that it nevertheless has poor spatial coverage of certain road sections, even parts of motorways were missing. In Figure 9, the M7 motorway (going through the south-western part of Hungary) is shown in red by the underlying open street map data, while the black line shows only those part of the road, which appears in the CORINE dataset.

Figure 9 - CORINE Land Cover dataset 1.2.2. Road and rail networks and associated land category

The advantage of the OSM vector dataset is that it classifies dense road networks and roads into many categories; the disadvantage is that it is vector data, so road surface still has to be calculated by applying the relevant road widths to the lengths given by the map (Figure 10). This, however, results in a fairly accurate estimation. Due to the classified road categories, high-traffic road types can be distinguished, such as primary, secondary, residential, footway, etc.

[42] https://download.geofabrik.de/
[43] http://alapterkep.termeszetem.hu/

Figure 10 - Detail of the OSM dataset at Lake Balaton, Hungary

NÖSZTÉP is a raster type data, but the coverage of the road surfaces is inhomogeneous. Category 1210 includes paved roads. Figure 11 shows that the data is also incomplete, compared to the OSM database.

Figure 11 - Figure 3 NÖSZTÉP dataset

Based on the above analyses, the OSM dataset was selected as the primary data source. From the OSM data, roads classified under the motorway and primary categories were extracted for analysis. Motorways represent the highest-priority, high-capacity road infrastructure, engineered explicitly to facilitate efficient, fast, and safe long-distance travel. Primary roads, on the other hand, serve as the backbone of regional and national transport networks, typically connecting major cities, economic hubs, and key logistical points. While they may not offer the same level of access control as motorways, primary roads are critical for enabling high-volume traffic and facilitating connectivity between urban centres and surrounding areas.

Land use polygons classified as residential were extracted from the OSM dataset to exclude road segments within residential areas. These polygons were then overlaid on the road network, and intersecting sections were removed to isolate arterial roads outside residential areas, and then roads were merged within each analytical unit. The resulting highways (blue color) and arterial roads outside residential areas (red color) are shown in Figure 12.

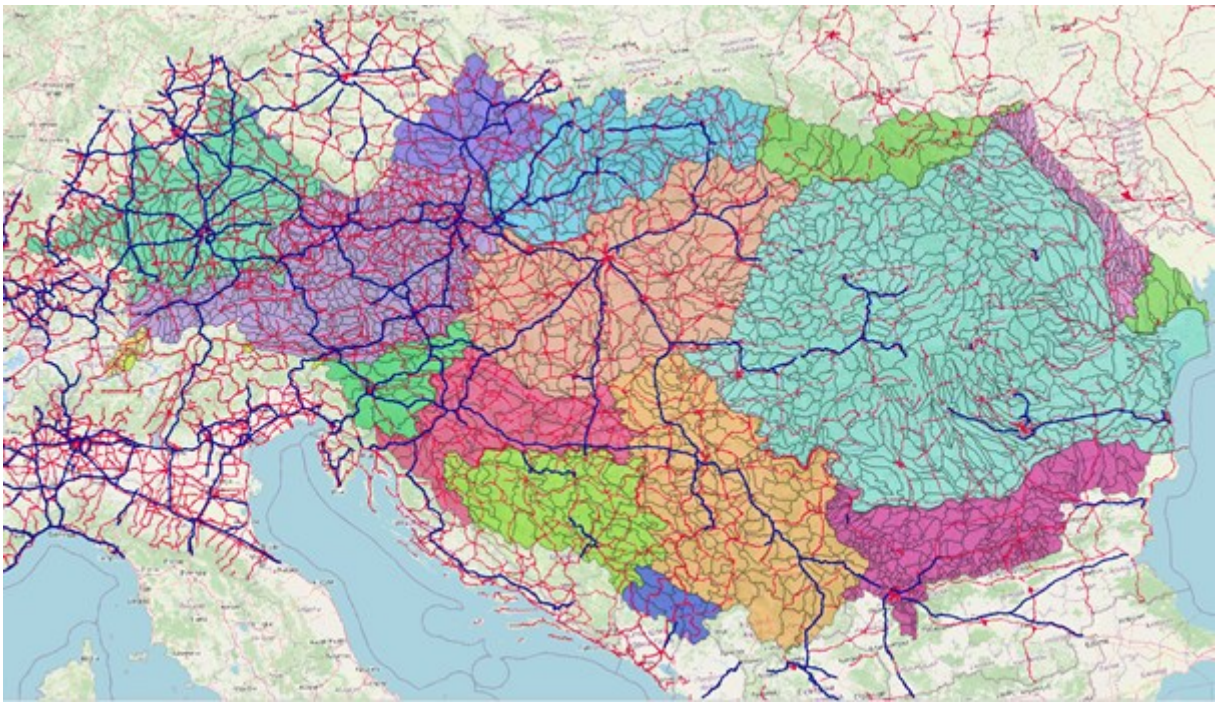

Figure 12 - Extracted non-urban roads (blue: motorway, red: primary road)

In addition to the derived road length, the average road width was also needed to calculate the total area. For motorways, the OSM dataset distinguishes between the two travel directions, so the area was calculated separately for each direction. This is not the case for primary roads, where the area was calculated for each road section. This approach ensures that the road areas accurately represent the road surfaces.

The formula used for motorways (for each analytical unit):

$$A_m = L_m \cdot W_m ,$$

where A_m is the calculated road area, L_m is the derived road length, and W_m is a constant for the average width for one direction (13.3 m).

The formula used for primary roads (for each analytical unit):

$$A_p = L_p \cdot W_p ,$$

where A_p is the calculated road area, L_p is the derived road length, and W_p is a constant for the average width of the road (25.6 m). The parts, data and dimensions of the roads are taken from the report of C. C. Schoon.[44]

The calculation steps of the road surface areas can be found on Figure 13.

[44] https://swov.nl/system/files/publication-downloads/r-94-07vii.pdf

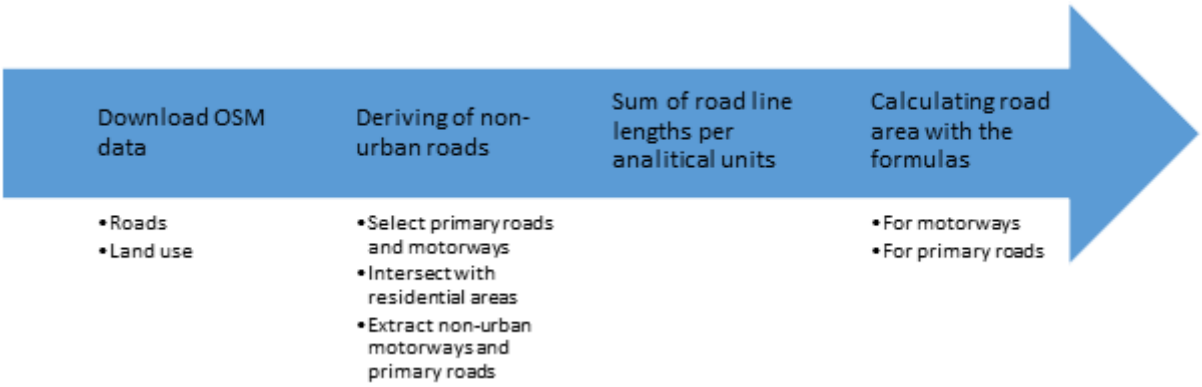

Figure 13 - Flowchart of the used methodology to derive open road surfaces

3.2.3 Population data

Population data at the level of analytical units is included among the model's input parameters. Two data sources were examined to calculate it.

Within the framework of the Copernicus program, the European Union monitors the Earth and its environment. Since 2021, the Copernicus Emergency Management Service has incorporated Exposure Mapping, designed to generate precise and up-to-date geospatial information on human settlements. A core element of this component is the Global Human Settlement Layer (GHSL), which provides highly accurate and continuously updated data on the presence of settlements and population, derived from satellite imagery and census data. https://human-settlement.emergency.copernicus.eu/ghs_pop.php Until 2018, we had the population figures assigned to the catchments, for this project we needed the 2019 and 2020 year data. This dataset is updated every five years, so we were able to download the 2020 data. The downloaded dataset was WGS84 projected at 30 arcsec resolution. Population data is provided as a spatial raster dataset with the population value in each cell.

The second dataset originates from Eurostat, which compiles data at the regional level within Europe based on the NUTS (Nomenclature of Territorial Units for Statistics) classification. The NUTS framework comprises four levels, with level 0 representing countries, followed by three hierarchical levels that further subdivide each country into progressively smaller territorial units. Among these, NUTS3 represents the lowest level and includes the smallest administrative units.

Each NUTS element in the dataset contains the total population for that particular region, alongside a vector-based shapefile (.shp) that defines the geometric boundaries of the NUTS regions. For the calculation, we utilized the 2021 dataset, which also incorporates data from previous years (2019 and 2020), as they were relevant for our analysis.[45]

Both datasets provide the total population corresponding to a specific geographic area, whether it is a raster cell or a NUTS 3 element. The initial step of the used method is calculating the population density, which is the population per unit area after the area data has been defined. After the population density has been calculated, the next step is to aggregate the values within the boundaries of the analytical units. This aggregation is carried out by summing the population density values per unit area within the boundaries of each analytical unit using zone statistics. The final result is the total population of each analytical unit.

[45]https://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/w/ks-gq-23-010
https://ec.europa.eu/eurostat/documents/3859598/20625712/KS-GQ-23-010-EN-N.pdf/d3a2b361-4e86-93c6-044f-dacbe2e4643a?version=1.0&t=1734081446125

3.2.4 Settlement data (sewerage)

Settlement data was taken from the MONERIS model and the ICPDR database. Databases were double checked with local project partners.

3.2.5 Erosion

Erosion data for agricultural soils is based on RUSLE 2015[46] data of the JRS ESDAC product, which was processed for the Moneris application in the RBPMP 2021. This data has been received and applied for the model; however, erosion data for the pastures and natural land has been revised as the originally received values were unrealistically high. The new values were gathered from two data sources: for EU countries, the RUSLE 2015 data has been used, while for non-EU countries the GLOSEM high resolution erosion database[47] was used. Example for the estimates for a pasture site is shown in Figure 11. The estimated erosion has many ”weak points” as quite frequently agricultural tables are classified as pastures by the Corine land use map. Pasture C factor values are generally low (below 1), despite these erosion values of 30-50 t/ha/y are shown, which might show that these were not estimated as pasture but as agricultural lands. Figure 14corine land use map.

[46] Panagos, P., Borrelli, P., Poesen, J., Ballabio, C., Lugato, E., Meusburger, K., Montanarella, L., Alewell, .C. 2015. The new assessment of soil loss by water erosion in Europe. Environmental Science & Policy. 54: 438-447. DOI: 10.1016/j.envsci.2015.08.012

[47] Borrelli, P., Ballabio, C., Yang, J.E. et al. GloSEM: High-resolution global estimates of present and future soil displacement in croplands by water erosion. Sci Data 9, 406 (2022). https://doi.org/10.1038/s41597-022-01489-x

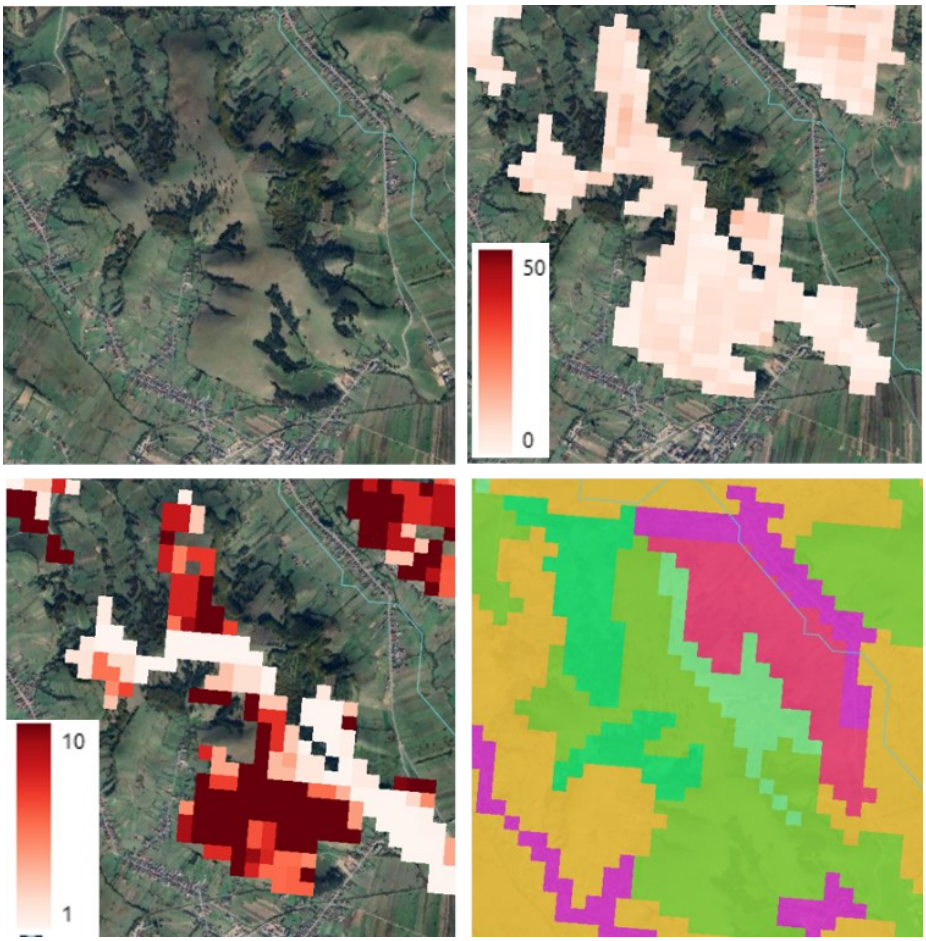

Figure 14 - Example location with pasture in Romania. Upper right estimation shows RUSLE 2015 estimate, lower left figure shows GloSEM estimate. Lower right masp shows the CLC2018 land use map, where dark green area shows pastures.

Estimates for natural areas are similar, natural land cover areas have been selected based on the Corine 2018 maps, which were then overlapped with the soil erosion estimate maps, producing erosion for the natural areas. These maps were then used to derive mean soil loss values in each analytical units using zonal statistics functions of QGIS software[48]. Examples of the resulting maps are shown on Figure 15 and Figure 16.

[48] "QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation Project

Figure 15 - GloSEM erosion estimates for the Western Balkan region. High erosion values are shown in many hilly places.

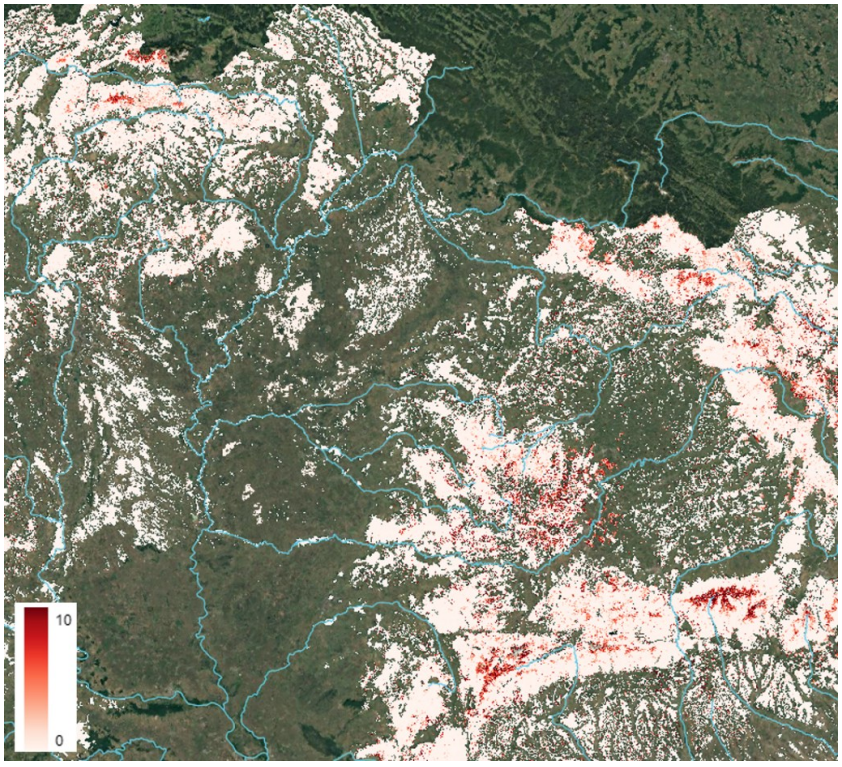

Figure 16 - Natural soil erosion according to the RUSLE 2015 model in the Tisza basin. High values are shown in open spaces (primarily high mountains).

3.3 POINT SOURCE DATA

3.3.1 Municipal waste water treatment plants

3.3.1.1 Available databases

Two kinds of sources were used to set up the wastewater treatment plant database for the model: (a) international and (b) national databases. Regarding (a), two international databases: the **UWWTD** database of the EEA, and the **ICPDR WWTP database** were merged. For EU countries, data from the UWWTD; for non-EU countries, data from the ICPDR database was applied. The merged database was compared again (b)-type national databases (where available).

Input data includes the following data:

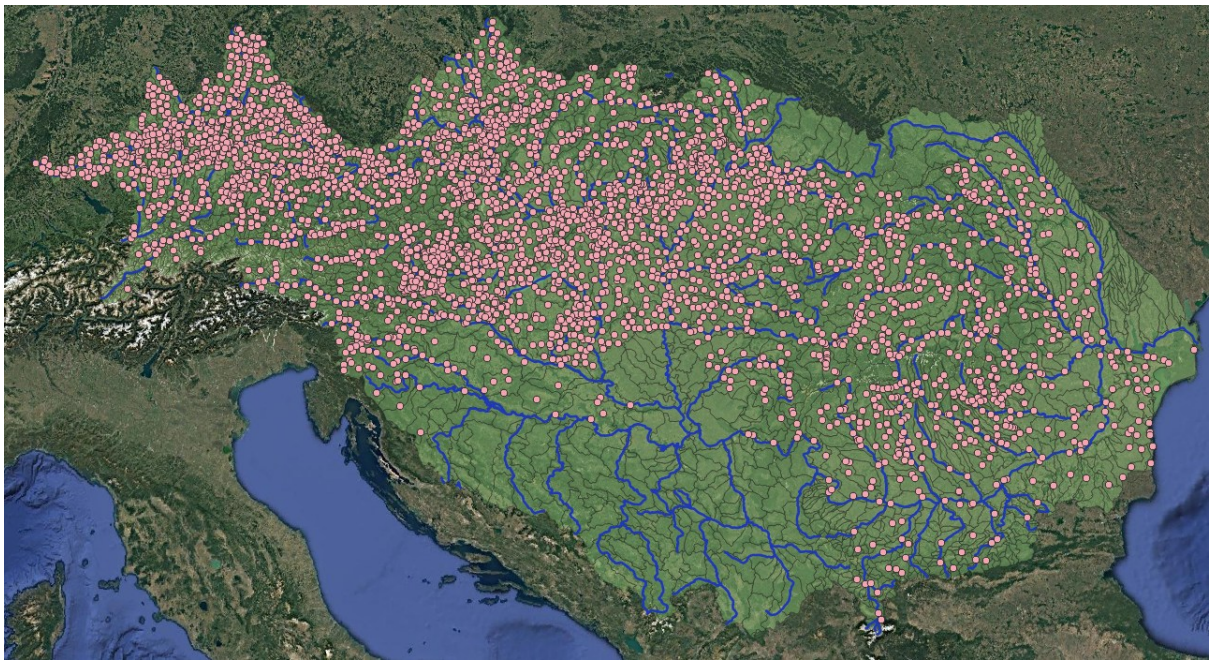

Figure 17 - Municipal and industrial WWTPs reported in the UWWTD database

Connection data

The connection data of the population is primarily available in water works operators, and in country databases. The secondary data source for the DRB is the ICPDR, which has collected these data for the earlier modelling work (years 2016-2018, varying country by country).

3.3.2 Industrial facilities

Two types of data source can be identified: (1) EU databases, comprehensive national reports on larger industrial facilities; and (2) national databases for non-EU members. In order to get the most accurate picture, we asked our partners for data reconciliation and data supplementation. In the data request, we published the data we had available (from EU reports) and asked for it to be supplemented (e.g. with smaller sites) if possible in order to score the model.

3.3.2.1 Available databases

According to national legislation, industrial facilities have a reporting obligation to the national administrative bodies and in some cases, they also have an EU data reporting obligation. The measurements

are (mostly) carried out on the basis of self-monitoring, i.e. their reliability is questionable, but at least we have measurement results that can inform the calculations of the industrial emission pathways of the MoRE model. However, the type and amount of measured and recorded data varies greatly from country to country. The following presents the accessible and used databases. If a country was not presented, we did not receive a national database from them, we could only use the E-PRTR data if the country reported there.

3.3.2.1.1 EU database

Name and availability and time scale: E-PRTR database: 'Industrial Reporting under the Industrial Emissions Directive 2010/75/EU and European Pollutant Release and Transfer Register Regulation (EC) No 166/2006 - ver. 10.0 Dec 2023 (Tabular data)' database. The used database short file name is *'Industrial_dataset_v10_December_2023.accdb'*. This can be downloaded from European Environment Agency (EEA) website[49][50].

This metadata refers to the geospatial dataset representing the status of the EEA Industrial Reporting database as of 15 December 2023 (version 10). The release and emissions data cover the period 2007-2022 as result of the data reported under the new E-PRTR facilities, 2017-2022 for IED installations and WI/co-WIs, and 2016-2022 for LCPs (Large Combustion Plans). These data are reported to EEA under Industrial Emissions Directive (IED) 2010/75/EU Commission Implementing Decision 2018/1135 and the European Pollutant Release and Transfer Register (E-PRTR) Regulation (EC) No 166/2006 Commission Implementing Decision 2019/1741. The dataset brings together data formerly reported separately under E-PRTR Regulation Art.7 and under IED Art.72. Additional reporting requirements under the IED are also included.

Spatial scale: The E-PRTR (and IED) databases are well harmonized across the EU, and even some non-EU countries report here.

Reporting scale: The database can be divided into two parts: 1) basic data on the most important industrial facilities that carry out activities listed in Annex I to Regulation 2006/61/EC (above a given production capacity/volume) and have direct emissions to water, soil or air; 2) annual emission load to water, soil or air, if the amount exceeds the reporting thresholds in Annex II to Regulation 2006/61/EC.

Disadvantages, problems: The E-PRTR database is the basis of the MoRE model database. Its biggest flaw is the high level of the reporting threshold, as it is clear from the national databases that many measurements of emissions are available, but they fall below the E-PRTR reporting threshold, so they are not included in the EU database. Therefore, the number of reported loads is extremely small, only the loads of the largest facilities are included in the database. Furthermore, the database does not contain discharge flow (Q) data. Another disadvantage is that the reports are made on an annual basis and the identification codes change many times (for company changes or other administrative reasons), which makes duplicate removal difficult. Furthermore, the database summarizes emissions at the facility level and not at the level of the receiving waterbody (stream) or discharge point. It contains the coordinates of the plant (mainly for central building), but not the coordinates of the discharge points, so the assignment to the AU unit is fraught with errors.

Advantage: The biggest advantage is that it has the best organized coherent database of the most important direct emitters, which also categorizes activities and marks plant closures.

[49] https://sdi.eea.europa.eu/catalogue/srv/eng/catalog.search#/metadata/63a14e09-d1f5-490d-80cf-6921e4e69551
[50] https://industry.eea.europa.eu/download

3.3.2.1.2 Austria

Name and availability, spatial and time scale: EMREG-OW database[51] contains discharge data from all over in Austria with basic metadata. We received the database exported on 12/22/2021 supplemented with the database exported on 4/30/2024 thanks to our partner TU Wien.

Disadvantages, problems: The export files received only included industrial facilities where the discharge flow (Q) was recorded. Several E-PRTR facilities were missing from the database. Concentrations or loads were not associated with the facilities. Facilities could not be linked to the E-PRTR and IED databases based on identifiers. The data link had to be performed manually. It is common for a single E-PRTR load to be associated with multiple discharge points if a facility has multiple discharge points. It is important to highlight that the discharges (Q) included here may not only be of technological processes, but there is also no information on the type of water used, e.g. cooling water, extracted uncontaminated groundwater, social wastewater.

Advantage: After linking with the E-PRTR, the receiving AU could be determined based on the coordinates of the discharge points. The national database provides data on the emitted discharge flows (Q). The E-PRTR database could be supplemented with a few more significant emitters.

3.3.2.1.3 Czechia

Name and availability, spatial and time scale: ISVA-VODA portal[52] contains discharge data from all over in Czechia, directly downloadable data in Czech language. The available data is from the 2020 collection.

Disadvantages, problems: The export includes several facilities that do not belong to the E-PRTR activities based on their scope of activity and whose metal/PFAS/pharmaceuticals emissions are not significant (e.g. hotel, sports facilities, gas station). The E-PRTR activity type could be inferred from the economic activity code, but this was fraught with uncertainty. The facilities did not have concentration or loads in the database. The facilities could not be linked to the E-PRTR and IED databases based on identifiers. The data link had to be performed manually. It is common for a single E-PRTR load to be matched to multiple emission points if a facility has multiple emission points. It is important to highlight that the water discharges included here may not only be of technological processes, but there is also no information on the type of water used, e.g. cooling water, extracted uncontaminated groundwater, social wastewater.

Advantage: After linking with the E-PRTR, the receiving AU could be determined based on the coordinates of the discharge points. The national database provides data on the emitted discharge flows (Q). The E-PRTR database could be supplemented with a few more significant emitters.

3.3.2.1.4 Germany, Baden-Württemberg

Name and availability: LUBW database[53], direct-downloadable.

Disadvantages, problems: Only coordinates available here. Discharge data was only stored at local authority level. The facilities did not have concentration or loads in the database.

Advantage: The facilities could be linked to the E-PRTR and IED databases based on identifiers. The E-PRTR database could be supplemented with a few more relevant emitters.

[51] https://maps.wisa.bmluk.gv.at/emreg

[52] https://www.voda.gov.cz/?page=vypousteni-do-povrchovych-vod

[53] https://rips-metadaten.lubw.de/trefferanzeige?docuuid=15bb18a0-7d6c-49c6-9075-875ad3da791f

3.3.2.1.5 Germany, Bavaria

Name and availability: LFU database[54], direct-download with geodata. Discharge data should be requested with a paid service, but it was not needed after all, because many countries could not give discharge data, and we need a homogenized model database.

Disadvantages, problems: Only coordinates available here. The facilities did not have concentration or loads in the database.

Advantage: The facilities could be linked to the E-PRTR and IED databases based on identifiers. The E-PRTR database could be supplemented with a few more relevant emitters.

3.3.2.1.6 Slovakia

Name and availability, spatial and time scale: The Slovakian Vyp VUVH industrial database shared by our Tethys partner from Water Research Institute from 2019 to 2023.

Disadvantages, problems: The export includes several facilities that do not belong to the E-PRTR activities based on their scope of activity and whose metal/PFAS/pharmaceuticals emissions are not significant (e.g. hotel, sports facilities). The E-PRTR activity type could be inferred from the economic activity code (NACE), but this was fraught with uncertainty. The facilities did have concentrations or loads, but in very small numbers (these loads were also included in the E-PRTR). The facilities could not be linked to the E-PRTR and IED databases based on identifiers. Moreover, the VUVH database did not contain coordinates, only the name of the receiving watercourse and the fkm, which could not be identified with the map file of the WFD spatial reports[55], since the smallest segment of the primary receiving water stream was indicated. The data link would have had to be done manually, but identifying the 2295 registered discharge points would have provided little additional information. It could have provided useful information on water discharges, indicating the origin of the water used[56], but this was not necessary, as many countries could not give discharge data, and we need a homogenized model database. The use of this water discharge data is recommended for national-level modeling, but requires lengthy preparation work. Another disadvantage is that the reports are made on an annual basis and the entered data changes frequently, and in the absence of appropriate identification codes (e.g. the identification code of the discharge point is missing) it is difficult to eliminate duplicates.

Advantage: The E-PRTR database could be supplemented with a few more significant emitters, and where there were measurement results (load), the concentrations could also be calculated using the water discharge flows, which provided useful data for estimating the emission factor. It is common that a single E-PRTR load can be paired with multiple emission points if a plant has multiple emission points.

3.3.2.1.7 Slovenia

Name and availability, spatial and time scale: The Slovenian national industrial emission database shared by our Tethys partner from Jožef Stefan Institute from 2000 to 2022.

---

[54] https://geoportal.bayern.de/geoportalbayern/suche/suche?0&q=kl%C3%A4ranlagen&f=true

[55] WFD spatial reports: WFD 2022 spatial report and WISE 5 report

[56]Type of water in Slovakian Vyp VUVH database: Sewage, Industrial wastewater, Urban wastewater (municipal), Rainwater, Mining, Cooling, Agricultural - crop production, Agricultural - animal production, Groundwater, Surface water, Water from chemical water treatment, Pure mineral water and swimming pool water, Thermal water, Water from washing ramp, Water from fish farming, Wipe from hydroelectric power plants, Water from tunneling, Pool water - filled with mineral water, Pool water - filled with drinking water, Other

Disadvantages, problems: The export includes several facilities that do not belong to the E-PRTR activities based on their scope of activity and whose metal/PFAS/pharmaceuticals emissions are not significant (e.g. hotel, sports facilities). The database does not contain the code of economic activity, i.e. it was not possible to conclude the E-PRTR activity type. The biggest problem was that the facilities could not be linked to the E-PRTR based on identifiers. Moreover, the database obtained between 2000 and 2021 did not contain coordinates, only site addresses, which in many cases could not be identified even with the help of Google Maps or Open Street Map. Except for the 2022 data, where the emission point coordinates helped with identification, so these could be linked manually with the E-PRTR database. In the case of plants from previous years, data linking was difficult, but feasible, since the reports were made on an annual basis and the entered data changed many times, and in the absence of the appropriate identification codes (e.g. emission point identification code, plant code missing), deduplication was difficult. It is common for a single E-PRTR load to be associated with multiple release points if a facility has multiple discharge points.

Advantage: A significant number of concentrations and water discharge flows (Q), loads, are associated with the facilities and their discharge points, indicating the recipient of the used water. All this data could be used for MoRE modeling.

3.3.2.1.8 Serbia

Name and availability, spatial and time scale: The Serbian national industrial emission database shared by our Tethys partner from Jaroslav Černi Water Institute from 2012 to 2023.

Disadvantages, problems: The export includes several facilities that do not belong to the E-PRTR activities based on their scope of activity and whose metal/PFAS/pharmaceuticals emissions are not significant (e.g. hotel, sports facilities). The biggest problem was that the plants could only be partially linked to the E-PRTR databases based on identifiers. Another disadvantage is that the reports are made on an annual basis and the entered data changes frequently, and in the absence of appropriate identification codes (e.g. the identification code of the emission point also changed), it was difficult to remove duplicates. The received database did not contain coordinates, only site addresses, which in many cases could not be identified even with the help of Google Maps or Open Street Map, only by city. The facilities and their discharge points do not have water discharge flows (Q), and there is no information about the type of water used, e.g. cooling water, extracted unpolluted groundwater, social wastewater. Thus, it is not possible to estimate emission factors or concentrations. It is important to emphasize that the database collects measurement results, i.e. it does not include those plants for which no concentration measurement was performed and therefore no report was made.

Advantage: There are a significant number of loads associated with the facilities and their emission points. All this data could be used for MoRE modeling, as the missing data was identified manually, so the data link could be implemented with the E-PRTR database. The E-PRTR database could be supplemented with a few more significant emitters.

3.3.2.1.9 Bulgaria

Name and availability, spatial and time scale: The Bulgarian national industrial emission database shared by our Tethys partner from Bulgarian Water Association from 2023.

Disadvantages, problems: The disadvantage is that only one year was available, with a relatively small number of measurement results. There are no water discharge flows (Q) associated with the facilities and their discharge points, and there is no information on the type of water used, e.g. cooling water, extracted uncontaminated groundwater, social wastewater.

Advantage: Loads were associated with the facilities and their discharge points, although not in significant numbers. The data link with the E-PRTR database could be implemented using identification codes. All this data could be used for MoRE modeling. The E-PRTR database could be supplemented with a few more significant emitters.

3.3.2.1.10 Hungary

Name and availability, spatial and time scale: The Hungaria national industrial emission database (OKIR) shared by our Tethys partner Budapest University of Technology and Economics from 2007 to 2023.

Disadvantages, problems: It is important to highlight that the database collects measurement results, i.e. it does not include plants that did not have concentration measurements and therefore no reports. The export includes several facilities that do not belong to the E-PRTR activities based on their scope of activity and whose metal/PFAS/pharmaceuticals emissions are not significant (e.g. hotel, sports facility). The disadvantage is that there is no information on the type of water used, e.g. cooling water, extracted uncontaminated groundwater, and social wastewater.

Advantage: Loads were associated with the facilities and their discharge points, although not in a significant number. The data link could be implemented with the E-PRTR database using identifying codes. All this data could be used for MoRE modeling. The E-PRTR database could be supplemented with a few more significant emitters.

3.3.2.1.11 Romania

Name and availability, spatial and time scale: The Romanian national industrial emission database shared by our Tethys partner from National Administration "Romanian Waters" from 2020 to 2021.

Disadvantages, problems: We only received information on wastewater discharges from 7 mines, and beyond that, only the E-PRTR report was available.

Advantage: The E-PRTR database could be supplemented with a few more significant emitters.

3.3.2.1.12 Ukraina

Name and availability, spatial and time scale: The Ukrainian national industrial emission database shared by our Tethys partner from Ukrainian Hydrometeorological Institute from 2015 to 2020.

Disadvantages, problems: We only received information on wastewater discharges from 2 industrial facilities and 5 mining sites, and beyond that, E-PRTR report was not available.

Advantage: The EU E-PRTR database could be supplemented with a few more significant emitters.

3.3.2.2 Selected industrial sites and their activity types

Time scale: We consider facilities that were in operation for at least one year between 2015 and 2023 based on E-PRTR or national databases.

Activity scale: All E-PRTR activities were considered in modelling, except for Intensive Rearing of Poultry or Pigs (farms), as metal emissions from these farms are irrelevant and usually discharge their wastewater to the public sewer system. These emissions are included in the emissions from municipal wastewater treatment plants. Many production facilities have more than one installation, with different production types or technologies. For modeling, we selected the main activity to which we could later assign an emission factor.

Size scale: We mainly incorporated industrial facilities covered by E-PRTR into the model database, supplemented with some of the more significant emitters obtained from national databases.

According to our original plan, we will also collect data on smaller facilities from countries (see above ‚national databases'). However, no country was able to provide properly systematized data on this, meaning that we cannot be used for modeling work. There were problems (listed above), for example, data provision without coordinates. But the main problem was that the activity could not be properly identified, e.g. the databases also contained gas stations, restaurants, manure storages and metal sur-

face treatment plants, but the economic activity code of many facilities was missing or clearly incorrect. Thus, at the project partner meeting in May 2025, we made a clear decision to only consider EPRTR facilities for international modeling, while smaller facilities could be considered in national models. The majority of smaller facilities discharge to municipal wastewater treatment plants through the public sewer network (indirect emission source). The municipal wastewater treatment plants pathway takes these loads into account to some extent, but the industrial pathway is still underestimated in the model, but based on our expert judgment, not to a significant extent.

Results: Totally of 4205 production facilities were considered for MoRE modeling in the Danube Basin (Figure 18 and Figure 19). Industrial data not available from Bosnia, Montenegro and Moldova.

Activity type categories: The facilities belonging to each industrial activity category belong to the same statistical group. The emission factor is applied to each such group. Based on the available information, the E-PRTR categories are the most suitable for categorization. In some cases, some restructuring was necessary, including separation and addition of new types, due to the nature of the activity. It was also important to establish a connection with the BREFs, as this is the main literature source regarding the expected emission levels. A category conversion was necessary, for example, in the case of waste incinerators, which we merged with LCPs (large combustion plans), since these incinerators also partially burn waste for energy production. The final activity list is shown in the table below.

Table 3 - Categories of the Industrial activities with abbreviations of BREF documents and the main industrial sectors

Industrial sector	Main activity of the facility	Notes, examples
1Energy sector	combustions_ID_LCP waste incinerations_ID_WI	+
	refinaries_ID_REF	
2Production and processing of metals	2Production of ferrous metals_ID_SF_FMP	
	2Processing of metals_ID_STMGSTS	Facilities with surface treatment of metal products, galvanization
	2Production and processing of non-ferrous metals_ID_NFM	
3Mineral industry & mining	cement_ID_CLM	
	ceramics_ID_CER	
	glass_ID_GLS	
	mining_open_ID_MIN_O	
	mining_underground_ID_MIN_U	
4Chemical industry	general_ID_CWW	Large chemical factories, except fertilizers, biocides or pharmaceuticals, explosives or polymers
	fertilizers_ID_LVIC_NP	
	biocids_ID_OFC_BIOCID	
	explosives_ID_OFC_EX	
	pharmaceautical_ID_OFC_Pharma	
	polymer_ID_POL	
5Waste management	Animal waste_ID_WTL_Animal	
	hazardous_ID_WTL_H	
	non hadordous_ID_WTL_NH	landfills
6Paper and wood production and processing	Paper and wood_ID_PP	
8Food sector	8Animal products ID_FDM_ME	
	8Milk and cheese products ID_FDM_MI	

Industrial sector	Main activity of the facility	Notes, examples
	8Vegetable products and beverage ID_FDM_V	
9Other activities	9Other activities using chemicals_ID_STS	Surface treatment with solvents

For some metals, emissions from smaller, non-E-PRTR plants may also be significant, which we cannot consider in international modelling due to lack of information. Among others, the utilization of thermal water (baths, gardens) or other groundwater utilization (drinking water for people or the food industry) should be highlighted. These activities may be particularly relevant in the case of arsenic, since the utilization of groundwater results in a continuous release of arsenic from these activities, which varies significantly in space depending on the geochemical background. The models calculate that other activities are not considered by the model.

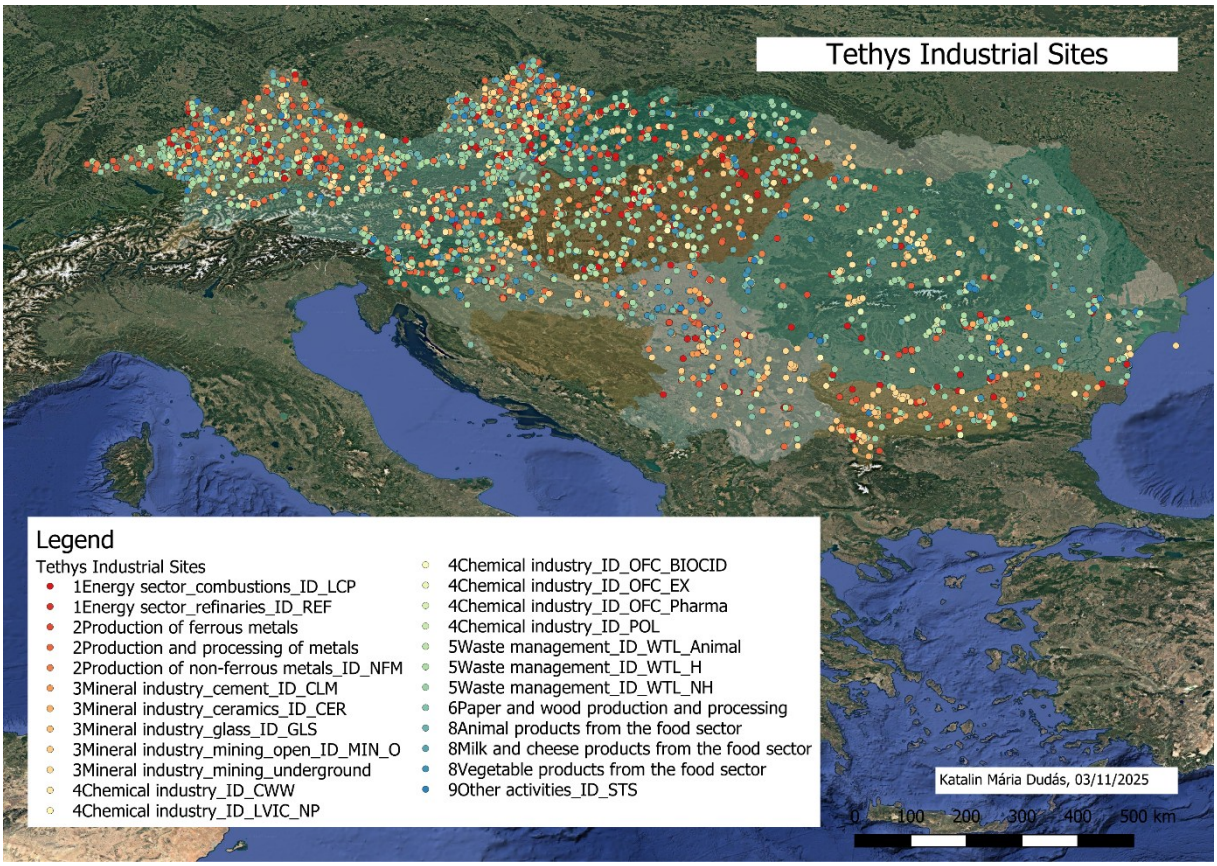

Figure 18 - Industrial facilities across the Danube River Basin based on IEPR database

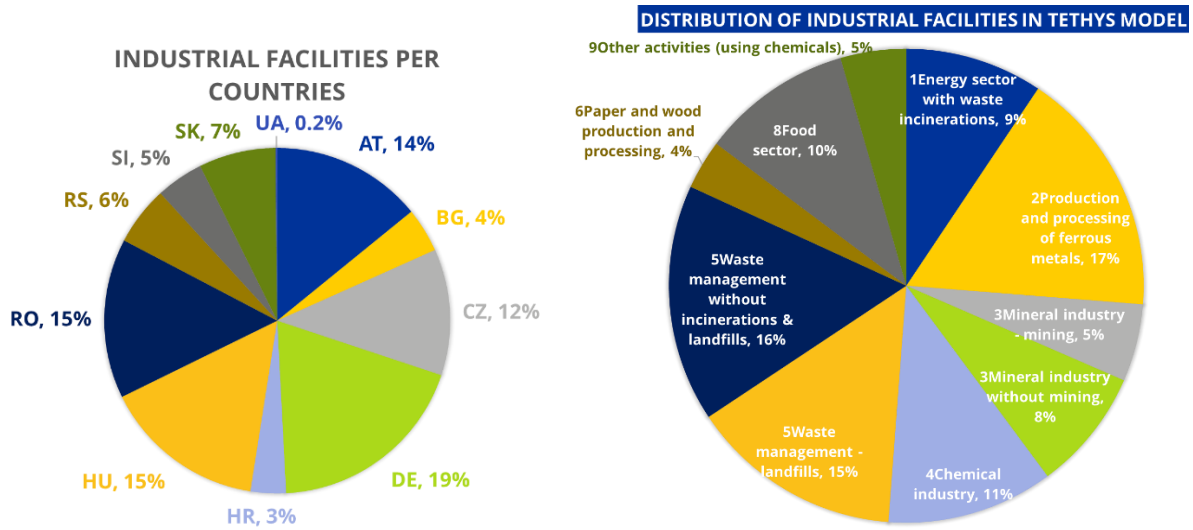

Figure 19 - Portion of Industrial facilities in the Danube Basin per countries (a) and per activities (b) after dataprocess

3.3.3 Landfill sites

The 'Industrial_dataset_v10_December_2023.accdb' database (see section above) contains 611 landfills without information for area and applied technology. This database was extended with landfill sites extracted from the OpenStreetMap database using the QuickOSM plugin in QGIS, where landfill sites are represented as polygons enabling direct area calculation.

For sites present in both datasets, the OpenStreetMap polygon geometries were used to assign area values to the corresponding sites in the industrial database. For landfill sites not represented in OpenStreetMap, site areas were estimated using median values derived from nearby landfill sites.

As a result of this data integration process, a total of 763 landfill sites were included in the model (Figure 20).

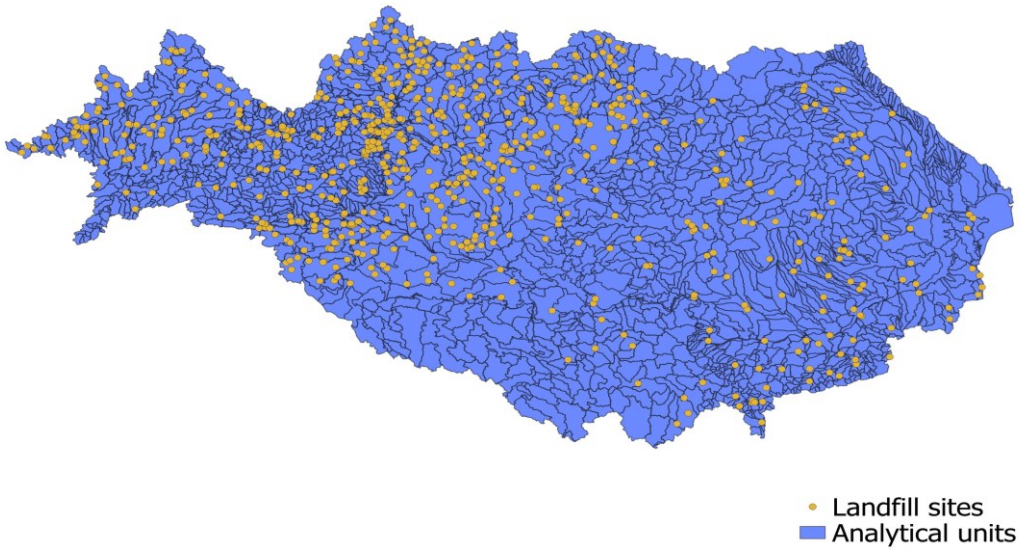

Figure 20 - Landfill sites represented on analytical units

3.3.4 Mining sites

The *Industrial_dataset_v10_December_2023.accdb* database (see Section above) contains mining-related entries for which information on the spatial extent and applied technologies is missing. Nevertheless, the national data collection process enables the total number of records to be increased to a maximum of 222.

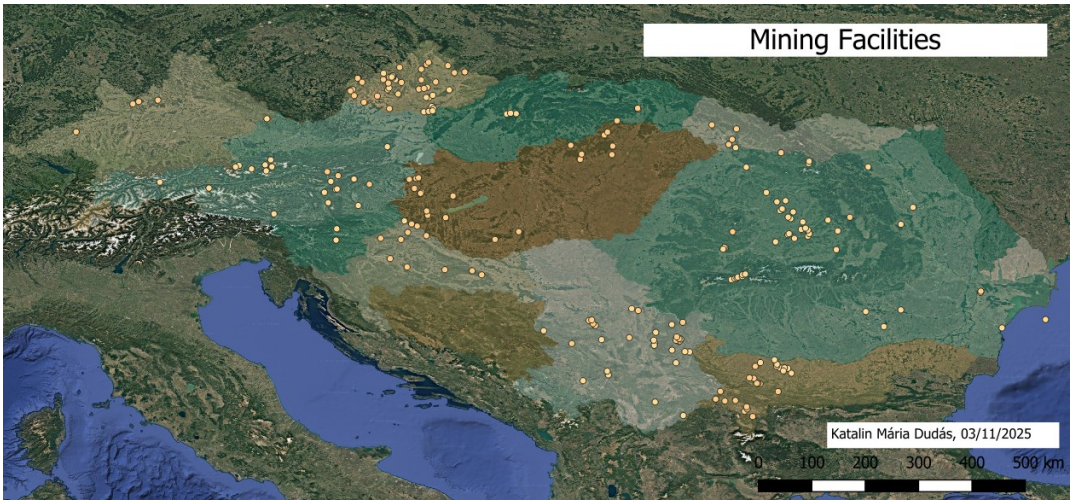

Figure 21 - Mining facilities in the Danube River Basin considered in Tethys modelling

Table 4 - Categories of mining sites

Country	opencast mining and quarrying	Numer of underground mining and related operations
AT	3	17
BA		No data
BG	15	10
CZ	26	21
DE	5	2
HR	9	
HU	22	
ME		No data
MD		No data
RO	4	43
RS	14	17
SI		3
SK	3	3
UA	5	
Total	106	116

3.4 WATER BALANCE

The water balance component of the model is heavily relying on existing methods, used in the MoRE model, but there are new approaches applied as the hydrological input to the model comes from a hydrological model, that makes it possible to directly use some water balance components from the model.

3.4.1 Introduction of the CWatM hydrological model

The Community Water Model (CWatM, (Burek et al., 2020)) model is an open source, freely available hydrological model developed by IIASA (International Institute for Applied Systems Analysis) for large scale water related applications with the specific aim to assess water availability, water demand and environmental needs. The model was designed to support water management decisions. The theoretical background of CWatM is based on the LISFLOOD model.

In terms of the hydrological processes, there is only one modelling option for each. The PET is modelled with the FAO's Penman-Monteith formula, Infiltration, water movement and Rainfall runoff volume are calculated on a three-layer soil accounting model. Groundwater is calculated as a two-layer parallel linear reservoir model (HBV). It is a grid-based model with the spatial resolution optimized to larger scale applications (over 1km) and continuous (daily) temporal resolution. This model is built to handle regional, global gridded data. The model handles all important water transfer options that are necessary for water balance calculations including lakes and reservoirs, water abstractions, irrigation and water return flows etc. It handles land use in a simplified way using only 6 land cover classes. The model has been used with the MODLFOW model to calculate groundwater levels in Beijing (Long, D, 2020), and also used for water quality simulations in China. The model can be run in Python 3 environment and is not supported by user interface.

3.4.2 Spatial resolution and methodological constraints of the two models (CWatM and MoRE)

The two models are using a different approach in terms of basin hydrology. As the MoRE approach is not a hydrological model, the analytical unit of the model does not necessarily have to be correct in hydrological terms. In some cases, the boarders of the analytical units coincide with administrative boarders, such as country or county boarders. In these instances, the hydrological description of the unit is not straight forward for any models. The CWatM model is a spatially distributed model, hence it needs a complete hydrological hierarchy to aggregate flows to the outlets of analytical units. In some cases, this means that multiple outlets belong to one analytical unit.

3.4.3 Ways to handle hydrological model input data

In the current modelling approach, two separate methods is used to describe the water balance:

1. Net total discharge is supplied by the CWatM model, but the distribution of the water balance components is done by the MoRE internal empirical approaches. This approach is following the original approaches, applied in the Moneris model[7], which estimates surface runoff from pervious areas, then estimates Tile drain runoff, atmospheric deposition generated runoff from water surfaces, urban runoff (impervious surfaces) and the rest is equal to groundwater flow (baseflow).
2. The CWatM model supplies the key water balance components: The model directly calculates surface runoff, interflow, and groundwater baseflow. These are associated with some or one of the six land use categories (Forest, grassland, irrigated crop fields, paddy irrigated rice fields, water surfaces, and sealed surfaces). Snow melt and glacier runoff are also directly estimated. These components can be used in the MoRE approach as inputs; however, the aggregation problem described above needs to be addressed. A corrected water balance can be used by the following approach:
    a. Each cell is defined to the corresponding AU
    b. The runoff components and evaporation loss are calculated in each grid cell.
    c. WB components are aggregated to the AU from the corresponding cells
    d. Losses are subtracted from the runoff
    e. Corrected runoff is supplied to the

In the MoRE model the following parameters are related to the CWatM outputs:

Figure 22 - Parameter connections of the two models

3.4.3.1 Testing the two model versions

Total runoff (first version) has been calculated to a subbasin of the Tisza Basin, above the station of Szolnok. The rasterized discharge value map is shown in Figure 23, the aggregated, AU level annual discharge is shown on Figure 24.

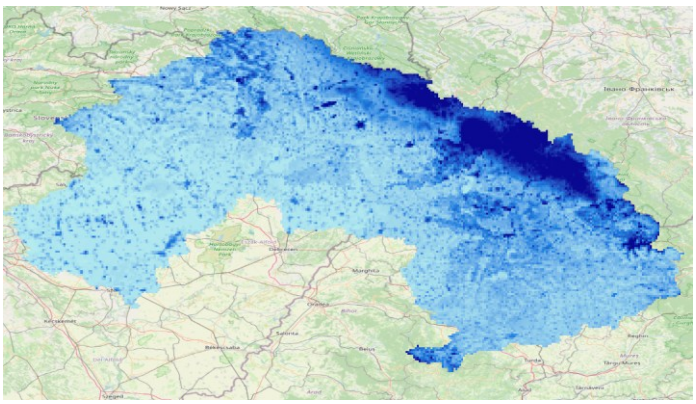

Figure 23 - Cell based runoff from the model output of CWatM

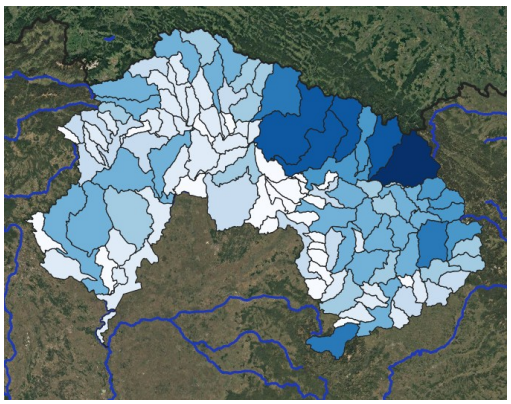

Figure 24 - Average annual runoff from MoRE analytical units above Szolnok station, Hungary

Testing the overall performance of the total runoff

Testing has been performed using the MoRE model itself. The measured discharge (Q_{meas}) data has been imported, then compared to the calculated aggregated runoff (TOT_FNE_Q). Q_{diff} has been calculated by the following equation:

$$Q_{diff} = Q_{meas} - TOT\_FNE\_Q$$

Apparently, the upper part of the Danube (German part) is slightly underestimated, while overall, the total annual runoff is underestimated.

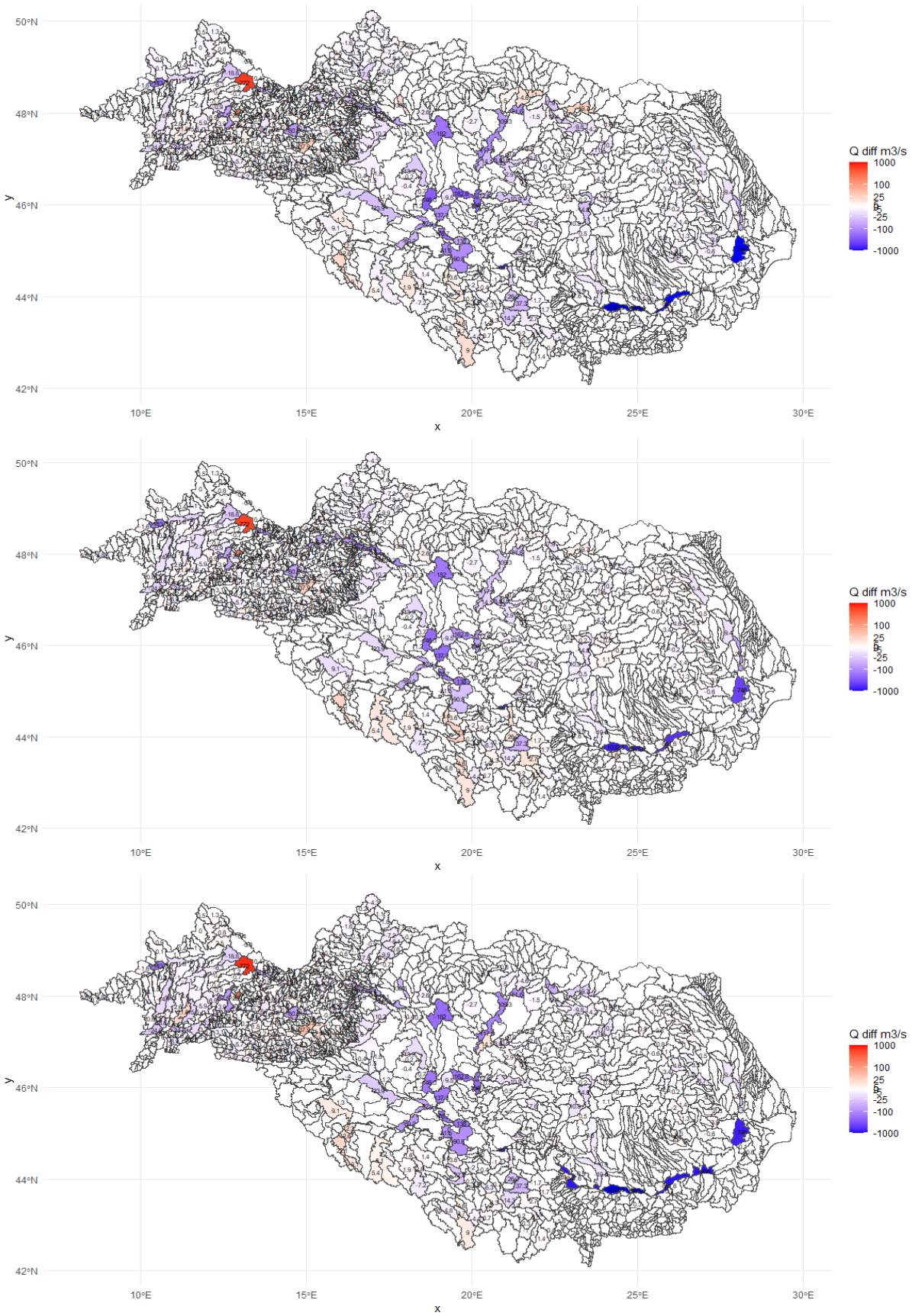

Figure 25 - Q differences in AUs

4 DERIVING EMISSION FACTORS AND CONCENTRATIONS FOR MODEL PARAMETERS

4.1 EFFLUENT CONCENTRATION FROM WWTP DIRECT DISCHARGERS

Municipal sewage discharges – whether treated or untreated – represent – very generally – the most important pathway after erosion; thus, the derivation of the WWTP emission factors is of utmost importance. As a starting point, existing literature on this topic was scrutinized, including, but not limited to

- The JRC report "Water quality in Europe" from year 2019[57]
- EEA report 2020, Report from the Workgroup Chemicals sub-group on emissions to water[58]
- Kardos 2025 has delivered "per capita" emission rates for the pharmaceuticals Carbamazepine and Diclofenac[59]

Data from previous uses of the MoRE model were considered, and – where needed – further developed, including

- MoRE application for Austria in 2021[60],
- MoRE modeling for 7 subcatchments of the Danube in frames of the Danube Hazard m^3c project in 2020-2022[61],[62],[63],
- MoRE model run for the Upper-Danube (Danube upstream to Budapest) within the PROMISCES project in 2024[64],[65].

[57] Pistocchi, A., Dorati, C., Grizzetti, B., Udias, A., Vigiak, O., Zanni, M., Water quality in Europe: effects of the Urban Wastewater Treatment Directive. A retrospective and scenario analysis of Dir. 91/271/EEC, EUR 30003 EN, *Publications Office of the European Union*, Luxembourg, 2019, ISBN 978-92-76-11263-1, doi:10.2760/303163, JRC115607

[58] Caroline Whalley (EEA) and Joost van den Roovaart (ETC-ICM) (2020), Report from the WG Chemicals sub-group on emissions to water, EU *CIRCABC platform*

[59] Kardos, M.K., Patziger, M., Jolánkai, Z. et al. The new urban wastewater treatment directive from the perspective of the receiving rivers' quality. *Environ Sci Eur* 37, 10 (2025). doi: 10.1186/s12302-024-01040-2

[60] Steffen Kittlaus, Manfred Clara, Jos van Gils, Oliver Gabriel, Marianne Bertine Broer, Gerald Hochedlinger, Helene Trautvetter, Gerold Hepp, Jörg Krampe, Matthias Zessner, Ottavia Zoboli, Coupling a pathway-oriented approach with tailor-made monitoring as key to well-performing regionalized modelling of PFAS emissions and river concentrations, *Science of The Total Environment*, 2022, https://doi.org/10.1016/j.scitotenv.2022.157764

[61] Zoboli, O., Kovacs, A., Kittlaus, S. et al. Spurenstoffmanagement im Donaueinzugsgebiet. *Österr Wasser- und Abfallw* 75, 558–571 (2023). https://doi.org/10.1007/s00506-023-00985-7

[62] Gabriel, O., Marianne Bertine Broer, Thomas Rosmann, Clemens Steidl, Steffen Kittlaus, Matthias Zessner-Spitzenberg, Ottavia Zoboli, Nikolaus Weber, Adrienne Clement, Máté Krisztián Kardos, Zsolt Jolankai, Katalin Maria Dudas, Galina Dimova, Radoslav Tonev, Dimitar Mihalkov, Dimiter Alitchkov, Ioana Nedelea, Corina Boscornea, Mugurel Sidau, Elvira Marchidan, Jos van Gils, Sibren Loos, Adam Kovacs: Danube Hazard m3c project output T2. 1 "Harmonized MoRE model adapted to specific territorial characteristics within the DRB" *Project Report* 2023.

[63] Oliver Gabriel, Marianne Bertine Broer, Thomas Rosmann, Clemens Steidl, Steffen Kittlaus, Matthias Zessner-Spitzenberg, Ottavia Zoboli, Nikolaus Weber, Adrienne Clement, Máté Krisztián Kardos, Zsolt Jolankai, Katalin Maria Dudas, Galina Dimova, Radoslav Tonev, Dimitar Mihalkov, Dimiter Alitchkov, Ioana Nedelea, Corina Boscornea, Mugurel Sidau, Elvira Marchidan, Jos van Gils, Sibren Loos, Adam Kovacs: Danube Hazard m³c project Output T2. 2: „Report on improved system understanding as basis for adapted transnational emission modelling at DRB scale" *Project Report* 2023.

[64] Liu, M., Saracevic, E., Oudega, T.J. *et al.* Investigating the extent of PFAS contamination in the Upper Danube Basin across environmental compartments. *Environ Sci Eur* **37**, 99 (2025). https://doi.org/10.1186/s12302-025-01141-6

[65] Liu, M., Saracevic, E., Kittlaus, S. et al. PFAS-Belastungen im Einzugsgebiet der oberen Donau. *Österr Wasser- und Abfallw* 75, 503–514 (2023). https://doi.org/10.1007/s00506-023-00973-x

As a third type of information source, existing data[66] was – as far as possible – complemented (especially from the metadata side); emission factors were newly derived. By the delivering EF, data is arranged along various dimensions. The most trivial is to subset the data according to treatment technology. This has the advantage that the delivered EF can be easily utilized when calculating future management scenarios. However, it might hide regional patterns. Another option is to subset the data along the size (constructed capacity) of the WWTP, as it was done for PFAS in the PROMISCES project. For some substances / regions, there might not be enough data to deliver such EF; in addition, the question of thresholds arises. Regionality was already mentioned, which can be an important factor both in the case of PFAS and in the case of pharmaceuticals. Instead of presenting all investigations, here we report the most interesting results. By this, we will give examples for almost all the mentioned background factors which by chance and not by intent.

The basic approach is that we deliver effluent concentration values as emission factors. The total emission of the plant / wastewater discharge point can then be calculated:

$$E = Q \cdot C$$

Where

E is the plant substance emission in [mass / time]

Q is the discharge of the plant / discharge point [volume / time] in the given year / period

C is the delivered emission factor for the given substance and treatment stage (and /or region, and/or plant size, etc.) [mass / volume].

Alternative ways would be to deliver EF for population equivalents / inhabitant values. These were also investigated, but seemed problematic. Population equivalent are standardly used in wastewater treatment by norming for BOD. Thus, they hide the industrial share in the wastewater, which very much influences the micropollutant emissions. The number of normal people (capita) would be a much more exact indicator of the pollution amount, but is not available in the large databases mentioned in Chapter 3.3.1.

4.1.1 Metals

There is a significant effect of the fact of whether the treatment technology contains chemical precipitation or not ("chemical precipitation" plots, Figure 26), although the data is scarce, but the difference is at least clearly visible.

4.1.1.1 Chromium and nickel

Emission factors (EF), as concentration [µg/l], for chromium and for nickel are determined by plant constructed capacity.

Table 5 - Statistics and suggested EF of HM concentration by plant size (constructed capacity). Mean ± standard deviation (number of >LOQ measurements / number of total measurements)

	Statistics		Suggested EF	
WWTP capacity class	**Cr**	**Ni**	**Cr**	**Ni**
		µg/l		
PE < 2000	3.4 +/- 2.3 (12/16)	3.1 +/- 2.6 (24/28)	3	3

[66] Kittlaus, S., Kardos, M.K., Dudás, K.M. *et al.* A harmonized Danube basin-wide multi-compartment concentration database to support inventories of micropollutant emissions to surface waters. *Environ Sci Eur* **36**, 52 (2024). https://doi.org/10.1186/s12302-024-00862-4

WWTP capacity class	Statistics Cr	Statistics Ni	Suggested EF Cr	Suggested EF Ni
		µg/l		
2000 <= PE < 5000	2.7 +/- 2.6 (17/29)	4.5 +/- 10.4 (41/51)	3	4
5000 <= PE < 10 000	5.7 +/- 10.4 (14/19)	4.2 +/- 2.8 (92/96)	5	4
10 000 <= PE < 100 000	4.9 +/- 7.8 (83/89)	5.4 +/- 6.0 (562/577)	5	5
PE >= 100 000	4.6 +/- 8.1 (80/98)	6.1 +/- 4.0 (517/536)	5	6

4.1.1.2 Copper, zinc, arsenic, cadmium and lead

The effluent concentration is determined by the presence / absence of chemical precipitation in the treatment technology.

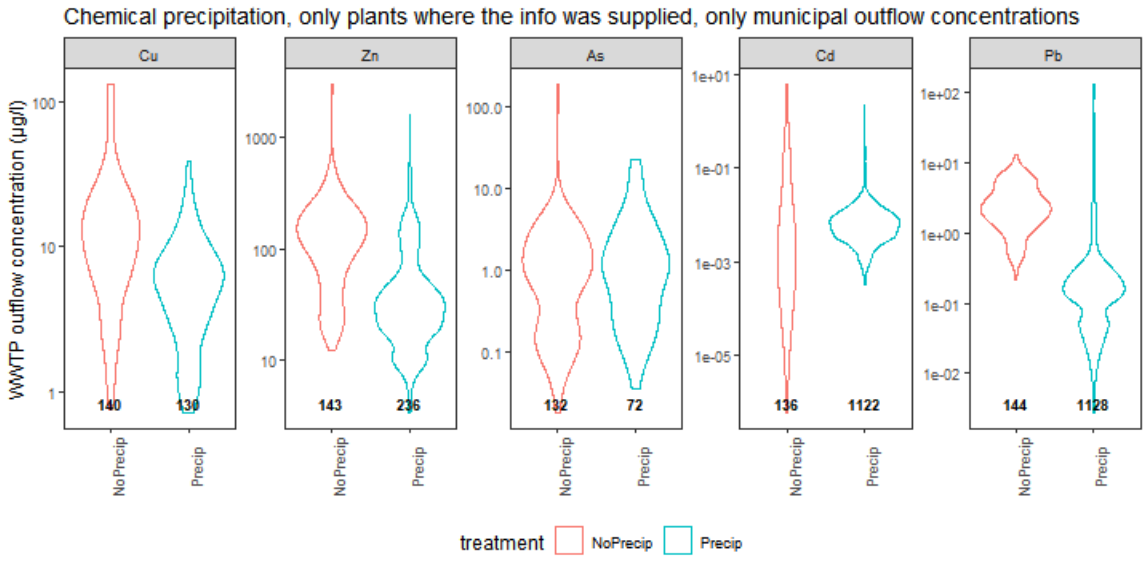

Figure 26 - Violin plots a WWTP effluent heavy metal concentrations by presence / absence of chemical precipitation in the treatment technology.

Table 6 - Statistics and suggested EF of HM concentration by the presence / absence of chemical precipitation. Mean ± standard deviation (number of >LOQ measurements / number of total measurements).

Parameter	Statistics No precipitation	Statistics With precipitation	Suggested EF NoPrecip.	Suggested EF Precip.
		µg/l		
Cu	17 +/- 21.7 (136/140)	6.6 +/- 5.8 (121/130)	17	7
Zn	157.6 +/- 260 (132/143)	61 +/- 133 (218/236)	160	60
As	2.7 +/- 15.9 (103/132)	2.2 +/- 4 (53/72)	2.7	2.2
Cd	0.10 +/- 0.58 (28/136)	0.018 +/- 0.125 (950/1122)	0.1	0.02
Pb	2.8 +/- 2.1 (130/144)	0.61 +/- 5.6 (747/1128)	2.8	0.6

Following concentrations were applied in the model for municipal WWTP effluents:

Table 7 - Applied municipal WWTP effluent concentration values in µg/l.

	As	Cd	Cr	Cu	Ni	Pb	Zn
				µg/l			
No treatment	6.00	1.000	50.0	36.00	180.0	15.00	420.0
Primary treat.	4.00	0.500	20.0	21.00	90.0	7.50	210.0
Sec	3.00	0.100	5.0	15.00	9.0	3.00	150.0
Tert	2.50	0.020	4.0	9.00	8.0	1.00	60.0
Quat.	0.10	0.001	0.4	0.90	0.8	0.10	6.0
small WWTPs	1.55	0.002	2.8	4.51	6.8	1.43	36.8

4.1.2 PFAS

Regarding PFAS compounds, the economical development seems to be a determining factor. Thus, the regionalization approach was applied. Countries in the Danube River Basin were classified into three categories.

- Upper Danube countries: Germany, Austria, Czech Republik and Slovenia. Having the highest economic potential, the use the most products containing PFAS, emitting the highest pollution.
- Lower Danube countries: Bulgaria, Republic of Serbia and Ukraine. Having the lowest economic potential, the use the least products containing PFAS, emitting lower.
- The rest of the countries (Slovakia, Hungary, Croatia and Romania) were classified into the third category standing in between the previous two categories.

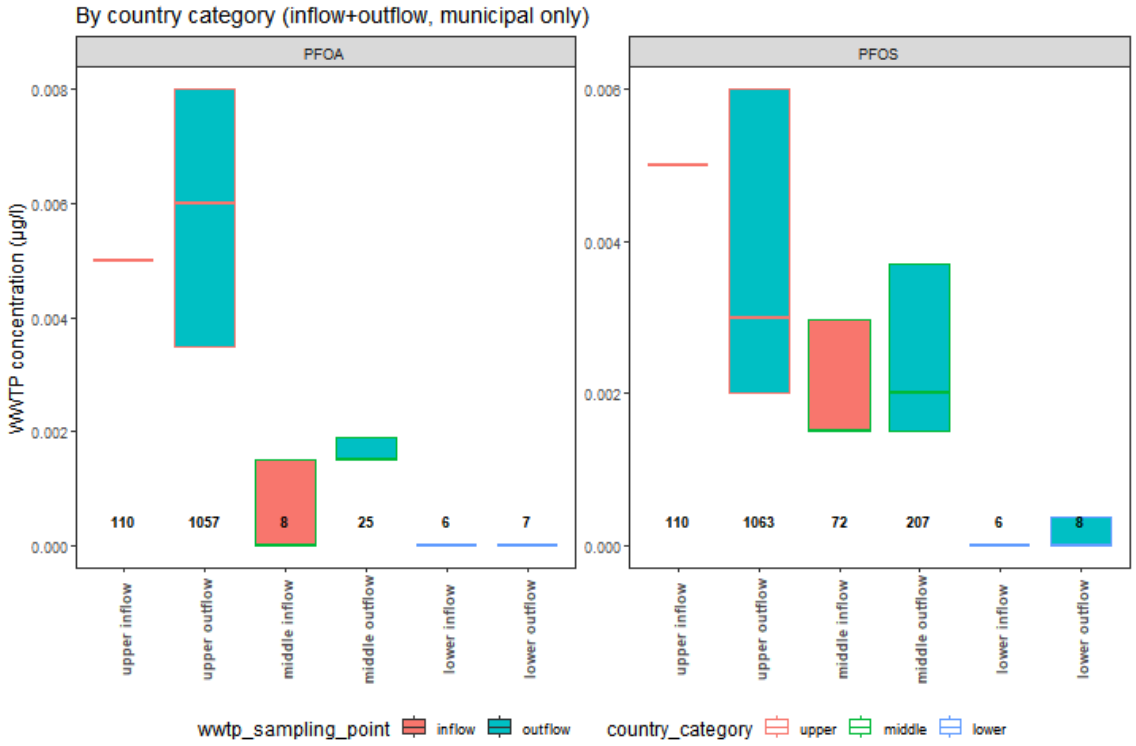

Figure 27 - WWTPinfluent and effluent concentrations by country category. "upper" countries: DE, AT, CZ and SI; "middle" countries: SK, HU, HR and RO; "lower" countries: RS, BG and UA. Number of measure- ments indicated.

Table 8 - PFAS influent and effluent concentrations in the three country groups.Mean ± standard deviation (with number of >LOQ samples) per country category.

country class		PFOS	PFOA
„upper" countries: DE, AT, CZ, SI	inflow	0.0126 ± 0.0385 (22)	NA
	outflow	0.0089 ± 0.0382 (896)	0.0094 ± 0.0239 (952)
"middle" countries: SK, HU, HR, RO	inflow	0.0062 ± 0.0274 (26)	NA
	outflow	0.0042 ± 0.0114 (126)	0.0017 ± 0.0021 (10)
„lower" countries: BG, RS, UA	inflow	NA	NA
	outflow	NA	NA

We suggest to apply the delivered EF based on EUSTAT Price level indices for year 2022[67], see the next table.

Table 9 - PFAS influent and effluent concentrations in the three country groups.

Country / region	GDP PPP price level indices, year 2022
Switzerland	153.9
Germany	111.5
Austria	111.4
EU27	100.0
Slovenia	83.4
Czechia	82.3
Slovakia	79.0
Croatia	67.4
Hungary	64.3
Bulgaria	59.4
Serbia	59.1
Romania	56.9
Montenegro	55.1
Bosnia and Herzegovina	54.5

As a second approach, the size of the wastewater treatment plant was considered.

[67] Eurostat Data Browser, Purchasing power parities, price level indices, nominal and real expenditures by analytical categories - based on COICOP 2018. Online data code: prc_ppp_ind_1 DOI: https://www.doi.org/10.2908/prc_ppp_ind_1 Last update: 07/JAN/2026 11:00.

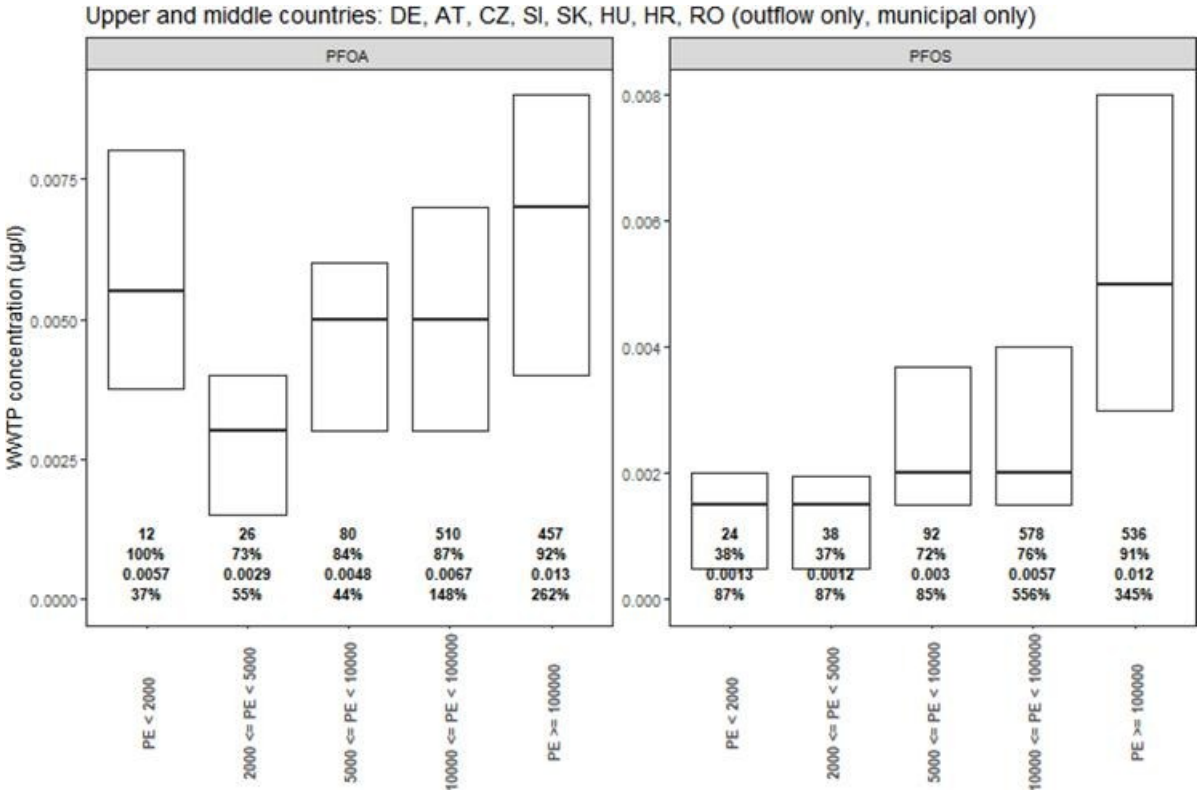

Figure 28 - WWTP effluent concentrations by plant constructed capacity in "upper" and "middle" countries (DE, AT, CZ, SI, SK, HU, HR and RO). Explanation of the grey figures (numbers): 1. number of measurements in the particular category; 2. fraction of >LOQ measure-ments; 3. mean value; 4. standard deviation relative to the mean value.

Table 10 - PFAS effluent concentrations in upper and middle DRB countries (i.e. excluding RS, BG and UA) by plant size. Mean ± standard deviation (number of >LOQ samples / total number of samples).

WWTP capacity class	Statistics of the municipal WWTP effluent data (µg l^{-1})		Suggested EF (ng l^{-1})	
	PFOS	PFOA	PFOS	PFOA
PE < 2000	0.0013 +- 0.0012 (9/24)	0.0057 +- 0.0021 (12/12)	1	5
2000 <= PE < 5000	0.0012 +- 0.001 (14/38)	0.0029 +- 0.0016 (19/26)	1	3
5000 <= PE < 10K	0.003 +- 0.0026 (66/92)	0.0048 +- 0.0021 (67/80)	3	5
10000 <= PE < 100K	0.0057 +- 0.032 (441/578)	0.0067 +- 0.01 (445/510)	6	7
PE >= 100K	0.0124 +- 0.043 (489/536)	0.0131 +- 0.034 (419/457)	10	10

N.B.: Concentrations are substantially more scattered in the larger plant categories (PE > 10K) for both PFAS compounds.

In the current model version, 10 ng/L effluent concentration is applied for both PFAS compounds for treatment technology up to tertiary. For quaternary treatment, 1 ng/L effluent concentration is ap-plied.

4.1.3 Pharmaceuticals

Two approaches were applied.

Approach 1: Effluent concentrations. No significant difference between secondary and tertiary treatment plants were found. There was insufficient data to derive concentrations for other treatment types (no treatment, primary, quaternary), teherfore, uniform (independent of treatment type) concentration value was applied for all treatment stages.

Approach 2: Per capita emission factors. Besides working with plant effluent concentrations (approach 1), "per capita" emission values were calculated by corresponding "discharge during sampling" and "number of plant-connected inhabitants" values to the effluent concentrations.

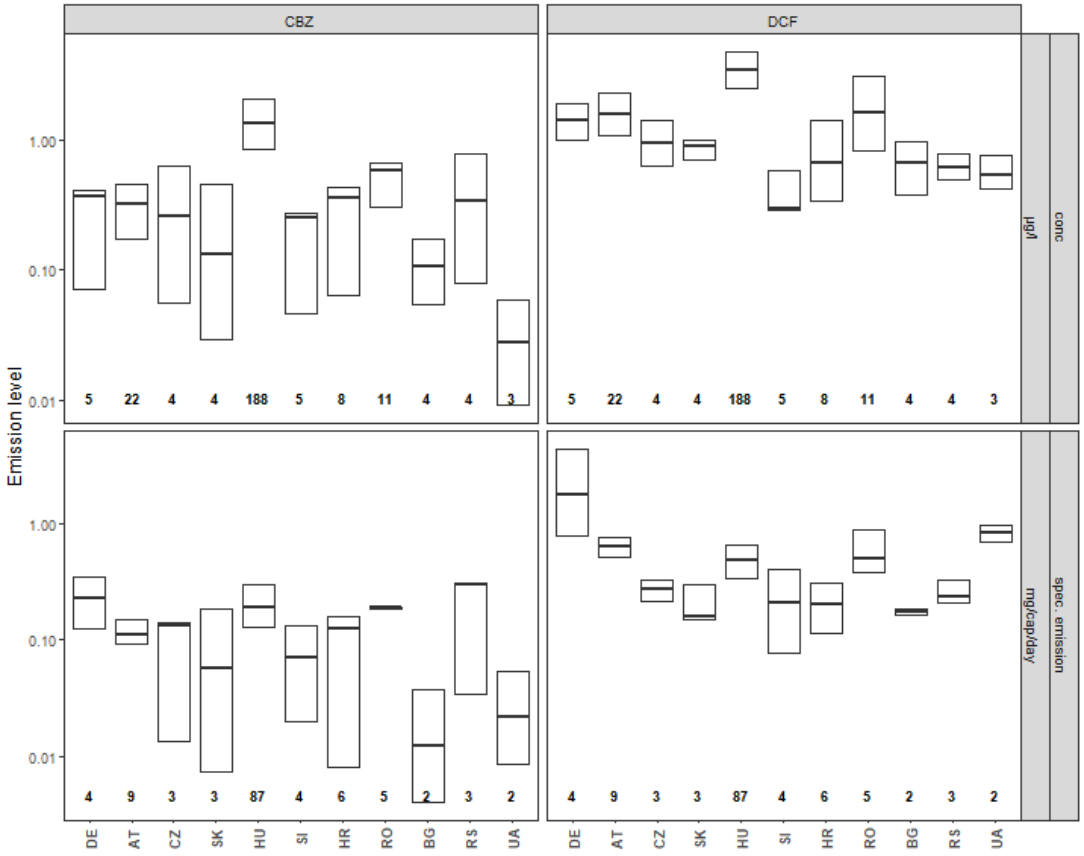

Figure 29 - Municipal WWTP effluent concentrations and specific emissions by country. Number of measurements indicated.

Table 11 - Pharmaceutical effluent concentrations and emission factors. Mean ± standard deviation (with number of >LOQ samples / total number of samples).

substance	concentration (µg l^{-1})	statistics of the measured values (mg cap^{-1} day^{-1})	Suggested emission factor (mg cap^{-1} day^{-1})
CBZ	0.5	0.22 +/- 0.21 (128/128)	0.20
DCF	1.0	0.63 +/- 0.83 (128/128)	0.60

4.2 METHODOLOGY FOR THE DERIVATION OF INDUSTRIAL EMISSION FACTORS

When determining the emission factors, it was necessary to be extremely circumspect, as industrial facilities are very diverse. The industrial sector and within it the type of goods produced, the size of the facility, the technological solutions, cleaning/recovery technologies and emission methods also different. We had to bear in mind that the most significant emitters are required to measure and report

emissions, but these largest plants carry out measurements within the framework of self-monitoring, i.e. the measurement results are subject to uncertainty.

4.2.1 Industrial emissions into the water

Several possible methods are available for estimating industrial emissions (loads) to water of unmeasured facilities:

1) One of the most trivial and accurate methods would be to multiply the „concentration specified for the given industrial activity“ (emission factor, EF as concentration) **by the discharged water flow of the used water passing through the technological processes of the given plant** (Q_{tech} as an annual volume of technological wastewater).
   Or multiply the „concentration specified for the given industrial activity“ (emission factor, EF as concentration) **by the total discharged water flow of the given plant** (Q as an annual volume of total wastewater).

A key element of this method is the knowledge of discharged water flow data, which are available from the following sources as listed below:

- ‚Annual volume of technological wastewater' (Q_{tech}) is available only from HU and SK;
- ‚Annual volume of total wastewater' (Q) is available from AT, CZ, SL, and DE;
- Neither ‚annual volume of technological wastewater“ (Q_{tech}) nor ‚annual volume of total wastewater' (Q) is available from RS, UA, RO, HR, BG, MD, ME, and MO.

The method was discarded because the ‚annual volume of technological wastewater' was only available from SK and HU. However, it was pointed out that in countries where only the total discharged water flow (Q) per discharge point is only available (without indicating its origin, such as AT, CZ and SL), then multiplying the discharged water flow (Q) by EF (conc.) would cause a huge error. The reason for this is that technological wastewater is usually only a small part of ‚other waters' (e.g. cooling water, transport water, heating water, rainwater, social wastewater, extracted and not used groundwater). Moreover, examining the SK and HU data, it turned out that often the mentioned ‚other waters' (except social wastewater) are discharged directly into the receiving waterbody and the technological wastewater is discharged into public sewers (mostly after pre-treatment). The Serbian, Romanian, Bulgarian, and Croatian territory itself is so large that a different method must be chosen if we want to maintain the international homogeneity and robustness of the MoRE model.

2) The other most accurate method would be to know the **production capacity of industrial facilities** and generate an emission factor per unit produced. However, production data is even less available than water discharge flow (Q) data.

Only in the case of waste incinerators and large combustion plans were there adequate production-specific data available. In the case of combustion plans, we found a correlation between the annual 'Thermal Input (MJ)' and the discharged metal loads via wastewater, based on the E-PRTR data. However, with the national data received, we had to discard the previous correlations, as it would have caused a significant error due to the smaller facilities, and the ‚normal approach' (detailed in point 4) proved to be more accurate.

3) Another possible method is to use **literature data, including reviewing BAT reference documents (BREFs)**. Articles that examine the emissions of individual activities may be prepared for a variety of reasons. BREF[68] documents also collect and process these articles and even request additional data from Member States, to estimate the expected emission levels of a modern facility using the best available technology. We have reviewed these BREFs carefully and compared them with the

[68] https://bureau-industrial-transformation.jrc.ec.europa.eu/reference

industrial emission data available at the time (BG, RS, SK and SL data were obtained later), see Annex II for the results. Experience has shown that smaller plants operate near to the lower end of the associated emission ranges estimated in the BREFs, as if they apply the best available technology, and their emissions are even lower than that. This literature method and studies also use water discharge flow data. A possible reason for the low emission levels may be the high proportion of dilution waters.

The other, real reason is that BREFs and often literature analyses examine the largest plants, as these are the ones that are reported in the E-PRTR and are the riskiest and most interesting for the authors of the articles. However, it appears that the measured and reported loads in Danube River Basin (using AT, HU, BG, RS, SK, and SL national data) are well below the expected emission levels range of the BREFs and the E-PRTR reporting threshold, also. (Figure 30). (BREF AEL ranges vary by activity, see Annex II for details.)

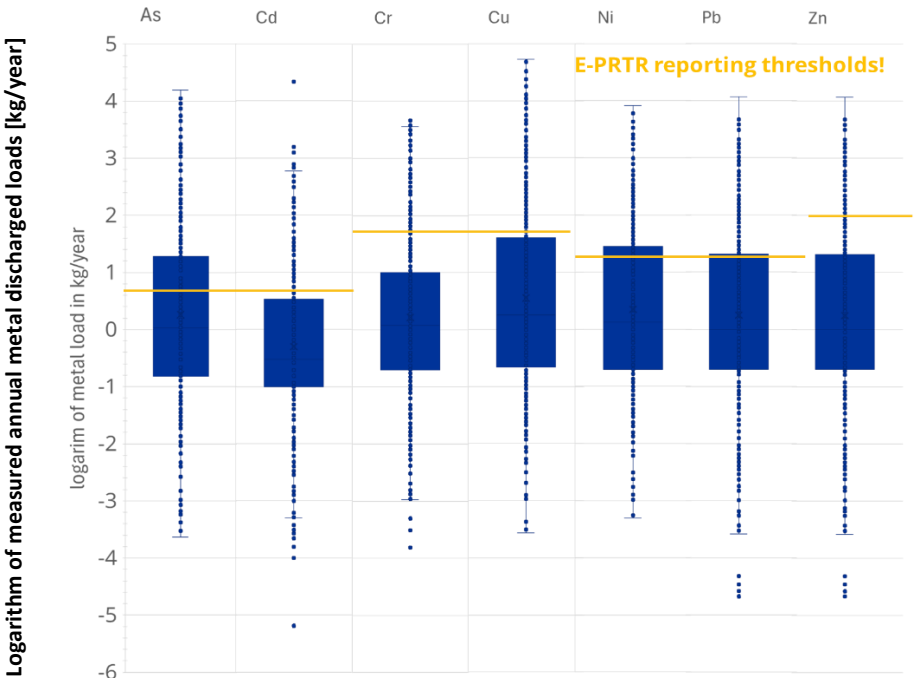

Figure 30 - Logarithm of measured annual metal discharged loads [kg/year] and the E-PRTR reporting the threshold

Based on the above analyses, we have established that it is not possible to estimate the smaller facilities based on the measurement results of the largest plants (even taking into account discharged water flows). An estimation must be given for the smaller plants based on the loads measured in such smaller facilities. This is how we came to apply the 4th methodology, which is the most robust and – with the use of variants – can provide the most accurate picture of the extent of industrial emissions.

4) Use ‚**Sectorial Emission Factor as an activity-specific load**', if site-specific data is not available. The Sectorial Emission Factor is based on the median emission values (loads) of the given activity type.
   - It is important to emphasize that the measured values are indeed included in the MoRE model in an unchanged form (except for the measured value 0, see later).
   - At those discharge points where there was no measurement in the given year, but there was in other years, the average of the measurement results is considered.
   - Where no emission load could be determined based on the above, the 'Sectoral Emission Factor' was applied, which was estimated using the following method.

For the variant 1 (normal case) to use median value instead of mean allows the exclusion of plants with high emissions, whose measured values added to the model anyway. The method uses the 20% percentile to estimate the 'best case' for calculating minimal industrial loads (variant 2). The national databases contains many 0 values, without indicating LOQ-s or other information related to 0 value. The huge number of zero have a significant statistical distorting effect, therefore variant 4 'modified normal case' omitted zero values from delivering EF (statistics) and the variant 4 uses 20% percentile to estimate emission load instead of 0; and uses sectorial median of the modified statistics. The variant 3 describe the 'worst case', uses 20% percentile to estimate emission load instead of 0; and uses sectorial 80% percentile of the modified statistics. (Table 12)

Table 12 - Variants

Used variants for modelling	Measured 0	Not measured facilities
VAR1 = Normal case	Considering	EF = Sectorial median
VAR2 = Best case (minimum)	for delivering EF	EF = Sectorial 20% percentile
VAR3 = Worst case (maximum)	Omitted from delivering EF,	EF = Sectorial 80% percentile
VAR4 = Modified normal case	& using Sectorial 20% percentile instead of measurements	EF = Sectorial median

*Table 13 - Statistics and suggested **Sectorial Emission Factor (load)** Arsenic concentration*
a) Number of measured facilities
b) ID_ps_E_HM_AS = Estimated annual metal discharged loads [kg/year]
c) Instead of measured 0 for VAR3 and VAR4

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
1Energy sector	Combustions_ LCP + waste incinerations_WI	164	835.9	2323	0.000	248	0.20	0.2	0	117	1172	2679	0.100	1465	23.20	0.1	1465	23.2
	refinaries_REF	23	6.4	10.6	0.000	8.3	0.40	0.4	0	12	12.4	11.9	5.380	21.9	6.86	5.4	21.9	6.9
2Metal ind.	Production of ferrous metals__SF_FMP	26	10.6	24.4	0.031	11	1.20	1.2	0.03	26	10.6	24.4	0.031	11.0	1.20	0.03	11	1.2
	Processing of metals STMGSTS	20	1.0	4.2	0.001	0.1	0.02	0.02	0.0008	18	1.1	4.4	0.007	0.1	0.02	0.01	0.1	0.02
	Production and processing of non-ferrous metals_NFM	48	1.1	3.2	0.000	0.4	0.13	0.1	0	35	1.5	3.7	0.034	0.4	0.38	0.03	0.4	0.4
3Mineral indus-try	cement_CLM	30	11.3	49.6	0.000	3.5	1.50	1.5	0	23	14.7	56.2	1.200	3.9	1.60	1.2	3.9	1.6
	ceramics_CER	5	0.0	0.1	0.000	0.0	0.00	0.3	0	1	0.2	0.0	0.200	0.2	0.20	0.2	0.2	0.2
	glass_GLS	10	0.5	0.6	0.180	0.4	0.30	0.3	0.18	10	0.5	0.6	0.180	0.4	0.30	0.2	0.4	0.3

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
	mining_open_MIN_O	41	40.2	123.3	0.000	22.8	2.50	2.5	0	31	53.2	139.3	0.700	66.9	8.20	0.7	66.9	8.2
	mining_underground_MIN_U	31	52.0	87.9	0.400	62.0	24.80	24.8	0.4	29	55.5	89.8	0.800	70.0	27.00	0.8	70	27
4Chemical industry	general_CWW	12	5.6	9.7	0.640	3.3	1.32	1.3	0.64	10	6.7	10.3	1.010	7.8	1.81	1	7.8	1.8
	fertilizers_ID_LVIC_NP	4	21.4	12.3	11.080	31.5	20.80	20.8	11.08	4	21.4	12.3	11.080	31.5	20.80	11.1	31.5	20.8
	biocid_OFC_BIOCID	9	0.0	0.0	0.000	0.0	0.00	0.3	0	0						0.3	0.4	0.3
	explosives_OFC_EX	0						0.3	0.31	0						1	7.8	0.3
	pharma_OFC_Pharma	4	0.4	0.1	0.314	0.4	0.32	0.3	0.31	4	0.4	0.1	0.314	0.4	0.32	0.3	0.4	0.3
	polimers_POL	3	1.2	0.3	1.022	1.4	1.24	1.2	1.02	3	1.2	0.3	1.022	1.4	1.24	1	1.4	1.2
5Waste management	animal_WTL_Animal	3	0.2	0.0	0.200	0.2	0.20	0.026	0.2	3	0.2	0.0	0.200	0.2	0.20	0.2	0.2	0.2
	hazardous_ID_WTL_H	0						0.026	0	0						0.2	0.2	0.1
	non hadordous_ID_WTL_NH	65	1.6	10.0	0.000	0.1	0.03	0.026	0	39	2.6	12.8	0.032	0.2	0.10	0.03	0.2	0.1
6Paper and wood production and processing_ID_PP		5	0.0	0.0	0.000	0.0	0.00	0.026	0	1	0.1	0.0	0.100	0.1	0.10	0.1	0.1	0.1

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
8Food sector	Animal products FDM_ME	9	1.0	0.7	0.635	1.5	0.77	0.8	0.64	9	1.0	0.7	0.635	1.5	0.77	0.6	1.5	0.8
	Milk and cheese products FDM_MI	8	1.0	0.8	0.400	1.3	0.80	0.8	0.4	8	1.0	0.8	0.400	1.3	0.80	0.4	1.3	0.8
	Vegetable products FDM_V	21	10.7	40.4	0.800	2.5	1.30	0.8	0.8	20	11.3	41.4	0.893	2.6	1.35	0.9	2.6	1.3
9Other activities using chemicals_ID_STS		9	0.9	1.3	0.000	2.2	0.00	0.02	0	3	2.6	0.6	2.220	3.0	3.00	2.2	3	3

*Table 14 - Statistics and suggested **Sectorial Emission Factor (load)** Cadmium concentration*
a) Number of measured facilities
b) ID_ps_E_HM_CD = Estimated annual metal discharged loads [kg/year]
c) Instead of measured 0 for VAR3 and VAR4

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%		c)	b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median		VAR3	VAR4
1Energy sector	Combustions_ LCP + waste incinerations_WI	197	153.5	1535.7	0.000	1.0	0.10	0.1	0	109	277.4	2056.1	0.100	41.1	0.30	0.1	41.1	0.3
	refinaries_REF	18	69.6	98.8	5.880	113.6	19.85	19.9	5.88	18	69.6	98.8	5.880	113.6	19.85	5.9	113.6	19.9
2Metal ind.	Production of ferrous metals__SF_FMP	46	8.9	16.5	0.001	13.8	0.30	0.3	0.0014	43	9.5	16.9	0.113	13.8	0.40	0.1	13.8	0.4
	Processing of metals STMGSTS	49	0.0	0.0	0.000	0.0	0.00	0.005	0	20	0.0	0.1	0.001	0.1	0.01	0.001	0.1	0.006
	Production and processing of non-ferrous metals_NFM	55	0.1	0.1	0.000	0.2	0.04	0.04	0	37	0.2	0.1	0.032	0.3	0.10	0.03	0.3	0.1
3Mineral industry	cement_CLM	17	0.2	0.2	0.000	0.3	0.10	0.1	0	10	0.3	0.3	0.100	0.4	0.20	0.1	0.4	0.2
	ceramics_CER	41	0.1	0.1	0.000	0.1	0.00	0.1	0	14	0.2	0.1	0.079	0.3	0.11	0.1	0.3	0.1
	glass_GLS	26	0.1	0.1	0.000	0.1	0.10	0.1	0	17	0.2	0.1	0.100	0.3	0.10	0.1	0.3	0.1
	mining_open_MIN_O	45	11.3	24.9	0.100	8.0	1.90	1.9	0.1	39	13.1	26.3	0.320	8.8	2.20	0.3	8.8	2.2

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
	mining_underground_MIN_U	49	40.0	188.1	0.102	5.2	0.80	0.8	0.1	46	42.6	193.9	0.153	7.4	1.10	0.2	7.4	1.1
4Chemical industry	general_CWW	18	0.3	0.5	0.000	0.7	0.09	0.1	0	10	0.6	0.6	0.148	0.9	0.43	0.1	0.9	0.4
	fertilizers_ID_LVIC_NP	4	20.7	9.8	14.960	26.8	21.65	21.7	14.96	4	20.7	9.8	14.960	26.8	21.65	15	26.8	21.7
	biocid_OFC_BIOCID	27	0.0	0.0	0.000	0.0	0.00	0.003	0	0						0.003	1.2	0.003
	explosives_OFC_EX	8	5.1	6.7	0.600	11.0	1.70	1.7	0.6	8	5.1	6.7	0.600	11.0	1.70	0.6	11	1.7
	pharma_OFC_Pharma	3	0.7	0.9	0.003	1.2	0.00	0.003	0.003	3	0.7	0.9	0.003	1.2	0.00	0.003	1.2	0.003
	polimers_POL	3	0.8	0.2	0.646	0.9	0.71	0.7	0.65	3	0.8	0.2	0.646	0.9	0.71	0.6	0.9	0.7
5Waste manage-ment	animal_WTL_Animal	7	0.0	0.1	0.000	0.0	0.00	0.01	0	1	0.3	0.0	0.300	0.3	0.30	0.3	0.3	0.3
	hazardous_ID_WTL_H	6	0.0	0.0	0.010	0.0	0.01	0.01	0.01	6	0.0	0.0	0.010	0.0	0.01	0.01	0.02	0.01
	non hadordous_ID_WTL_NH	73	0.7	2.5	0.000	0.0	0.00	0.01	0	32	1.6	3.6	0.004	0.1	0.02	0.004	0.1	0.015
6Paper and wood production and processing_ID_PP		5	32	4.8	5.6	0.400	9.0	2.85	2.9	0.4	28	5.5	5.7	1.136	11.0	3.79	1.1	11
8Food	Animal products FDM_ME	1	5.6	0.0	5.600	5.6	5.60	0.01	5.6	1	5.6	0.0	5.600	5.6	5.60	5.6	5.6	5.6

Main activity of the facility	Normal statistics results								Modified statistics results								
	a)			20%	80%		b)		a)			20%	80%			b)	
	n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
Milk and cheese products FDM_MI	9	0.3	0.3	0.100	0.3	0.20	0.01	0.1	9	0.3	0.3	0.100	0.3	0.20	0.1	0.3	0.2
Vegetable products FDM_V	15	0.3	0.4	0.000	0.3	0.10	0.1	0	11	0.4	0.4	0.100	0.6	0.10	0.1	0.6	0.1
9Other activities using chemicals_ID_STS	9	26	0.4	0.6	0.000	0.7	0.20	0.2	0	14	0.7	0.6	0.260	0.8	0.65	0.3	0.8

*Table 15 - Statistics and suggested **Sectorial Emission Factor (load)** Chromium concentration*
a) Number of measured facilities
b) ID_ps_E_HM_CR = Estimated annual metal discharged loads [kg/year]
c) Instead of measured 0 for VAR3 and VAR4

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
1Energy sector	Combustions_ LCP + waste incinerations_WI	227	204.4	718.7	0.000	7.8	0.20	0.2	0	170	272.9	819.2	0.100	79.6	0.70	0.1	79.6	0.7
	refinaries_REF	18	213.9	201.9	8.885	361.0	160.25	160.3	8.88	18	213.9	201.9	8.885	361.0	160.25	8.9	361	160.3
2Metal ind.	Production of ferrous metals__SF_FMP	84	92.1	187.6	0.139	118.4	24.20	25.1	0.14	83	93.2	194.4	0.145	118.6	25.07	0.1	118.6	25.1
	Processing of metals STMGSTS	80	0.9	2.8	0.015	0.6	0.11	0.1	0.01	68	1.0	3.1	0.087	0.8	0.20	0.1	0.8	0.2
	Production and processing of non-ferrous metals_NFM	36	14.0	66.6	0.000	1.7	0.65	0.7	0	28	18.0	75.0	0.300	1.9	1.00	0.3	1.9	1
3Mineral industry	cement_CLM	33	3.6	4.9	0.280	5.5	1.40	1.4	0.28	28	4.2	5.1	0.940	6.6	2.45	0.9	6.6	2.5
	ceramics_CER	50	0.5	0.8	0.000	0.7	0.20	0.2	0	37	0.7	0.9	0.100	0.9	0.50	0.1	0.9	0.5
	glass_GLS	11	1.0	1.0	0.200	1.3	0.70	0.7	0.2	11	1.0	1.0	0.200	1.3	0.70	0.2	1.3	0.7
	mining_open_MIN_O	35	44.5	110.9	0.000	35.7	4.00	4	0	26	59.9	125.1	1.700	44.2	11.90	1.7	44.2	11.9

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
	mining_underground_MIN_U	40	13.9	29.1	0.480	12.7	3.35	3.4	0.48	37	15.1	29.9	0.840	14.3	4.60	0.8	14.3	4.6
4Chemical industry	general_CWW	29	6.7	16.0	0.056	5.4	1.62	1.6	0.06	24	8.1	17.2	0.184	6.3	2.64	0.2	6.3	2.6
	fertilizers_ID_LVIC_NP	4	301.7	107.8	218.480	370.9	266.60	266.6	218.48	4	301.7	107.8	218.480	370.9	266.60	218.5	370.9	266.6
	biocid_OFC_BIOCID	26	0.0	0.0	0.000	0.0	0.00	0	0	3	0.1	0.0	0.100	0.1	0.10	0.1	0.1	0.1
	explosives_OFC_EX	9	86.5	105.8	11.740	120.2	42.40	42.4	11.74	8	97.3	107.4	22.500	143.7	55.45	22.5	143.7	55.5
	pharma_OFC_Pharma	16	3.6	5.2	0.017	4.1	1.49	1.5	0.02	16	3.6	5.2	0.017	4.1	1.49	0.017	4.1	1.5
	polimers_POL	3	13.3	5.6	9.345	17.5	14.33	14.3	9.34	3	13.3	5.6	9.345	17.5	14.33	9.3	17.5	14.3
5Waste manage-	animal_WTL_Animal	9	0.3	0.5	0.000	0.3	0.10	0.1	0	5	0.5	0.6	0.100	0.7	0.20	0.1	0.7	0.2
	hazardous_ID_WTL_H	6	0.1	0.0	0.054	0.1	0.06	0.1	0.05	6	0.1	0.0	0.054	0.1	0.06	0.1	0.1	0.1
	non hadordous_ID_WTL_NH	84	30.4	117.2	0.000	1.9	0.20	0.2	0	64	40.0	132.9	0.038	2.6	0.37	0.038	2.6	0.4
6Paper and wood production and processing_ID_PP		5	19	22.2	30.3	1.540	54.3	8.50	8.5	1.54	18	23.4	30.7	3.180	54.9	8.50	3.2	54.9
8Food	Animal products FDM_ME	1	56.1	0.0	56.100	56.1	56.10	0.5	56.1	1	56.1	0.0	56.100	56.1	56.10	56.1	56.1	56.1

Main activity of the facility	Normal statistics results								Modified statistics results								
	a)			20%	80%		b)		a)			20%	80%			b)	
	n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
Milk and cheese products FDM_MI	8	0.9	1.0	0.240	1.1	0.45	0.5	0.24	8	0.9	1.0	0.240	1.1	0.45	0.2	1.1	0.5
Vegetable products FDM_V	16	14.0	36.9	0.100	10.3	2.20	2.2	0.1	14	16.1	39.0	0.680	11.4	3.15	0.7	11.4	3.2
9Other activities using chemicals_ID_STS	9	20	1.0	1.4	0.000	1.7	0.35	0.4	0	12	1.7	1.5	0.620	1.9	1.55	0.6	1.9

*Table 16 - Statistics and suggested **Sectorial Emission Factor (load)** Copper concentration*
a) Number of measured facilities
b) ID_ps_E_HM_CU = Estimated annual metal discharged loads [kg/year]
c) Instead of measured 0 for VAR3 and VAR4

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
1Energy sector	Combustions_ LCP + waste incinerations_WI	209	1169.5	4852.8	0.000	115.6	0.60	0.6	0	166	1472.4	5404.1	0.100	226.1	1.75	0.1	226.1	1.8
	refinaries_REF	11	90.3	82.2	1.100	153.0	98.60	98.6	1.1	11	90.3	82.2	1.100	153.0	98.60	1.1	153	98.6
2Metal ind.	Production of ferrous metals__SF_FMP	81	170.0	352.7	0.079	208.0	6.30	6.4	0.08	80	172.1	368.0	0.109	215.6	6.35	0.1	215.6	6.4
	Processing of metals STMGSTS	75	0.8	1.9	0.015	0.6	0.20	0.2	0.02	62	0.9	2.1	0.100	0.8	0.37	0.1	0.8	0.4
	Production and processing of non-ferrous metals_NFM	64	66.6	222.5	0.044	9.2	0.68	0.7	0.04	55	77.5	238.2	0.209	11.4	1.47	0.2	11.4	1.5
3Mineral industry	cement_CLM	18	9.3	11.8	0.080	17.7	3.55	3.6	0.08	14	11.9	12.1	1.340	22.0	5.50	1.3	22	5.5
	ceramics_CER	62	1.1	3.4	0.000	0.9	0.10	0.1	0	48	1.4	3.8	0.100	1.2	0.30	0.1	1.2	0.3
	glass_GLS	19	4.2	9.7	0.100	0.6	0.30	0.3	0.1	18	4.4	10.0	0.100	0.7	0.30	0.1	0.7	0.3
	mining_open_MIN_O	55	9855.5	23091.3	26.420	13727.8	287.00	287	26.42	54	10038.0	23264.8	29.240	13795.8	484.75	29.2	13795.8	484.8

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
	mining_underground_MIN_U	72	64.7	134.7	1.140	69.8	6.25	6.3	1.14	70	66.6	136.2	1.380	72.4	6.90	1.4	72.4	6.9
4Chemical industry	general_CWW	41	61.5	224.8	0.000	39.0	4.03	4	0	31	81.3	255.4	0.236	54.1	12.64	0.2	54.1	12.6
	fertilizers_ID_LVIC_NP	6	175.5	94.9	92.500	283.6	147.80	147.8	92.5	6	175.5	94.9	92.500	283.6	147.80	92.5	283.6	147.8
	biocid_OFC_BIOCID	25	0.2	0.4	0.000	0.1	0.00	0	0	7	0.5	0.6	0.100	1.3	0.10	0.1	1.3	0.1
	explosives_OFC_EX	7	668.8	1020.2	38.140	800.3	178.40	178.4	38.14	6	780.3	1061.8	54.300	883.1	323.70	54.3	883.1	323.7
	pharma_OFC_Pharma	12	10.2	20.4	0.029	4.5	2.29	2.3	0.03	12	10.2	20.4	0.029	4.5	2.29	0.029	4.5	2.3
	polimers_POL	3	7.6	1.6	6.455	8.7	7.11	7.1	6.46	3	7.6	1.6	6.455	8.7	7.11	6.5	8.7	7.1
5Waste manage-	animal_WTL_Animal	11	0.5	0.5	0.100	0.9	0.10	0.1	0.1	9	0.6	0.5	0.100	0.9	0.70	0.1	0.9	0.7
	hazardous_ID_WTL_H	5	1.8	0.3	1.538	2.0	1.81	1.8	1.54	5	1.8	0.3	1.538	2.0	1.81	1.5	2	1.8
	non hadordous_ID_WTL_NH	82	38.8	166.8	0.000	0.6	0.10	0.1	0	57	55.9	197.7	0.049	0.9	0.22	0.05	0.9	0.2
6Paper and wood production and processing_ID_PP		5	45	48.5	78.6	0.100	46.2	18.50	18.5	0.1	37	59.0	83.0	1.795	80.0	25.00	1.8	80
8Food	Animal products FDM_ME	0						6.9	3.04	0						3	11.6	6.9

Main activity of the facility	Normal statistics results								Modified statistics results								
	a)			20%	80%		b)		a)			20%	80%			b)	
	n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
Milk and cheese products FDM_MI	8	8.5	6.4	3.040	11.6	6.90	6.9	3.04	8	8.5	6.4	3.040	11.6	6.90	3	11.6	6.9
Vegetable products FDM_V	20	30.3	82.5	0.860	23.0	5.30	5.3	0.86	19	31.9	84.4	1.260	26.0	6.30	1.3	26	6.3
9Other activities using chemicals_ID_STS	9	29	5.8	17.2	0.160	3.8	1.70	1.7	0.16	28	6.0	17.5	0.240	4.1	1.70	0.2	4.1

*Table 17 - Statistics and suggested **Sectorial Emission Factor (load)** Nickel concentration*
a) Number of measured facilities
b) ID_ps_E_HM_NI = Estimated annual metal discharged loads [kg/year]
c) Instead of measured 0 for VAR3 and VAR4

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
1Energy sector	Combustions_ LCP + waste incinerations_WI	213	257.1	969.9	0.000	5.9	0.10	0.1	0	147	372.5	1149.0	0.100	117.6	0.40	0.1	117.6	0.4
	refinaries_REF	32	109.9	85.6	34.440	198.4	76.60	76.6	34.44	32	109.9	85.6	34.440	198.4	76.60	34.4	198.4	76.6
2Metal ind.	Production of ferrous metals__SF_FMP	90	83.9	211.6	0.025	78.0	27.26	27.9	0.02	82	92.1	226.0	0.157	87.2	30.96	0.2	87.2	31
	Processing of metals STMGSTS	95	1.4	4.4	0.000	1.4	0.15	0.2	0	72	1.9	4.9	0.103	1.8	0.50	0.1	1.8	0.5
	Production and processing of non-ferrous metals_NFM	62	5.2	13.9	0.079	1.8	0.40	0.4	0.08	53	6.0	14.8	0.220	2.4	0.50	0.2	2.4	0.5
3Mineral industry	cement_CLM	26	4.8	5.7	1.200	6.1	2.35	2.4	1.2	25	5.0	5.7	1.280	6.7	2.50	1.3	6.7	2.5
	ceramics_CER	52	0.5	1.1	0.000	0.4	0.10	0.1	0	39	0.7	1.2	0.100	0.5	0.30	0.1	0.5	0.3
	glass_GLS	9	0.5	0.5	0.200	0.8	0.20	0.2	0.2	9	0.5	0.5	0.200	0.8	0.20	0.2	0.8	0.2
	mining_open_MIN_O	42	54.1	90.2	0.300	77.9	9.35	9.4	0.3	39	58.3	92.3	0.300	87.3	12.10	0.3	87.3	12.1

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
	mining_underground_MIN_U	42	17.5	27.0	0.104	29.5	6.60	6.6	0.1	38	19.3	27.8	0.298	31.3	9.45	0.3	31.3	9.5
4Chemical industry	general_CWW	44	48.8	105.7	0.000	41.2	3.52	3.5	0	29	74.0	122.8	2.686	59.9	21.15	2.7	59.9	21.2
	fertilizers_ID_LVIC_NP	6	85.7	50.2	30.500	126.5	78.95	79	30.5	6	85.7	50.2	30.500	126.5	78.95	30.5	126.5	79
	biocid_OFC_BIOCID	18	0.0	0.0	0.000	0.0	0.00	0	0	0						0.1	32.4	1.4
	explosives_OFC_EX	5	34.4	33.8	6.240	45.2	31.60	31.6	6.24	5	34.4	33.8	6.240	45.2	31.60	6.2	45.2	31.6
	pharma_OFC_Pharma	17	17.5	29.0	0.103	32.4	1.40	1.4	0.1	17	17.5	29.0	0.103	32.4	1.40	0.1	32.4	1.4
	polimers_POL	3	7.6	1.6	6.455	8.7	7.11	7.1	6.46	3	7.6	1.6	6.455	8.7	7.11	6.5	8.7	7.1
5Waste manage-	animal_WTL_Animal	7	0.3	0.4	0.020	0.4	0.20	0.2	0.02	5	0.4	0.4	0.180	0.5	0.20	0.2	0.5	0.2
	hazardous_ID_WTL_H	7	5.4	12.5	0.159	0.4	0.21	0.2	0.16	7	5.4	12.5	0.159	0.4	0.21	0.2	0.4	0.2
	non hadordous_ID_WTL_NH	83	30.3	118.0	0.000	0.5	0.10	0.1	0	63	39.9	134.0	0.055	0.8	0.26	0.1	0.8	0.3
6Paper and wood production and processing_ID_PP		5	27	62.2	72.3	9.740	119.6	18.40	18.4	9.74	26	64.6	72.6	9.900	123.0	18.60	9.9	123
8Food	Animal products FDM_ME	1	33.6	0.0	33.600	33.6	33.60	2.2	33.6	1	33.6	0.0	33.600	33.6	33.60	33.6	33.6	33.6

Main activity of the facility	Normal statistics results								Modified statistics results								
	a)			20%	80%		b)		a)			20%	80%			b)	
	n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
Milk and cheese products FDM_MI	8	2.6	1.5	1.360	3.4	2.20	2.2	1.36	8	2.6	1.5	1.360	3.4	2.20	1.4	3.4	2.2
Vegetable products FDM_V	19	4.7	6.3	0.260	7.8	2.00	2	0.26	18	5.0	6.3	0.420	8.6	2.30	0.4	8.6	2.3
9Other activities using chemicals_ID_STS	9	25	17.0	22.5	0.000	28.2	10.20	10.2	0	19	22.3	23.4	4.920	28.3	24.90	4.9	28.3

*Table 18 - Statistics and suggested **Sectorial Emission Factor (load)** Lead concentration*
a) Number of measured facilities
b) ID_ps_E_HM_PB = Estimated annual metal discharged loads [kg/year]
c) Instead of measured 0 for VAR3 and VAR4

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%		c)	b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median		VAR3	VAR4
1Energy sector	Combustions_ LCP + waste incinerations_WI	217	258.5	1047	0.000	3.1	0.20	0.2	0	164	342.1	1192.6	0.100	11.9	0.45	0.1	11.9	0.5
	refinaries_REF	28	114.0	74.5	54.360	189.2	112.00	112	54.36	28	114.0	74.5	54.360	189.2	112.00	54.4	189.2	112
2Metal ind.	Production of ferrous metals__SF_FMP	71	349.6	915.7	0.100	218.0	26.20	26.3	0.1	70	354.6	929.3	0.100	219.0	26.32	0.1	219	26.3
	Processing of metals STMGSTS	66	1.4	4.4	0.004	0.4	0.10	0.1	0.0039	54	1.7	4.8	0.076	0.8	0.10	0.1	0.8	0.1
	Production and processing of non-ferrous metals_NFM	65	21.0	51.8	0.201	18.4	0.80	0.8	0.2	55	24.8	55.4	0.412	20.4	1.26	0.4	20.4	1.3
3Mineral industry	cement_CLM	17	2.9	6.0	0.000	3.7	0.30	0.3	0	11	4.4	7.0	0.300	4.3	1.80	0.3	4.3	1.8
	ceramics_CER	52	2.2	5.9	0.044	1.3	0.49	0.5	0.04	42	2.7	6.4	0.307	1.5	0.65	0.3	1.5	0.7
	glass_GLS	27	0.6	0.6	0.100	1.4	0.30	0.3	0.1	27	0.6	0.6	0.100	1.4	0.30	0.1	1.4	0.3
	mining_open_MIN_O	38	134.7	495.7	0.200	78.4	17.45	17.5	0.2	34	150.5	521.7	0.480	87.0	24.20	0.5	87	24.2

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
	mining_underground_MIN_U	56	135.6	294.9	0.500	150.0	9.55	9.6	0.5	54	140.6	299.1	0.679	156.6	10.80	0.7	156.6	10.8
4Chemical industry	general_CWW	32	13.5	29.2	0.009	4.1	0.45	0.4	0.01	25	17.2	32.0	0.173	17.1	1.67	0.2	17.1	1.7
	fertilizers_ID_LVIC_NP	4	99.5	38.8	82.660	123.3	116.95	117	82.66	4	99.5	38.8	82.660	123.3	116.95	82.7	123.3	117
	biocid_OFC_BIOCID	27	0.0	0.0	0.000	0.0	0.00	0	0	5	0.1	0.0	0.100	0.1	0.10	0.1	0.1	0.1
	explosives_OFC_EX	5	9.0	2.8	6.360	11.1	9.50	9.5	6.36	5	9.0	2.8	6.360	11.1	9.50	6.4	11.1	9.5
	pharma_OFC_Pharma	6	12.0	22.6	0.007	10.2	0.02	0	0.007	6	12.0	22.6	0.007	10.2	0.02	0.01	10.2	0.02
	polimers_POL	4	6.3	2.7	4.531	8.2	6.56	6.6	4.53	4	6.3	2.7	4.531	8.2	6.56	4.5	8.2	6.6
5Waste manage-	animal_WTL_Animal	7	0.1	0.1	0.020	0.2	0.10	0.1	0.02	5	0.2	0.1	0.100	0.2	0.10	0.1	0.2	0.1
	hazardous_ID_WTL_H	9	0.0	0.1	0.000	0.0	0.00	0	0	2	0.1	0.1	0.111	0.2	0.14	0.1	0.2	0.1
	non hadordous_ID_WTL_NH	80	31.0	120.1	0.000	0.2	0.01	0.01	0	50	49.6	148.9	0.012	0.4	0.10	0.01	0.4	0.1
6Paper and wood production and processing_ID_PP		5	38	37.5	35.3	7.320	70.2	24.50	24.5	7.32	36	39.6	35.1	10.100	71.0	24.90	10.1	71
8Food	Animal products FDM_ME	1	112.1	0.0	112.100	112.1	112.10	0.5	112.1	1	112.1	0.0	112.100	112.1	112.10	112.1	112.1	112.1

Main activity of the facility	Normal statistics results								Modified statistics results								
	a)			20%	80%		b)		a)			20%	80%			b)	
	n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
Milk and cheese products FDM_MI	9	1.0	1.6	0.200	0.9	0.50	0.5	0.2	9	1.0	1.6	0.200	0.9	0.50	0.2	0.9	0.5
Vegetable products FDM_V	20	4.5	5.0	0.280	7.6	2.75	2.8	0.28	18	5.0	5.0	0.720	8.7	3.10	0.7	8.7	3.1
9Other activities using chemicals_ID_STS	9	29	1.9	3.4	0.000	2.4	0.40	0.4	0	21	2.7	3.8	0.200	3.0	1.10	0.2	3

*Table 19 - Statistics and suggested **Sectorial Emission Factor (load)** Zinc concentration*
a) Number of measured facilities
b) ID_ps_E_HM_ZN = Estimated annual metal discharged loads [kg/year]
c) Instead of measured 0 for VAR3 and VAR4

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
1Energy sector	Combustions_ LCP + waste incinerations_WI	211	2659	8358	0.100	424.1	1.00	1	0.1	175	3206	9082	0.200	1774	4.03	0.2	1774	4
	refinaries_REF	33	261.6	303.0	0.000	338.4	205.0	205	0	25	345.3	303.7	150.97	522.2	268.00	151	522.2	268
2Metal ind.	Production of ferrous metals__SF_FMP	120	1582	3614	0.409	1714	136.8	139	0.41	120	1582	3649	0.409	1714	136.78	0.4	1714	136.8
	Processing of metals STMGSTS	101	9.1	24.1	0.005	8.0	1.11	1.1	0.005	82	11.3	26.3	0.300	14.3	1.95	0.3	14.3	2
	Production and processing of non-ferrous metals_NFM	77	73.0	162.7	0.200	102.2	2.70	2.7	0.2	64	87.9	174.8	0.598	129.2	6.83	0.6	129.2	6.8
3Mineral industry	cement_CLM	18	13.4	14.4	0.160	25.9	9.10	9.1	0.16	14	17.2	14.2	3.280	28.6	13.45	3.3	28.6	13.5
	ceramics_CER	70	6.0	10.4	0.000	12.0	0.85	0.9	0	55	7.6	11.2	0.300	13.8	1.95	0.3	13.8	1.9
	glass_GLS	26	2.4	3.1	0.700	3.7	1.30	1.3	0.7	26	2.4	3.1	0.700	3.7	1.30	0.7	3.7	1.3
	mining_open_MIN_O	56	2483	5544	9.300	2481	121.3	121	9.3	55	2529	5584	10.020	2491	132.10	10	2491	132.1

	Main activity of the facility	Normal statistics results								Modified statistics results								
		a)			20%	80%		b)		a)			20%	80%			b)	
		n	mean	stdev.	percentile		median	VAR1	VAR2	n	mean	stdev	percentile		median	c)	VAR3	VAR4
	mining_underground_MIN_U	97	93.0	262.6	0.920	115.2	7.70	7.7	0.92	94	96.0	266.2	1.000	124.8	8.80	1	124.8	8.8
4Chemical industry	general_CWW	54	304.4	684.0	0.000	163.5	59.81	59.8	0	42	391.4	753.3	3.232	179.9	112.48	3.2	179.9	112.5
	fertilizers_ID_LVIC_NP	6	739.3	462.0	348.0	969.2	714.7	715	348	6	739.3	462.0	348.0	969.2	714.70	348	969.2	714.7
	biocid_OFC_BIOCID	27	0.0	0.1	0.000	0.0	0.00	0	0	5	0.2	0.0	0.180	0.2	0.20	0.2	0.2	0.2
	explosives_OFC_EX	7	1463	1975	213	2277	364	364	213.3	7	1463	1975	213.3	2277	364.00	213	2277	364
	pharma_OFC_Pharma	9	34.2	85.8	0.08	8.4	6.03	6	0.08	9	34.2	85.8	0.079	8.4	6.03	0.1	8.4	6
	polimers_POL	17	1111	2162	0.000	2280.0	0.00	0	0	8	2361.0	2642.0	21.459	4369.8	1450.86	21.5	4369.8	1450.9
5Waste manage-ment	animal_WTL_Animal	11	1.2	1.2	0.100	2.4	0.60	0.6	0.1	9	1.4	1.1	0.100	2.6	1.60	0.1	2.6	1.6
	hazardous_ID_WTL_H	5	2.8	1.4	1.878	3.2	2.01	2	1.88	5	2.8	1.4	1.878	3.2	2.01	1.9	3.2	2
	non hadordous_ID_WTL_NH	82	0.7	1.3	0.008	0.7	0.20	0.2	0.0077	66	0.8	1.4	0.100	1.2	0.30	0.1	1.2	0.3
6Paper and wood production and processing_ID_PP		5	52	196.2	301.2	2.343	326.8	42.60	42.6	2.34	51	200.0	302.9	2.514	336.9	43.20	2.5	336.9
8Food	Animal products FDM_ME	0						11.5	5.62	0						5.6	16	11.5

Main activity of the facility	Normal statistics results								Modified statistics results								
	a)						b)		a)							b)	
	n	mean	stdev.	20% percentile	80% percentile	median	VAR1	VAR2	n	mean	stdev	20% percentile	80% percentile	median	c)	VAR3	VAR4
Milk and cheese products FDM_MI	8	11.1	6.3	5.620	16.0	11.50	11.5	5.62	8	11.1	6.3	5.620	16.0	11.50	5.6	16	11.5
Vegetable products FDM_V	20	66.1	125.5	1.656	39.5	17.70	17.7	1.66	19	69.6	127.8	2.260	41.1	17.90	2.3	41.1	17.9
9Other activities using chemicals_ID_STS	9	32	16.0	31.1	0.100	17.7	3.00	3	0.1	26	19.7	33.4	1.700	18.8	4.80	1.7	18.8

PFOS and PFOA

The EU and national databases do not contain any PFOA or PFOS measurements, only some other related parameters, such as PFCs to the air. Therefore we need to use concentration values from literature (PROMISCES monitoring and IWR inventory 2024) and multiply them by discharges (Q). Where the Q was not available, there the sectorial median Q value was used.

variable	variant	value	unit
ID_CONC_PFAS_PFOA	1	2.13	ng/l
ID_CONC_PFAS_PFOA	2	0.15	ng/l
ID_CONC_PFAS_PFOA	3	23.3	ng/l
ID_CONC_PFAS_PFOS	1	3.5	ng/l
ID_CONC_PFAS_PFOS	2	1.33	ng/l
ID_CONC_PFAS_PFOS	3	23	ng/l

4.2.2 Industrial emission into the air

Atmospheric deposition from the chimneys of the industrial facilities based on the discharged metal load, the calculations are described in section 2.2.5. In this section we described how was the estimation of the ID_AD_ps_HM_i = individual industrial point source emitters of HM.

- We add industrial emissions only to each ‘large combustion plans’ and ‘waste incinerations’.
- In the case of measured atmospheric emissions (reported in the E-PRTR), we naturally use the measured results.
- At those facilities where there was no measurement in the given year, but there was in other years, the average of the measurement results is considered.
- Where no emission load could be determined based on the above, the 'AD Emission Factor' was applied, which was estimated using the following method.

'AD Emission Factor’

If metal and/or PFOS/PFOA emissions were not measured, an emission factor must be applied. Several studies were performed, one of them is the most suitable for load estimation by EF: ‘Estimate metal emissions to air based on NOx emissions to air with corrections’.

The NOx emissions are usually well measured in all facilities, and it has relations to dust cleaning technologies and the total volume of the dust emitted by the facility, but not for all metals. (See the following figures.) Other corrections were made for lower and upper borders:

- If the estimated metal emission loads based on NOx were higher than median value of the measurements, then use median value.
- If the estimated metal emission loads based on NOx were lower than 5% percentile of measurements/10, then use 5% percentile of measurements/10 value.
- Otherwise use estimated metal emission loads based on NOx by the following EFs.

Table 20 - Statistics of heavy metal measurements and their correlation with Nox measurements.

pollu-tant	Statistics of measurements		Statistics of the correlation of the measurements to the emitted NOx		
	Median of the measure-ments	5% percentile of measure-ments	EF (kg metal /y /kg emitted NOx)	R^2 of the EF correla-tion	Number of facilities used for corr.
Cr	202	102	0.1117	0.6532	65
Pb	538	230.8	2.0936	0.7445	95
Cd	21.4	10.7	0.02	0.7041	130
Ni	117	52.77	0.1312	0.5173	168
Zn	721	228.1	1.1913	0.6147	111
Cu	198	110	0.1615	0.6227	56
As	46.9	21.15	0.0511	0.324	93

4.3 DERIVING SOIL CONCENTRATIONS

The soil metal data were collected for different types of areas: agricultural, natural vegetation, rocks and glaciers. Land use data was extracted from the CORINE Land Cover dataset (https://land.copernicus.eu/en/products/corine-land-cover/clc2018), using the latest 2018 data. The following tables show the land use categories for each type, starting with agricultural land (Table 21), then continuing with other land uses (Table 22). QGIS software was used to perform geospatial tasks. After loading the data, the appropriate land use records were selected by selection and exported to new layers for future work. The figures shows the different land uses in the areas of the river basins (the light blue color indicates the river basin area; the other colors in each figure indicate the corresponding category of the land use).

Table 21 - CORINE Land Cover Agriculture categories

Label	Name	Value
211	Non-irrigated arable land	12
212	Permanently irrigated land	13
213	Rice fields	14
221	Vineyards	15
222	Fruit trees and berry plantations	16
223	Olive groves	17
241	Annual crops associated with permanent crops	19
242	Complex cultivation patterns	20
243	Land principally occupied by agriculture with significant areas of natural vegetation	21
244	Agro-forestry areas	22

Table 22 - CORINE Land Cover Natural vegetation, rocks, glaciers categories

Label	Name	Value
	Natural vegetation	
311	Broad-leaved forest	23
312	Coniferous forest	24
313	Mixed forest	25
321	Natural grasslands	26
322	Moors and heathland	27
323	Sclerophyllous vegetation	28

Label	Name	Value
324	Transitional woodland-shrub	29
	Rock	
332	Bare rocks	31
333	Sparsely vegetated areas	32
	Glaciers	
335	Glaciers and perpetual snow	34

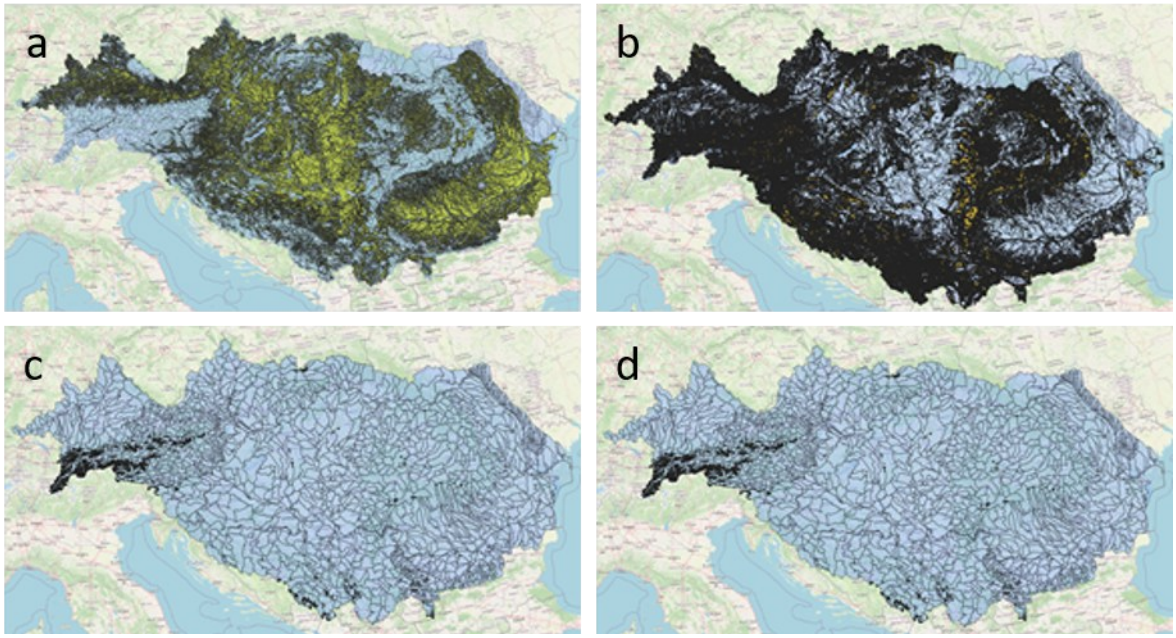

Figure 31 - Land use coverage in the DRB (a Agricultural land use, b Natural vegetation land use, c Rocks land use, d Glaciers land use

The next step involves pairing the different categories of land use with the analytical units. By spatial associating the selected soil types to the identifier of the analytical units, the result is a merged polygon per identifier, where each soil type is linked to an analytical unit based on the best fit. In Figure 32, we provide a closer view of several analytical units, shown in blue. The land use area of that category is marked in orange, of which the highlighted yellow land use area belongs to the overlapping analytical unit.

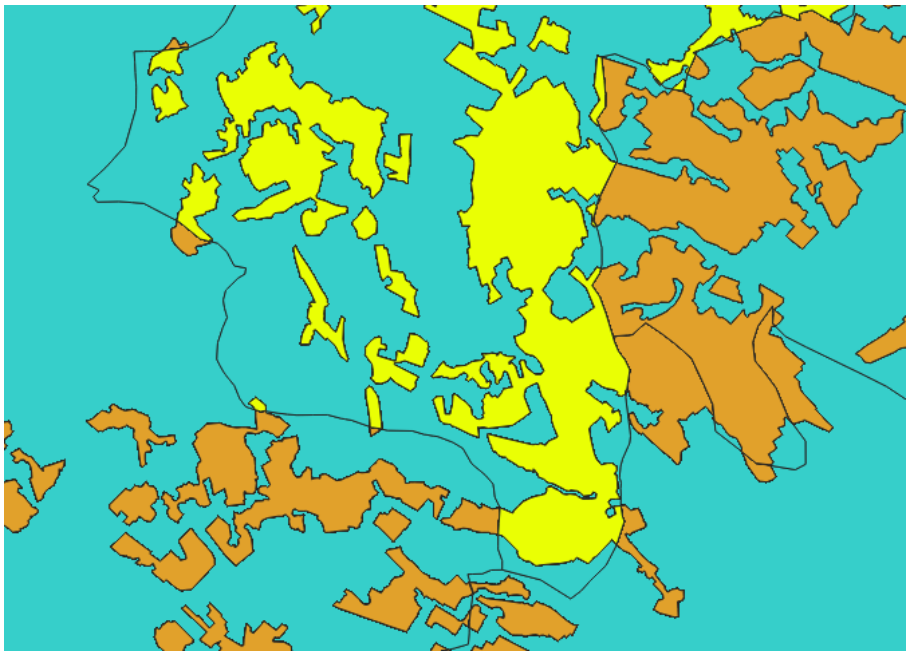

Figure 32 - Land use connected to the analytical units

The soil metals analysed were arsenic, cadmium, chromium, copper, nickel, lead and zinc. Several data sources were used, point and surface data from the GEMAS dataset (https://gemas.eurogeosurveys.org/) were available in vector format, the JRC data were in raster format.

The GEMAS (Geochemical Mapping of Agricultural soils of Europe) provide data on metals in arable and pasture soils in two categories. Among the downloadable data, we can distinguish between different types, both in terms of sampling (e.g. ICP-MS after aqua regia extraction) and in terms of the number of displayed levels. The number of displayed levels of soil metal values is divided into 7 or 72 categories and thus determines the level of detail of the data, for the calculation we used the larger level data, because it contained more detailed metal data due to more categories. The category values represent ranges, the average of which is calculated. Using the resulting numerical value, a raster conversion was performed. This provided the input for calculating the Zonal Statistics. Zonal Statistics gives different statistical values of raster data calculated on a vector layer (the layer of analytical units). After performing the Zonal Statistics, the result was generated, which in our case included the average value for the metal of interest within each analytical units.

The JRC ESDAC LUCAS (Joint Research Centre, European Soil Data Centre, Land Use/Cover Area frame statistical Survey) data is the result of a collaborative work of European Statistical Office (EUROSTAT) and the Directorate General responsible for Agriculture and the technical support of the JRC[69]. The data is available a raster file (.tif) with 1000 m resolution, which was processed without any conversion. Another modified data set for arsenic was considered, which was derived as the previous had high detection limits[70]. The two databases overlapped for most of the analytical units, but there were areas where only one or the other type of data was available. From the JRC raster data, zonal statistics values were calculated for each of the metals tested, based on the analytical units, using the above method.

The above procedures gave the metal content for each analytical unit from the two data sources. Neither of the JRC - LUCAS and GEMAS data covered the entire Danube river basin, and therefore it was necessary to examine the correlation of the data in order to assign a suitable value to each analytical

[69] https://esdac.jrc.ec.europa.eu/projects/lucas

[70] https://www.sciencedirect.com/science/article/pii/S0160412024001302?via%3Dihub

unit. We analysed the data using Excel. The values for each analytical unit were plotted from the two sets of source data, and the relationship between them was examined by fitting linear and polynomial trend lines (Figure 33).

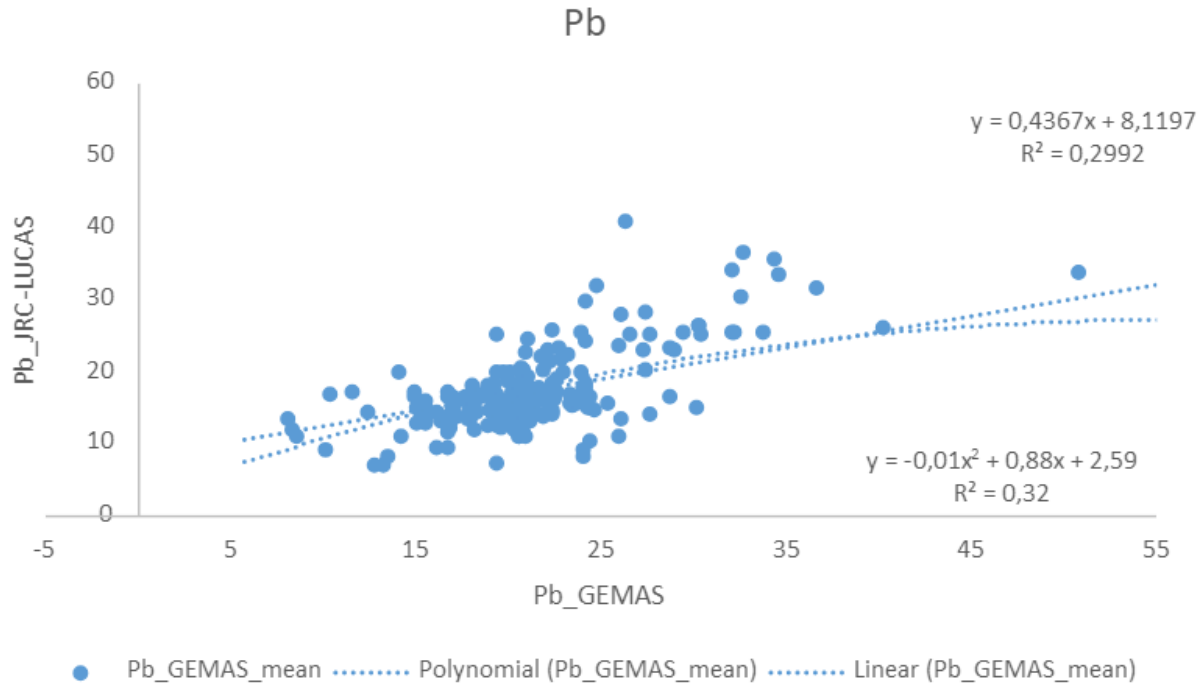

Figure 33 - Plots of lead content from the two data sources with fitted linear and polynomial trend lines

There were some differences in the results obtained from the two data sets, which can be attributed to variations in sampling points, sampling methods, or other factors. For the assignment of the final values, priority was given to the data from the JRC-LUCAS data source, as it is regarded as the most reliable and widely accepted official data source within the EU, using much higher soil sample density as the GEMAS program. For arsenic, where no JRC values were available, the modified data source was applied. Where a reliable correlation between the two data sets was identified, the GEMAS values were adjusted using the derived equation.

Deriving soil concentrations for PFAS substances

Based on various observations, PFAS contamination has been found to correlate with economic potential, expressed as gross domestic product (GDP), which partially reflects population density. For GDP calculations within the European Union, data from the Eurostat NUTS database were used[71], while for non-EU countries, national GDP data were retrieved from the World Bank[72].

The NUTS (Nomenclature of Territorial Units for Statistics) classification divides the territory of the European Union into statistical regions at multiple hierarchical levels[73]. To improve spatial accuracy, the smallest available regional level was used, which corresponded to the NUTS 3 level in most cases.

GDP data were available at the NUTS regional level in units of € / year. During data processing, GDP values assigned to NUTS regions were spatially distributed proportionally to analytical units based on spatial extent. Subsequently, GDP values were normalized by the area of each analytical unit, resulting

[71] https://ec.europa.eu/eurostat/databrowser/view/nama_10r_3gdp/default/table
[72] https://data.worldbank.org/indicator/NY.GDP.MKTP.CDend=2023&locations=UA&start=2011&view=chart
[73] https://ec.europa.eu/eurostat/web/nuts

in GDP values expressed per square meter. The investigated period covered the years from 2015 to 2020, and the final analysis was performed using the six-year average GDP values.

For certain regions, including Ukraine, Moldova, and Bosnia and Herzegovina, NUTS regional data were not available. In these cases, national annual GDP values were distributed proportionally across analytical units based on surface area. For Kosovo, GDP values were adopted from adjacent regions. All GDP values were ultimately converted to units of € / m² through normalization by area.

The results indicate that several areas—typically associated with large urban centers—exhibited exceptionally high GDP values, which caused distortion in the spatial patterns. To mitigate this effect, a threshold adjustment was applied: all GDP values exceeding 5 €/m² were reduced to this upper limit, effectively eliminating distortions. (Figure 34)

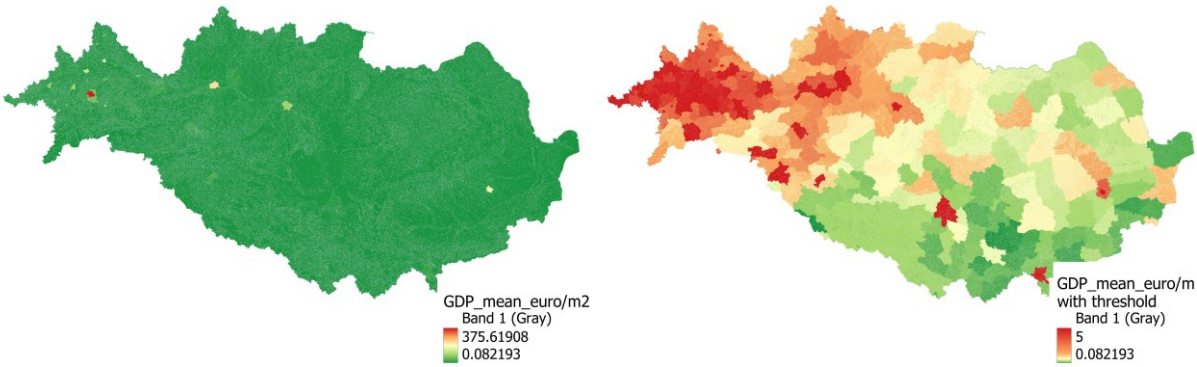

Figure 34 - GDP values in € / m² before and after thresholding

In the land-use analysis, four primary categories were distinguished: agricultural areas, forests, grasslands, and other land types. These categories were defined according to the CORINE Land Cover database. Empirical relationships between GDP and soil concentrations of PFOA and PFOS were established based on Table 23, where values are results from analysis of Kardos et al.,2024.

Table 23 - Empirical relationships between GDP and soil concentrations of PFOA and PFOS

	Land use			
Substance	**Forest areas**	**Grasslands**	**Agricultural area**	**Other**
PFOA	Y = 0.16 x	Y = 0.14 x	Y = 0.051 x	Y = 0.10 x
PFOS	Y = 0.085 x	Y = 0.070 x	Y = 0.032 x	Y = 0.060 x

where

Y represents the soil concentration of PFOA or PFOS in µg/kg DM, and
x represents the GDP characteristic of the area in € / year / m².

All calculations were performed using a raster calculator, after which the results were exported for further analysis. Substance-specific aggregation was also carried out to improve visual interpretability. (Figure 35and Figure 36) Values associated with analytical units were derived using zonal statistical methods.

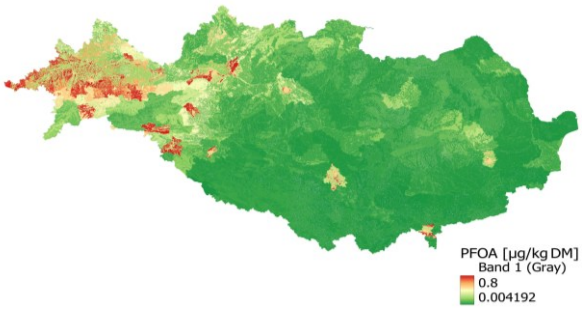

Figure 35 - Soil concentration of PFOA in µg/kg DM

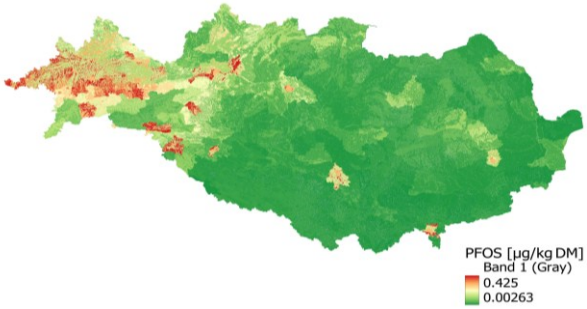

Figure 36 - Soil concentration of PFOS in µg/kg DM

4.4 SURFACE RUNOFF

Runoff concentrations are based on the concentrations of soils or suspended sediments, and the values are given based on the relevant solid phase-water phase distribution coefficients (Table 24).

Table 24 - $K_{sed,w}$ for heavy metals (Allison & Allison, 2005). Values reported are representing the median values of a thorough review of reported values of over hundreds of studies.

Compounds	**Log $K_{d,SS/water}$ (L/kg)**	**Log $K_{d,soil/water}$ (L/kg)**
Arsenic	4.0 (2.0-6.0)	3.4 (0.3-4.3)
Cadmium	4.7 (2.8-6.3)	2.9 (0.1-5.0)
Chromium	5.1 (3.9-6)	3.9 (1-4.7)
Copper	4.7 (3.1-6.1)	2.7 (0.1-3.6)
Nickel	4.6 (3.5-5.7)	3.1 (1-3.8)
Lead	5.6 (3.4-6.5)	4.2 (0.7-5.0)
Zinc	5.1 (3.5-6.9)	3.1 (-1.0-5.0)

In the calculation of surface runoff concentration, the concentration of either the original soil or the transported and enriched soil is transformed to water concentration. For the latter, this means that the relationship is related to suspended matter carried by the water and not to soils. Therefore, the Log $K_{soil/water}$ values are used when deriving from soils (EQ1) and Log $K_{d\ SS/water}$ values (EQ2) are used form the transformation for HM concentration to water from suspended sediment concentration.

SR_CONC_CALC_HM = ER_agrl_CONT_SOIL_top_AL_HM/ KD_SOIL_WATER_HM EQ(1)

SR_CONC_CALC_HM = ER_agrl_CONT_SOIL_top_AL_HM*ER_ENR_HM/ KD_SED_WATER_HM EQ(2)

Kd values for PFAS is of course also different for soils and suspended sediments in runoff (ICTR - 5.3.4.2 Surface Water/Sediment Interactions, Mussabek, 2019, 2020). Long chain PFAS compounds have higher adsorption potential. For PFAS there are no trustable information on the enrichment of the substances during transport, (Jolankai et al., 2026) highlighted that there are no clear enrichment patterns for these substances. For this reason, runoff is estimated directly from soil concentration and is used as it were the concentration of the suspended sediment also. The Log $K_{oc,\ SPM}$ values are used for both modelled PFAS (Table 25).

Table 25 - Koc,SPM,soil for PFAS -K_{OC}: ((ITRC), The Interstate Technology & Regulatory Council, 2023), Kd: ***(Mussabek et al., 2020)***

Compounds	Log $K_{oc,\ SPM}$ (L/kg)	Log $K_{oc,\ soil}$ (L/kg)	Log $K_{d,SS/water}$ (L/kg)
PFOA	3.17 (2.44 - 3.68)	1.98-2.92 2.9*	0.6
PFOS	4.3 (3.8-4.8)	2.4-2.6 3.6*[74]	3.3[75]

Surface runoff concentration for PFAS is calculated using the following equation.

SR_CONC_CALC_PFAS = ER_agrl_CONT_SOIL_top_AL_PFAS / KD_OC_PFAS / GW_CONC_CALC_FOC_PFAS * SR_FCT_CALIB_PFAS

Where

ER_agrl_CONT_SOIL_top_AL_PFAS is the PFAS concentration of soils (estimated),

KD_OC_PFAS is the partition coefficient of PFAS between organic carbon and water,

GW_CONC_CALC_FOC_PFAS is the fraction of organic carbon in soils and

SR_FCT_CALIB_PFAS is the calibration factor for the surface runoff concentration estimation.

4.5 ATMOSPHERIC DEPOSITION

4.5.1 Heavy metals

The EMEP programme is a Co-operative Programme for Monitoring and Evaluation of the Long-range Transmission of Air Pollutants in Europe. This programme is focusing on substances causing acidification and eutrophication, but later it also started to model the aerial concentration of persistent organic pollutants, heavy metals, and particulate matter. The modelling development for heavy metals and POPs is the responsibility of the Meteorological Synthesizing Centre -East (MSC-E). The heavy metal deposition maps are not available as downloadable grids, therefore the model centre has been contacted for the data.

Total annual emission data is available for Cd and Pb (Figure 37) for whole of Europe in a gridded format, with 0.1 arcmin resolution from 2015 to 2022. These serve as input for the EMEP model, which calculates long range transport.

[74] (Hunter Anderson et al., 2019)

[75] Mussabek et al., 2020

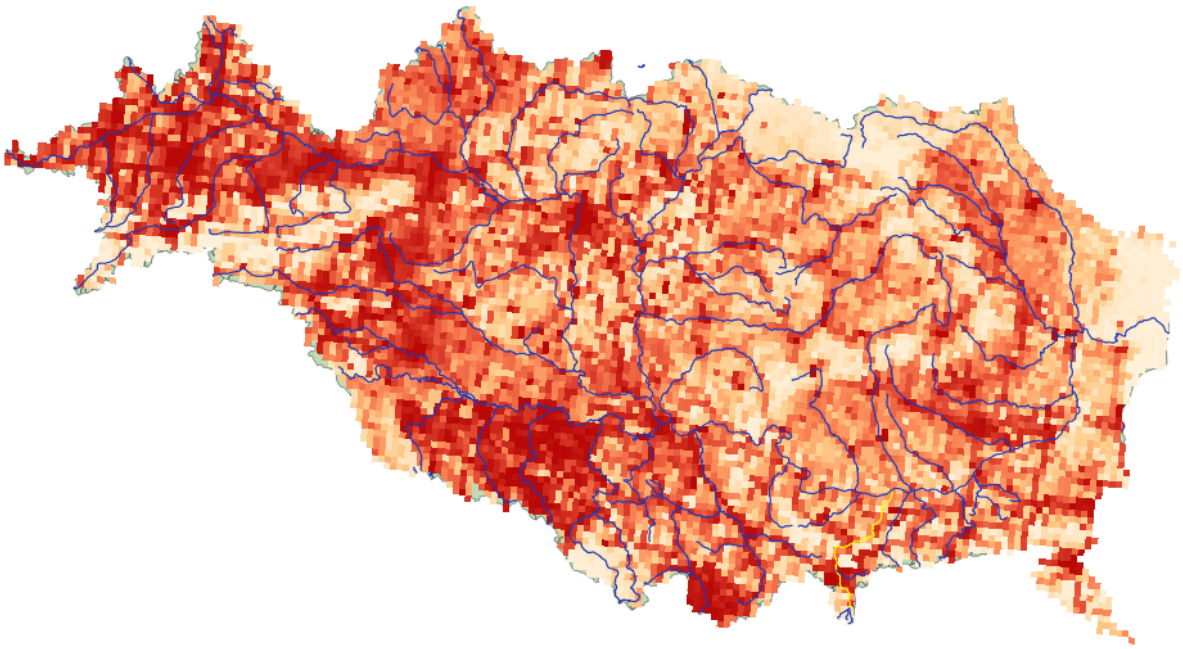

Figure 37 - EMEP - Total Pb emission (2022) grid infor the DRB from all sectors.
(Cell values represent g/cell/y, lightest color represents less then 1 mg/cell/y, most dark red color refers to 0.6 g/cell/y)

4.5.1.1 Deriving AD values from pilot site measurements

Bulk deposition values were measured in the Danube Hazard m^{3}c project[14,76] in three monthly composite samples at seven pilot catchments (2 in Austria, 2 in Hungary, 2 in Romania and 1 in Bulgaria). The deposition data was analysed to identify heterogeneity across the catchments. For the different substances some hot spot catchments have been identified with significantly higher concentration values in the rainwater (e.g. As is two orders of magnitude higher in Koppany samples, for Zinc see Figure 38). These outlier values were removed from further analysis for those substances, where no correlation was found between soil concentrations and atmospheric deposition values.

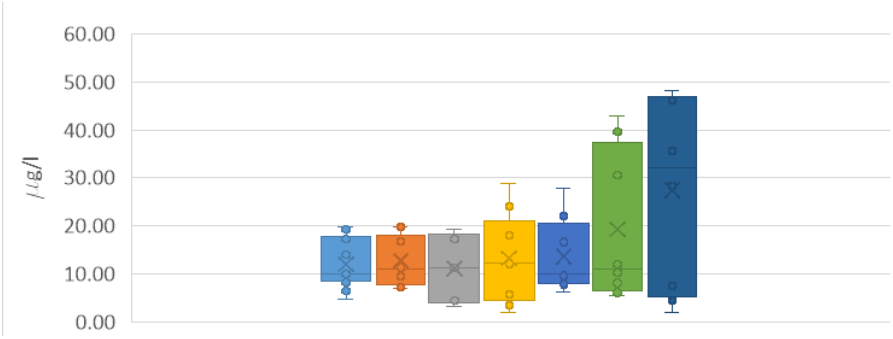

Figure 38 - Specific atmospheric deposition of Zinc

There are two approaches to distribute HM concentration in atmospheric deposition.

[76] Kittlaus, S., Kardos, M. K., Dudás, K. M., Weber, N., Clement, A., Petkova, S., Sukovic, D., Kučić Grgić, D., Kovacs, A., Kocman, D., Moldovan, C., Kirchner, M., Gabriel, O., Krampe, J., Zessner, M., & Zoboli, O. (2024). A harmonized Danube basin-wide multi-compartment concentration database to support inventories of micropollutant emissions to surface waters. *Environmental Sciences Europe*, *36*(1), 52. https://doi.org/10.1186/s12302-024-00862-4

1. Distributing the values based on empirical observations
    a. If no correlation was found between atmospheric deposition and soil HM concentrations, then a general average rate will be applied. For example, the rate of ATD_RATE and HM_MASS_SOIL is fairly constant for Cadmium (except the Vit catchment, where two strong outliers altered the averages, which may be caused by a close clinker-cement factory). It might mean that deposition is actually related to the local soils (resuspension of the soul particles), that is the soils are the primary source of deposition. Similar pattern was found for As, Ni and Cu, but with less homogeneous values across the pilots, which means that other factors might influence the atmospheric concentrations. For these chemicals, the deposition rate will be linked to soil concentrations by an average factor that is given by the current analysis.

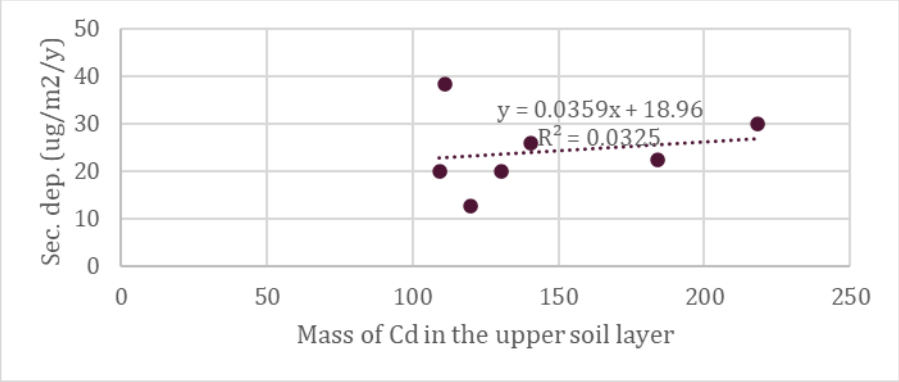

Figure 39 - Cadmium in atmospheric deposition vs. in soil

AD_RATE_HM = ER_agrl_CONT_SOIL_top_AL_HM * AD_FACT_SOIL_RATE_dep_HM

The factor values for the four chemicals is found in the table below:

Table 26 - Conversion factors to estimate Metal deposition from soil concentrations.

Variable name	Value (-)
AD_FACT_SOIL_RATE_dep_HM_AS	49.3
AD_FACT_SOIL_RATE_dep_HM_CD	7.1
AD_FACT_SOIL_RATE_dep_HM_CR	194
AD_FACT_SOIL_RATE_dep_HM_CU	6.3
AD_FACT_SOIL_RATE_dep_HM_NI	61.3
AD_FACT_SOIL_RATE_dep_HM_PB	49.9

    b. Two substances (Zn and Pb) show positive correlation with soil concentrations. In case of Zinc, this correlation is very strong (R^2 = 0.94), with no outlier (hot spot). In case of Pb, the correlation is also strong (R^2 = 0.63) when three outlier catchments (probable hot spots with extra anthropogenic influence) are removed from the analysis. If all catchments are considered, no correlation is found. As there is no evidence of anthropogenic effects at the outlier pilots, Pb is also given a general factor described in the first option.

In case of Zn, the following relationship is used to scale the deposition rate.

AD_RATE_HM_ZN = 0,783 * ER_agrl_CONT_SOIL_top_AL_HM_ZN-19985

II. The second variant for this pathway is a constant distribution of values measured in the pilot catchments of the Danube Hazard m^{3}c project (Table 27), based on the vicinity of the given AUs. Each AU receives the value from the closest pilot catchment.

Table 27 - Mean deposition rate values of 7 pilot catchment in the Danube Hazard model

AD_RATE_HM (g/ha/a)	Wulka	Ybbs	Vit	Koppany	Zagyva	Somes	Viseu
As	1.0	1.5	3.2	1.0	0.8	0.8	1.6
Cd	0.3	0.2	0.4	0.1	0.2	0.2	0.3
Cr	1.0	2.9	2.2	4.5	4.8	2.2	0.4
Cu	14.5	30.2	36.6	13.7	25.1	13.2	36.3
Ni	1.4	9.4	3.8	2.4	4.1	5.5	1.4
Pb	3.4	5.3	1.7	0.9	5.6	7.8	3.4
Zn	59.2	156.8	107.4	66.0	59.7	113.7	216.8

4.5.2 PFAS

Baseline $\log_{10}$ values reported in international literature are in the range of -3 to 2 pg/m^3 in the atmosphere for PFOA and PFOS, -0.5 to 2.7, -0.5 to 3 ng/g in settled dust for PFOA, PFOS respectively and -2 to 3 ng/l in the precipitation for both substances.

Bulk deposition values were measured in the Danube Hazard m3c project (Kardos et al., 2024, Kittlaus et al., 2024) in three monthly composite samples at seven pilot catchments (2 in Austria, 2 in Hungary, 2 in Romania and 1 in Bulgaria). Values of PFOA range between <LOQ values (20 samples) to 1.1 ng/l (7 samples), for PFOS, the values range from <0.15 (11 samples) to 7.3 ng/l (15 samples). Due to small sample numbers, significant regional differences were not spotted. Mean and median values (excluding <LOQ data) for PFOA were 0.4 and 0.33 ng/l, and for PFOS the mean and median values were 2.8 and 1.8 ng/l.

Soil concentrations, however, showed a strong correlation to runoff specific GDP (Kardos et al., 2024). In the current approach we assume that PFAS on soil surface horizon is originated from atmospheric deposition only (excluding pesticides and other agricultural sources). It is known that PFOS and PFOA, as long chain substances tend to attach to particles but also move through the soil column to groundwater. This partitioning makes it tricky to estimate the actual deposition on soils, but the pattern of the deposition seems to be obvious. This pattern is used to regionalize the atmospheric deposition.

In the Danube Hazard m^{3}c sampling campaign, mean PFOS concentration in the seven sampled catchment was 46 ng/kg and mean PFOA concentration was 81 ng/kg. Presuming a constant deposition rate over years, the average accumulated PFOA in the upper 30 cm of agricultural soils is **20** times more than the amounts accumulating by average deposition rates annually. The same ratio for PFOS is **203**. These ratios suggest differences between the two substances and reflect an order of magnitude difference in the K_d value for the matrix. This can be compared to K_d values given in the literature, which also suggest an even larger difference in case of sediments (0.6 for PFOA and 3.3 for PFOS), also suggested in batch experiments[77].

In the current approach, the above-described ratios are used to distribute deposition rates across the basin based on the mapped soil PFAS concentrations. The uncertainty of these data is high, as both the

[77] Rayner, J. L., Slee, D., Falvey, S., Kookana, R., Bekele, E., Stevenson, G., Lee, A., & Davis, G. B. (2022). Laboratory batch representation of PFAS leaching from aged field soils: Intercomparison across new and standard approaches. *Science of The Total Environment*, *838*, 156562. https://doi.org/https://doi.org/10.1016/j.scitotenv.2022.156562

measured atmospheric deposition data has high uncertainty, while soil concentrations are also based on a model with small sample numbers.

AD_SOIL_RATE_PFAS_PFOA = 20 (-)

AD_SOIL_RATE_PFAS_PFOS = 200 (-)

AD_RATE_PFAS_PFOA = ER_agrl_CONT_SOIL_top_PST_PFT_PFOA * AD_SOIL_RATE_PFAS_PFOA · 10^-6 ·BI_PREC_YEAR · 10^4 · 10^-3 (mg/ha/a)

AD_RATE_PFAS_PFOS = ER_agrl_CONT_SOIL_top_PST_PFT_PFOS * AD_SOIL_RATE_PFAS_PFOS · 10^-6 · BI_PREC_YEAR · 10^4 · 10^-3 (mg/ha/a)

4.6 GROUNDWATER

4.6.1 Heavy Metals in groundwater

Baseline: groundwater concentration in the baseline version contains a constant concentration value derived from GW datasets in earlier model applications (DHm^{3}c project database). This dataset is available and installed in the model.

Regionalized HM groundwater concentrations:

Monitoring data of heavy metal groundwater concentrations is available in many countries through the WISE reporting system. Measurement data is available either as single measurements or as aggregated concentration data. This data contains groundwater data covering different land uses and a large variation of depths, but depth data is missing for many wells. The Danube Hazard m^{3}c database was the starting dataset for our analysis. The spatial coverage of the dataset covered only six countries (AT, HU, DE, SI, RO (partial), SRB), therefore further data request has been sent to the partners, Romania provided further data.

Calculating groundwater concentrations:

- Based on the available number of wells, local hot-spots cannot be dealt with.
- Instead of analytical units, groundwater concentration is calculated for larger sub-catchments of the Danube, primarily first order catchments (Figure 40), which were further divided in the case of Tisza. The Drava has not been sub-divided as a very low number of groundwater wells were available in the whole Drava catchment.
- Where most measurements were below LOQ values, three variants were calculated.

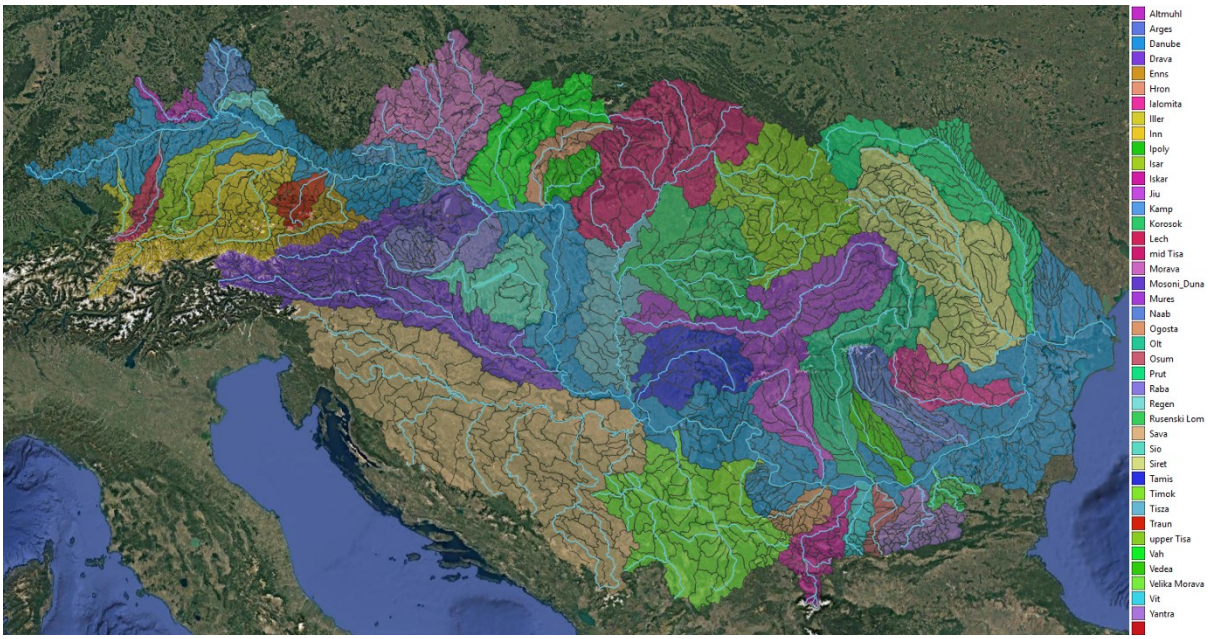

Figure 40 - First order subcatchments of the Danube (Tisza is further subdivided)

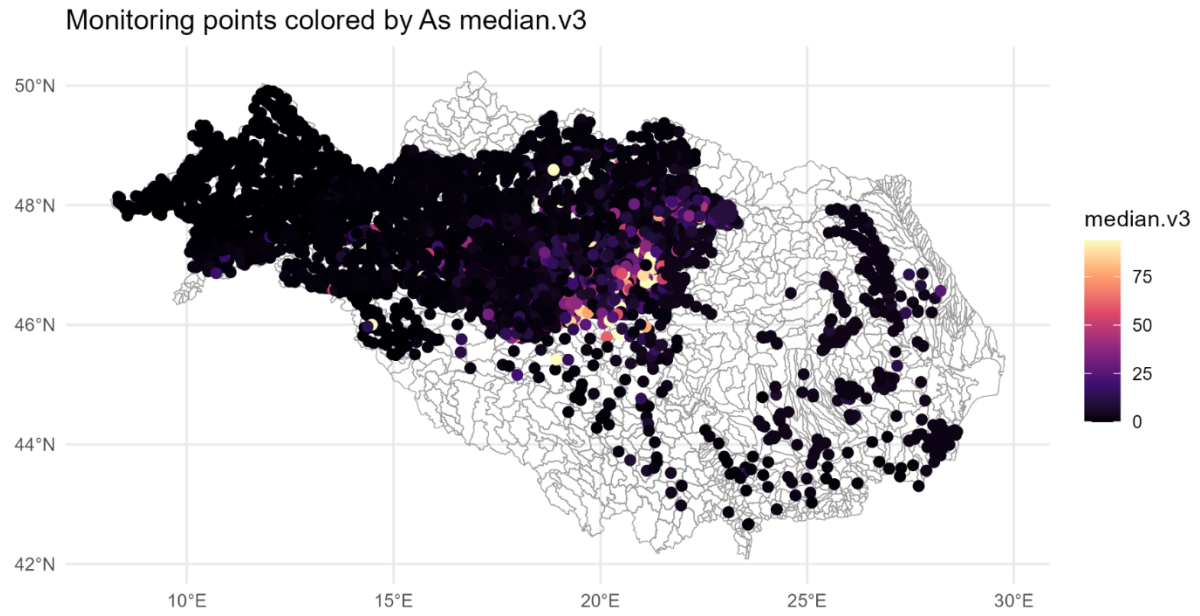

Figure 41 - Groundwater wells used to calculate concentrations in the DRB. Figure shows median Arsenic concentrations.

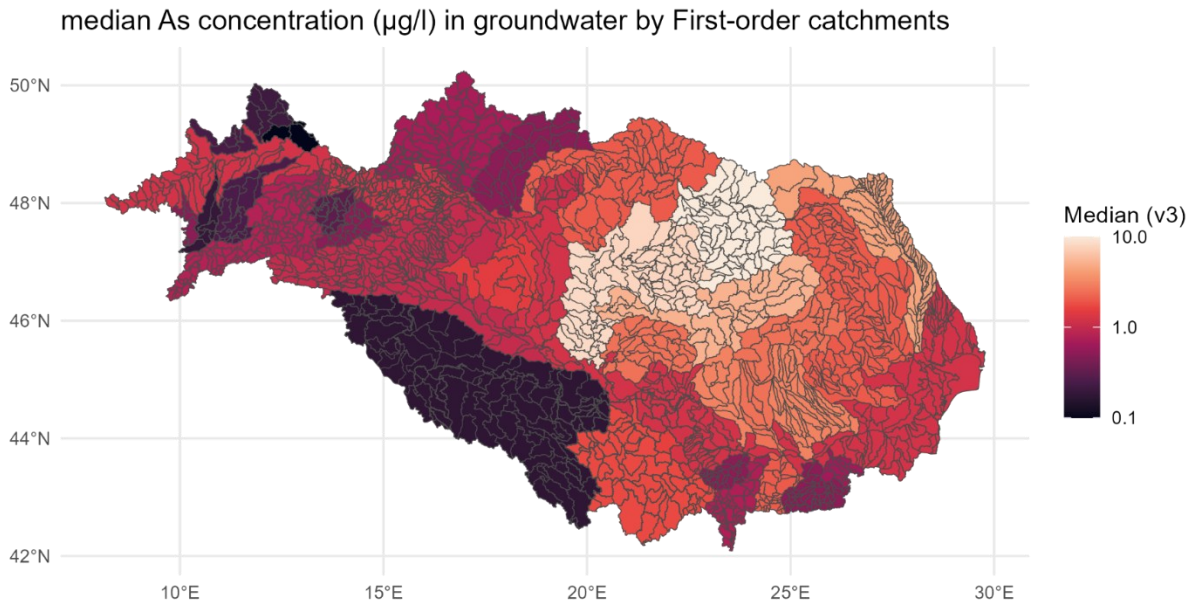

Figure 42 - median As concentrations for the Danube sub-catchments

4.6.2 PFAS in groundwater

The availability of PFAS measurements in groundwater is very limited and sporadic. Most countries have not started to routinely monitor these components in groundwater, Austria and Germany has a fairly good spatial coverage of some parts of their countries (Figure 43, Figure 44). Background values are considered to be zero as this is a fully anthropogenic substance; however diffuse and pointwise (hot-spot) emissions are both present through the groundwater pathway. Diffuse source emissions are linked to all land uses, soil contamination has been found in forests, pastures and agricultural lands in all 7 pilots of the DRB in the Danube Hazard m^3c project[14,78]. These are also supported by the groundwater measurements, that were analyses by regression on order statistics (NADA package, R[79]). Results show that detectable and quantifiable concentrations can be found in the groundwater in all land use cases (Figure 45). Concentrations under forests show an even higher value compared to agriculture, which is also in alignment with the soil measurements of the Danube Hazard m3c project, where higher PFAS concentrations were detected in forest and pasture soils[38]. It is also evident that there is an increased pollution at urban and industrial sites, therefore groundwater loads at these areas might be higher, however, recharge in urban areas is limited, therefore actual loads might not be eventually higher then loads at agricultural or natural sites.

[78] Jolankai, Z., Clement, A., Kardos, M. K., Kittlaus, S., Weber, N., Gabriel, O., Broer, M. B., Braun, K., Milačič Ščančar, R., Kozlica, K., Ščančar, J., Bordós, G., Zessner, M., & Zoboli, O. (2026). Occurrence and fate of PTE, PAH, and PFAS trace contaminants in soils and river suspended particulate matter in three DANUBEAN river catchments. *Journal of Environmental Quality*, *55*(1), e70116. https://doi.org/https://doi.org/10.1002/jeq2.70116

[79] R Core Team. (2021). R: A language and environment for statistical computing.

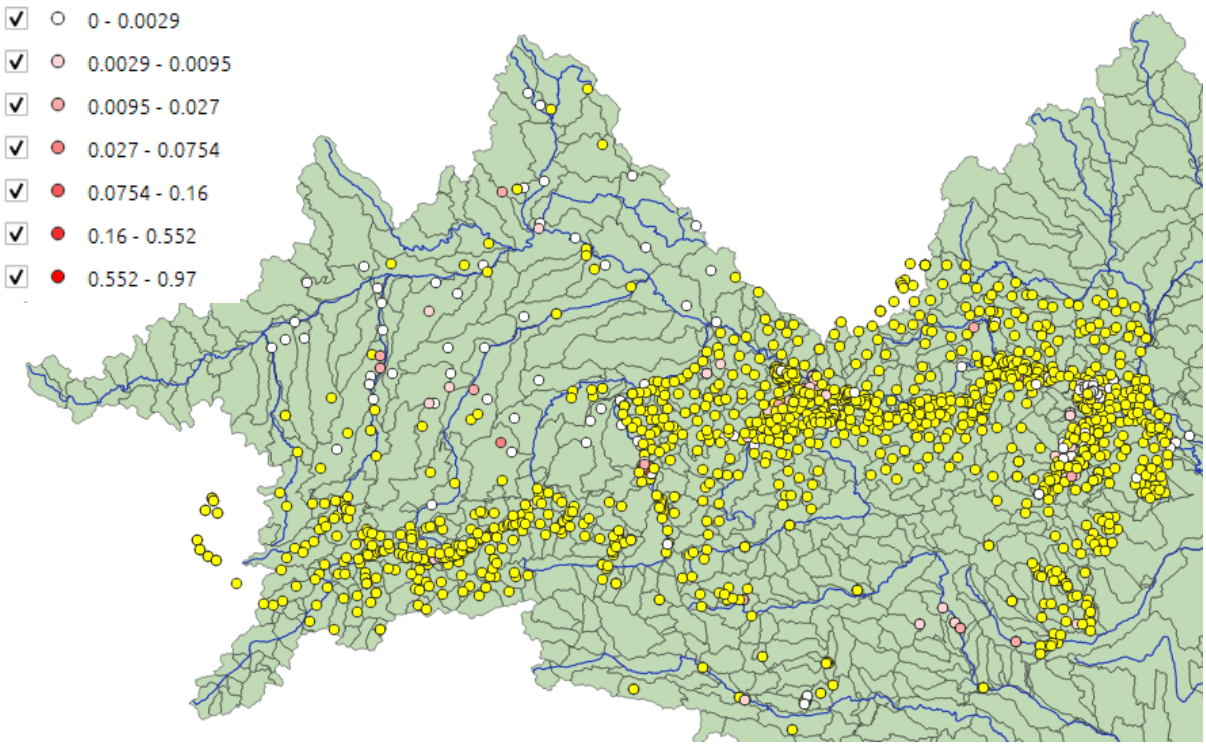

Figure 43 - PFOS measurements in the upper Danube region (DH m³c database. Yellow circles show measurements with below LOQ values.

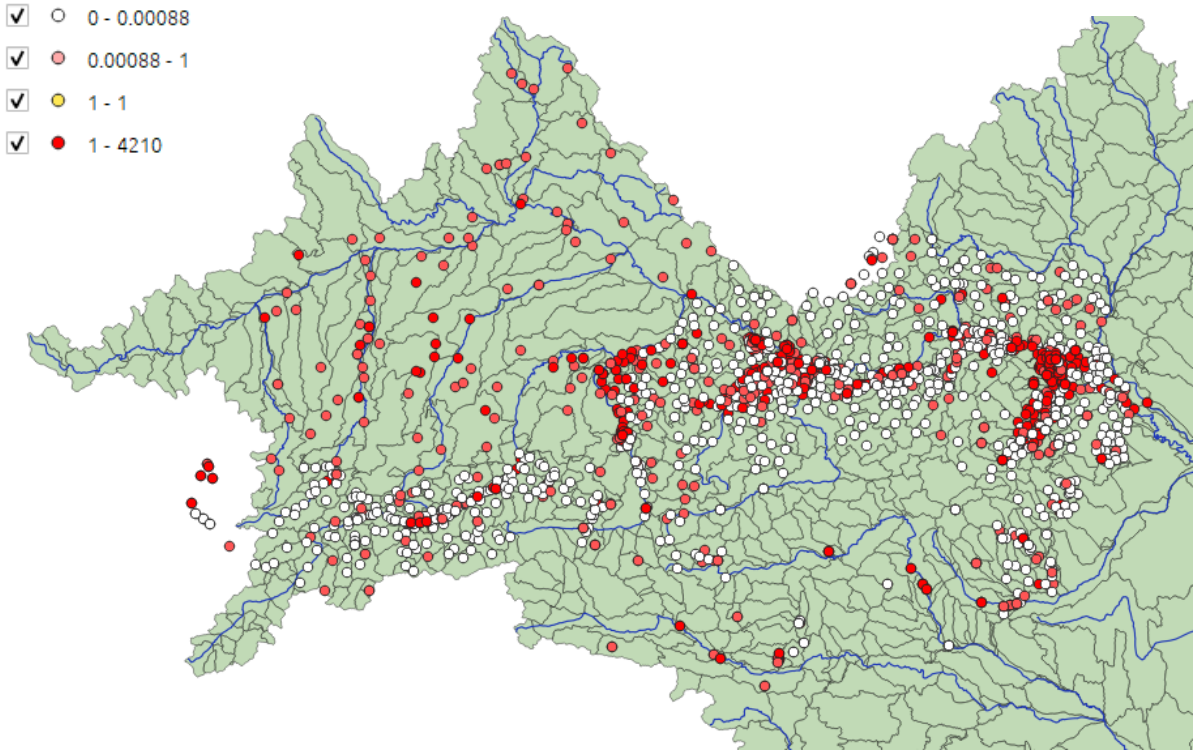

Figure 44 - PFOA measurements in the upper Danube region (DH m³c database). LOQ values were 0.25 ng/l at most samples, these are represented with white circles.

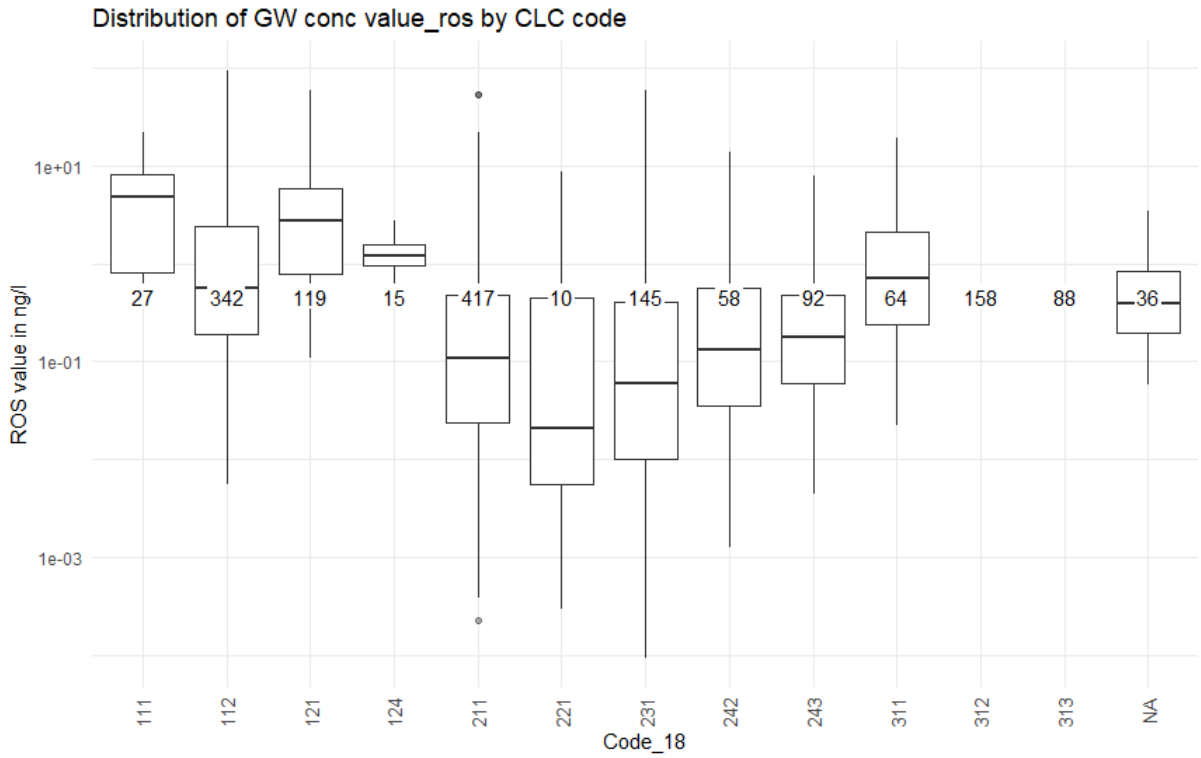

Figure 45 - PFOA concentration in groundwater per land use based on the Danube Hazard database. For code description, see table 20

Table 28 - Groundwater PFOA concentrations based on DH database. Figure represent values of the Upper Danube, Austria and Germany.

CLC code	Description	n	nq	share-AboveLOQ	mean (ng/l)	sd (ng/l)	relSD	me-dian (ng/l)
112	Discontinous urban fabric	342	175	0.49	5.08	9.19	1.81	2.47
121	Industrial and com-mercial units	119	34	0.71	7.11	10.40	1.46	4.21
124	Airports	15	3	0.80	1.43	0.61	0.42	1.20
211	Non-irrigated arable land	413	318	0.23	4.40	8.39	1.91	1.83
221	Vineyards	10	7	0.30	3.37	4.49	1.33	0.78
231	Pastures	143	113	0.21	8.16	13.70	1.68	3.54
242	Complex cultivation patterns	56	41	0.27	3.50	4.42	1.26	1.92
243	Mixed: agriculture and natural	90	70	0.22	1.85	1.21	0.65	1.81
311	Broad leaved forests	62	33	0.47	5.08	5.72	1.13	2.42
312	Coniferous forests	158	141	0.11	2.00	1.47	0.74	2.40
313	Mixed forests	88	79	0.10	6.19	8.71	1.41	2.31

Concentrations (median, lower and upper quartiles) in the groundwater have been set up in the Promisces project[18] for the upper Danube region. In the current project, those data are used as a baseline; however, further investigation is carried out to distribute concentrations within the different regions of the DRB. The starting point of the analysis is the soil PFAS concentration map prepared in this project. The leaching of the two examined PFAS compounds can be estimated based on the $K_{d,OC}$ values (distribution coefficient, see values in Table 25). Final concentrations are of course different to the leachate concentrations as leachate is mixed with a groundwater "reservoir" in a spatially and temporally varying manner. This mixing is difficult to evaluate, therefore in our analysis very rough assumptions are given, providing one single value based on the Kd value of the substance.

The partition coefficients for PFAS substances between soil and water (Kd,soil-gw) are taken from partition coefficients between organic carbon and water (Kd,OC) by considering the soil organic carbon content fOC (fraction):

$$Kd,soil\text{-}gw = KOC \cdot fOC$$

Groundwater concentration is calculated as follows:

$$Cgw = Csoil/KOC/fOC$$

Calculation in the model:

$$GW\_CONC\_PFAS = GW\_FCT\_CALIB\_PFAS \cdot ER\_agrl\_CONT\_SOIL\_top\_AL\_PFAS / KD\_OC\_PFAS / F_{OCG} \cdot CF\_unit$$

Where GW_FCT_CALIB_PFAS is a calibration factor used to set the order of magnitude right, if obvious differences are found compared to measured groundwater concentrations. KD_OC_PFAS is the KD value between soil organic matter and water in the soil water (and eventually in the groundwater). CF_unit is a conversion factor that converts between mg and µg or other units, if necessary.

To derive GW concentration, f_{OC} values are required. Soil organic matter content is available from various sources: the European soil database, the ISRIC soilgrids database[80], or the FAO HWSD database[81].

In our model calculation, the soil organic matter (SOM[82]) content from the European soil database is used for the covered countries, and the ISRIC database is used for those areas that are not covered.

4.7 TILE DRAINAGE

Tile drainage has not been applied earlier in the model; therefore concentration data had to be derived from the available datasets. Just as for groundwater, the concentrations can be derived from soil concentration data and Kd coefficients for the given substances. Also, similarly to the methods used for groundwater pathway, the concentrations have to be linked to some measured information. In this case, literature data provides these values.

[80] Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., and Rossiter, D.: SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty, SOIL, 7, 217–240, 2021. DOI

[81] Nachtergaele, F., Velthuizen, H., Verelst, L., Wiberg, D., Henry, M., Chiozza, F., Yigini, Y., Fischer, G., Tramberend, S., Batjes, N., Montanarella, L., Jones, A., Aksoy, E., Boateng, E., & Shi, X. (2023). Harmonized World Soil Database version 2.0. https://doi.org/10.4060/cc3823en

[82] Lugato, E., Lavallee, J. M., Haddix, M. L., Panagos, P., & Cotrufo, M. F. (2021). Different climate sensitivity of particulate and mineral-associated soil organic matter. *Nature Geoscience*, *14*(5), 295–300. https://doi.org/10.1038/s41561-021-00744-x

4.7.1 Heavy Metals

Limited studies are available with direct measurements of metal concentrations in the effluents from tile drain systems. One of the most relevant studies was carried out in the Netherlands[83], where a low-land plot was measured by an experimental setup of small reservoirs installed to collect the drain flows. Concentrations from the study are shown in Table 29. In this example, the ratio of the concentrations of the two compartments is close to 2-3 for Cu, Ni, Zn and Pb and over one order of magnitude for Cadmium.

Table 29 - HM concentrations in tile drain effluents and groundwater from a low-land agricultural field (Roezemeijer et al 2010)

Solute	*n*	Tile drains	Groundwater	Wilcoxon matched pairs test p-level
Cd (µg L^{-1})	69	0.14 (0.12–0.15)	0.045 (0.039–0.050)	9.82×10^{-13}
Cu (µg L^{-1})	69	10.2 (9.13–11.4)	3.13 (2.63–3.63)	2.31×10^{-11}
Ni (µg L^{-1})	69	30.0 (26.7–33.3)	15.7 (13.9–17.6)	7.09×10^{-8}
Pb (µg L^{-1})	69	0.44 (0.38–0.51)	0.20 (0.16–0.23)	1.29×10^{-8}
Zn (µg L^{-1})	69	17.5 (15.6–19.3)	8.05 (7.14–8.95)	1.41×10^{-11}

The approach is that the above given concentrations are used as a guiding range for the tile drain concentrations, that are derived from the HM concentrations in soils.

4.7.2 PFAS

PFAS concentrations in tile drain runoff is poorly known. A small-scale study shows that PFOS and PFOA both appear in the tile drain flow in varying concentrations[84] from an agricultural site that were treated with biosolids containing PFAS (this produces higher concentrations compared to “normal” contamination levels in soils). PFOS in tile drain flow shows the concentration magnitude of 1 ng/l, cca. 30% higher than groundwater concentrations, PFOA showed the magnitude of 5-20 ng/l, cca. 10 times higher concentration compared to groundwater concentrations. The appearance of PFOA in the tile drain runoff is much faster than the appearance of PFOS. This result is supported by the difference between K_d values of the two compounds. Another research is also emphasizing the potential significance of tile drainage in PFAS transport[85].

In our modelling approach, concentration in tile drainage flow is calculated with the same assumption that we use for groundwater, the only difference is the use of a different calibration factor, which describes the retention of PFAS in the upper soil column. This factor has not been calibrated in lack of TS runoff data, so this was set to 0.5 for PFOA and 0.25 for PFOS as a rough estimation. This pathway would need monitoring data to calibrate the coefficients.

---

[83] Rozemeijer, J. C., van der Velde, Y., van Geer, F. C., Bierkens, M. F. P., & Broers, H. P. (2010). Direct measurements of the tile drain and groundwater flow route contributions to surface water contamination: From field-scale concentration patterns in groundwater to catchment-scale surface water quality. *Environmental Pollution*, *158*(12), 3571–3579. https://doi.org/https://doi.org/10.1016/j.envpol.2010.08.014

[84] Gottschall, N., Topp, E., Edwards, M., Payne, M., Kleywegt, S., & Lapen, D. R. (2017). Brominated flame retardants and perfluoroalkyl acids in groundwater, tile drainage, soil, and crop grain following a high application of municipal biosolids to a field. *Science of The Total Environment*, *574*, 1345–1359. https://doi.org/https://doi.org/10.1016/j.scitotenv.2016.08.044

[85] Peter, L. G., & Lee, L. S. (2025). Sources and Pathways of PFAS Occurrence in Water Sources: Relative Contribution of Land-Applied Biosolids in an Agricultural Dominated Watershed. *Environmental Science & Technology*, *59*(2), 1344–1353. https://doi.org/10.1021/acs.est.4c09490

TD_CONC_PFAS = TD_F_CALIB_PFAS · ER_agrl_CONT_SOIL_top_AL_PFAS/ KD_OC_PFT_PFOS/fOC · TD_Q_spec · CF

Where GW_Q_spec is the specific runoff through groundwater that is equivalent to the annual recharge. GW_KD_PFAS is the KD value between soil and water in the groundwater. CF is a conversion factor that converts mm/a recharge to m3/a (1 m2 surface), µg/kg

5 FURTHER UPGRADE AND EXTENSION OF THE MODEL

5.1 R MIRRORING

To provide flexibility to the model calculations and implement uncertainty analysis of certain model parameters, the calculation engine of the MoRE model has been mirrored in R language.

Due to the dynamic structure of the MoRE model—where calculation equations are stored in the same database as the data—the underlying algorithm may change over time. A static implementation in R would therefore require continuous and labor-intensive synchronization with updates to the original model. To avoid this, the R implementation is designed to be fully dynamic and is based on an exported formula heap from MoRE.

The formula heap contains the equations defining all relevant calculated variables in an undefined order, while ensuring completeness. At a minimum, it stores the name of each calculated variable and the corresponding equation. Prior to loading the heap into R, minor preprocessing steps are required, including translating conditional expressions into valid R syntax and renaming specific function calls to ensure compatibility.

Following preprocessing, the formula heap is analyzed automatically. Variable identifiers are extracted from the equations, and dependency relationships between variables are derived by linking each left-hand-side variable to all variables appearing on the right-hand side. Based on these dependencies, a directed calculation graph (Figure 46) is constructed. This graph is used to identify input variables, terminal variables, and to determine a valid calculation sequence through dependency-aware ordering.

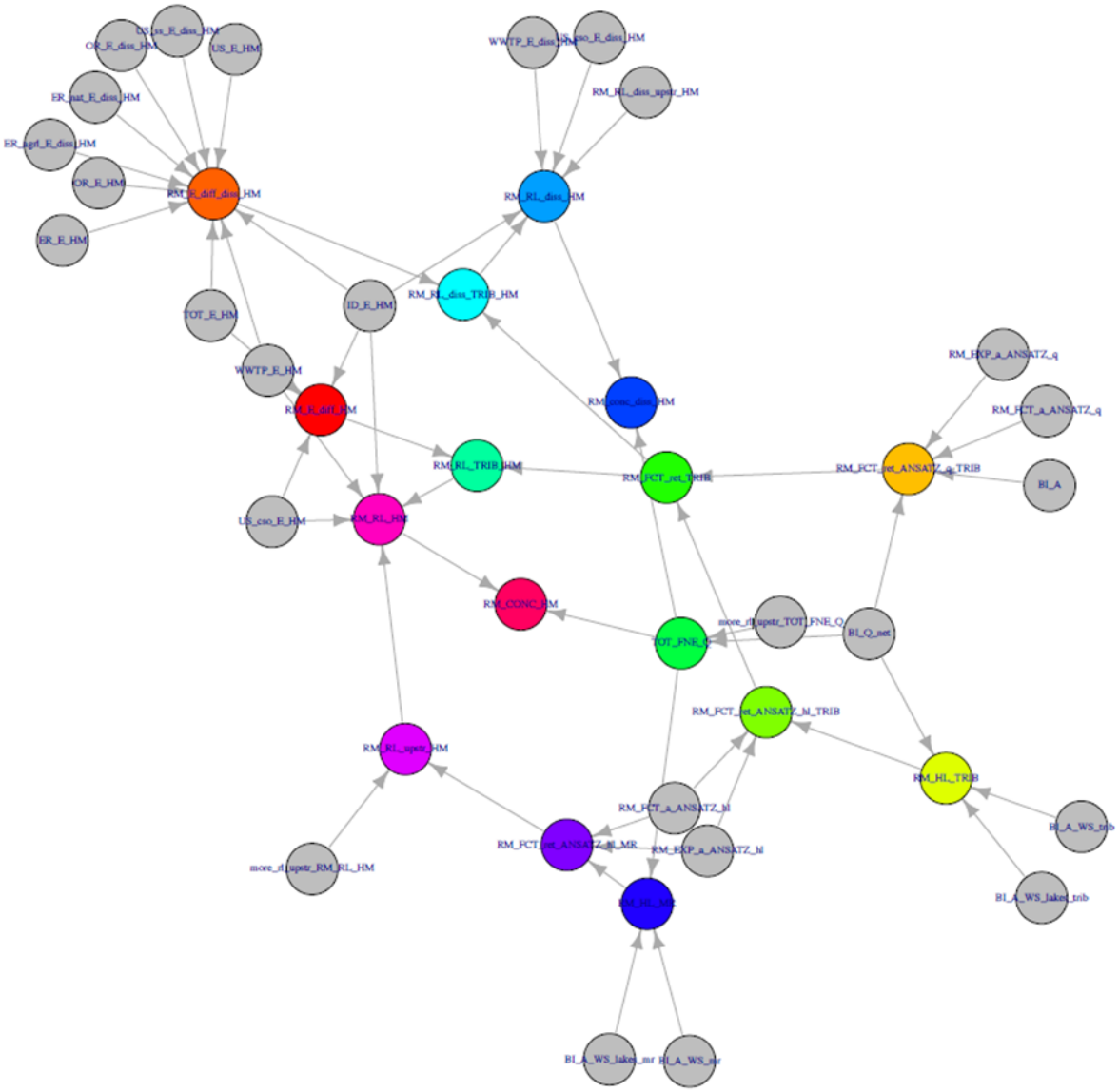

Figure 46 - Dependency graph of the MoRE-R engine

The analyzed formula heap and the resulting graph structure are stored in a compact data object that contains all algorithmic information required to evaluate the model for a single spatial unit (AU). At this stage, neither the input data nor the interconnections between AUs are specified. To fully mirror the MoRE model, the algorithm must therefore be executed for all AUs in a logical dependency order, with the required inputs provided externally. This task is handled by a separate routine capable of iterating over AU collections and selecting AU-specific inputs from a pooled data source. For further details, see Appendix I.

Testing of the MoRE-R model has been done using a result dataset from the MoRE model as reference. The relevant input data has been loaded into the R model, and the retention calculation heap has been run.

5.2 IMPLEMENTATION OF THE RETENTION CALCULATION

The MoRE model has a built-in retention algorithm that was developed for nutrients and is based on hydraulic loads in case of nitrogen and is based on area-specific runoff and hydraulic loads in case of phosphorus.

$$R_{TN} = 1/(1 + a * HL^b)$$

$$R_{TP} = 1/(1 + a * q^b) + 1/(1 + c * HL^d)$$

Where R_{XX} is the retention of the relevant substance, q is the area-specific runoff (V T -1 M -2), HL is the hydraulic load (L T -1), k_{ri} – are retention constants.

The implementation of the retention algorithm within the MoRE model is through a set of equations (Figure 47), including the aggregation of flow and the calculation of hydraulic loads. This approach uses different factors (FCT) and exponents (EXP) for tributaries (TRIB) and main rivers (MR).

step	active	formula	formula content
01	☑	RM_FCT_ret_ANSATZ_q_TRIB (...	1 / (1 + RM_FCT_a_ANSATZ_q * (BI_Q_net / BI_A * 1000)^ RM_EXP_a_ANSATZ_q)
02	☑	RM_HL_TRIB (variant 1)	if (BI_A_WS_trib + BI_A_WS_lakes_trib <= 0, 100 * 1000 * 1000, (BI_Q_net * 365 * 86400) / ((BI_A_WS_trib + BI_A_WS_lakes_trib) * 1000 * 1000))
03	☑	RM_FCT_ret_ANSATZ_hl_TRIB (...	1 / (1 + RM_FCT_a_ANSATZ_hl * RM_HL_TRIB ^ RM_EXP_a_ANSATZ_hl)
04	☑	RM_FCT_ret_TRIB (variant 1)	(RM_FCT_ret_ANSATZ_hl_TRIB + RM_FCT_ret_ANSATZ_q_TRIB) / 2
05	☑	TOT_FNE_Q (variant 1)	if(more_rl_upstr(TOT_FNE_Q) < 0 , 0 , more_rl_upstr(TOT_FNE_Q)) + if(IM_Q < 0, if(BI_Q_net < 0 , 0 , BI_Q_net), IM_Q)
06	☑	RM_HL_MR (variant 2)	if((BI_A_WS_mr + BI_A_WS_lakes_mr)<=0,100*1000*1000,TOT_FNE_Q * 365 * 86400 / ((BI_A_WS_mr + BI_A_WS_lakes_mr) * 1000 * 1000))
07	☑	RM_FCT_ret_ANSATZ_hl_MR (v...	1 / (1 + RM_FCT_a_ANSATZ_hl * RM_HL_MR ^ RM_EXP_a_ANSATZ_hl)

Figure 47 - Retention of particulate phase pollutants in the river system

These equations may be suitable for other pollutants, but the model parameters have to be set for the relevant pollutant groups. Instead of solely doing this exercise, a more stringent analysis has been done, where polluter characteristics are taken into account (see Appendix I for details). The application and validation of the relevant methods are delivered on river reach level in the first phase then transferred to AU level by a lumped approach. The workflow of the river reach approach is presented in Figure 48.

In this approach, so called bulk degradation parameters are used, that are derived from process based parameters. Three process is taken into account with respect to retention of the modelled substances: photo degradation, biotransformation and hydrolysis. These are all estimated by first order kinetics, where all sub-processes have a separate rate-constant depending on characteristics of the respected river segment, such as the surface-specific unit mass of the active sediment layer, suspended sediment concentrations, river water depth or the partitioning coefficient between liquid and solid phase. The equations are described in the Appendix I.

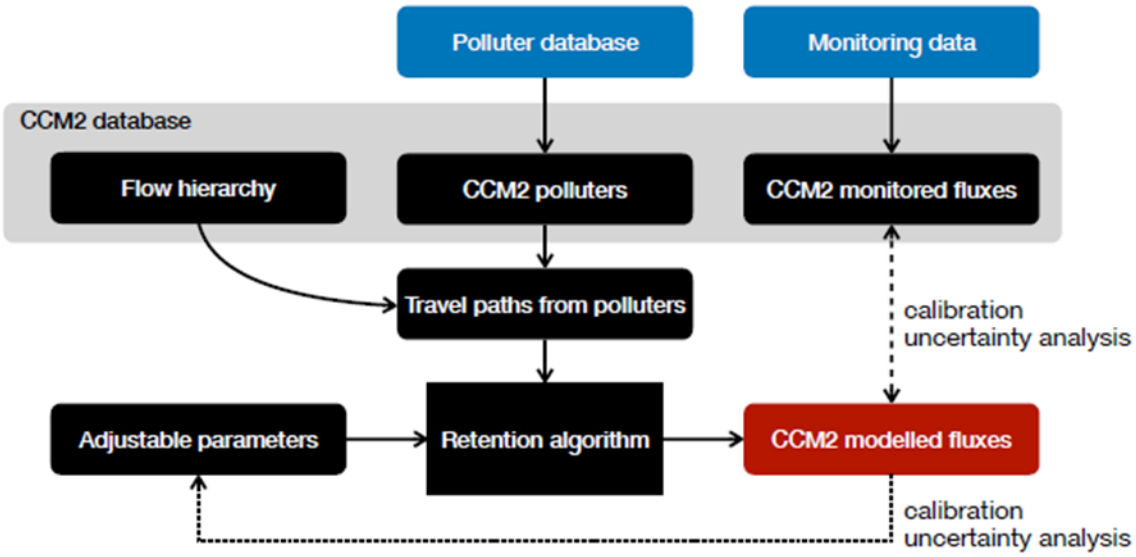

Figure 48 - Workflow of the river retention approach in river reach level. (outside of the MoRE model)

In the first phase, the retention parameters of the relevant equations are fitted for river segments for a calibration database (In our case, the Inn catchment). This is carried out in a database that is not fully aligned with the analytical units of the MoRE model, therefore aggregation can only be done for the so-called hydrologically coherent analytical units (HCAU, Figure 49).

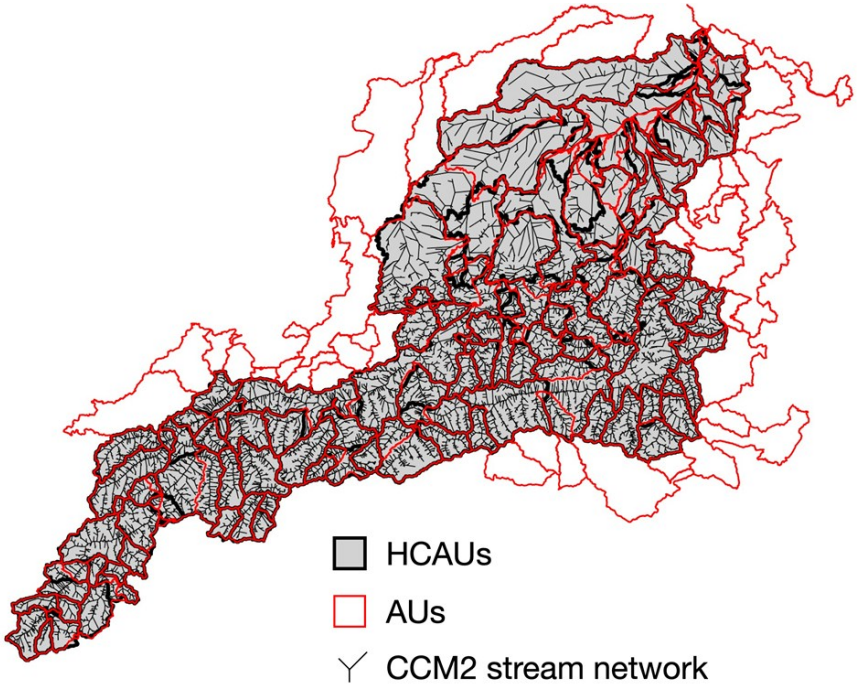

Figure 49 - Hydrologycal network and AUs vs. HCAUs in th eInn catchment.

In the next phase these parameters are upscaled to AUs, which can be achieved in different approaches:

1. *GIS-based aggregation*: Segment retention efficiencies can be attached to the subcatchments of the stream segments, thereby covering the entire basin area. This can be simply intersected with AU

boundaries, or a load-weighed procedure can be applied. This way the hydrological coherence of calculations is violated, yet the predominantly good match between AUs and HCAUs warrants a limited error.

2. *HCAU overlap*: Calculate retention on the HCAU level and then transfer HCAU retention to the corresponding AU, patch missing AUs with value from the most overlapping HCAU.

3. *Empirical model emulation*: estimate HCAU retention with a simple linear function of area, geometric mean of travel distance in main and minor streams, etc. and use that function to calculate retention in AUs or HCAUs directly.

Once the aggregation is done using any of the above options, the results can be used in MoRE. MoRE uses separate forwarding factors (1-retention) for main rivers and tributaries:

- RM_FCT_ret_ANSATZ_hl_MR for main rivers
- *RM_FCT_ret_TRIB* for tributaries, which is the average of *RM_FCT_ret_ANSATZ_hl_TRIB* (hydraulic load approach) and *RM_FCT_ret_ANSATZ_q_TRIB* (q approach).

Integration can take place along two paths:

1. No change in equations: the new retention factor can be readily plugged into *RM_FCT_ret_ANSATZ_hl_MR* and *RM_FCT_ret_TRIB*, but it will not apply to local loads of main rivers (CSOs, WWTP effluents).

2. Modify equations to reflect that we apply a uniformised, AU- and pollutant-specific retention factor to all incoming and local loads.

The overall workflow can be seen on Figure 50.

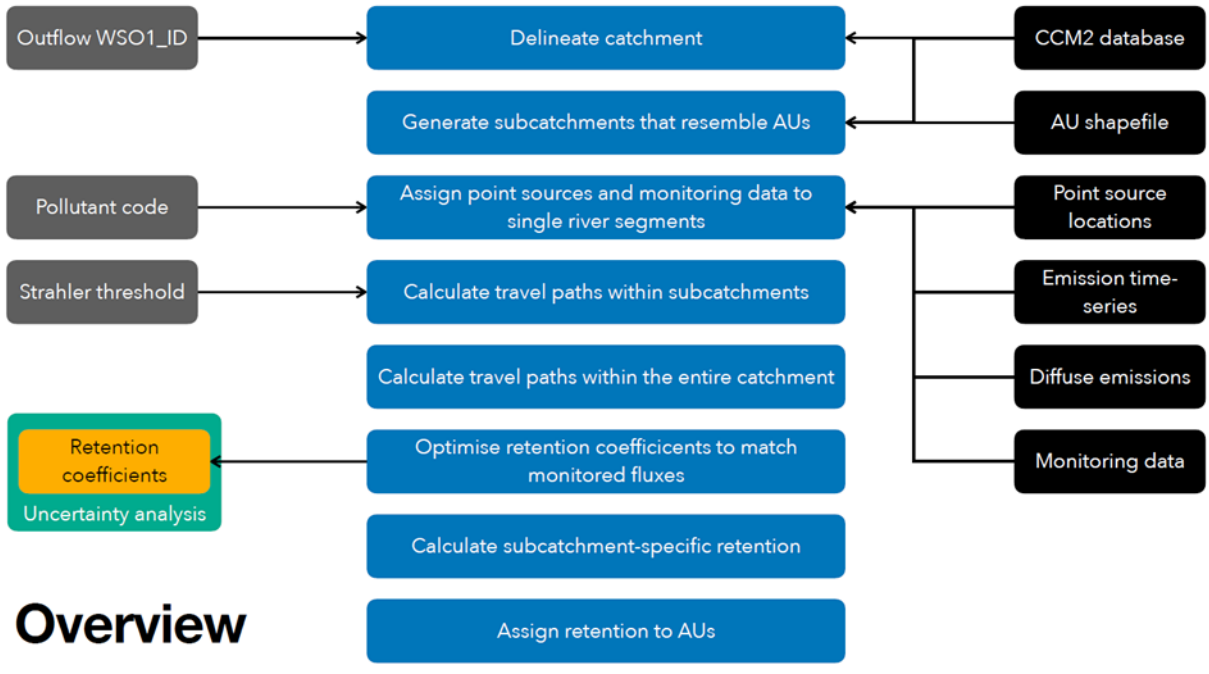

Figure 50 - Workflow of the retention calculation and adaptation to the AU level

An example results of cadmium of the heavy metal calculations are shown in Figure 51. Given the complex and dynamic process of riverine loads and retention, the approximation by a static model can be considered successful at this level of fitness.

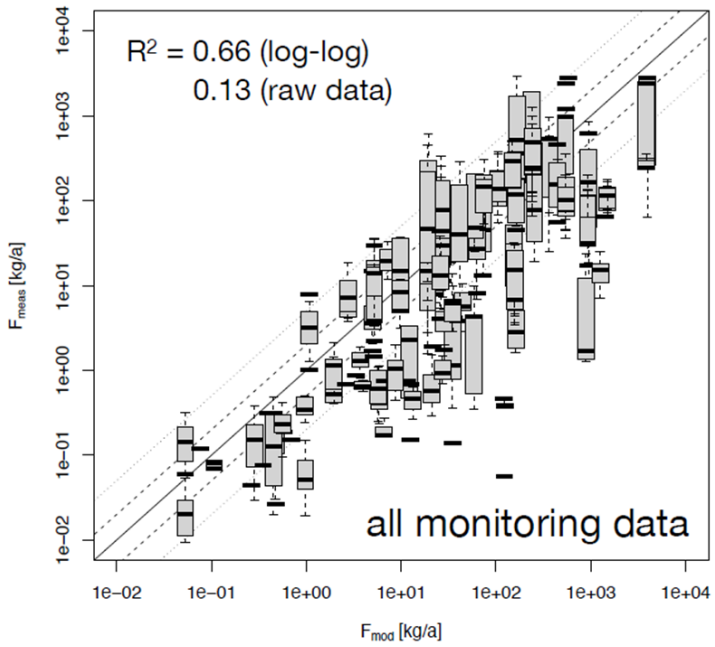

Figure 51 - Outcome of calibration of Cd retention in the Inn catchment (2018). Fmod and Fmeas are modelled and measured river loads at the specific monitoring sites, respectively. Vertical spread indicates the variability of measurement results. R2 is shown for site means of monitoring data.

Regression of the given retention parameters were also examined and good correlation was found with travel distance, which is a promising way to transfer segment level retention to AU level retention.

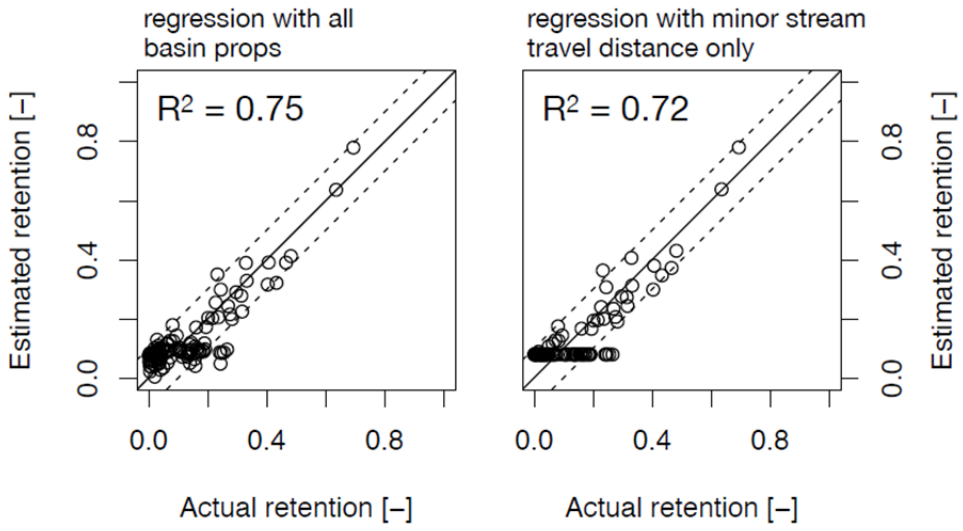

Figure 52 - Empirical estimation of HCAU-level retention efficiency of Cd in the Inn catchment. Left panel: Regression with all basin properties (area, sum of point emissions, sum of diffuse emissions, geometric mean travel distances of point and diffuse source pollution in minor and major streams), right panel: regression as a function of the geometric mean of travel distance in minor streams alone.

5.3 DEVELOPMENT OF THE STOCHASTIC FRAMEWORK

Model uncertainty arises from multiple sources throughout the modeling process. Input data are affected by measurement errors and limited spatial representativeness (input uncertainty), model algorithms are simplified representations of complex natural systems (structural uncertainty), and model parameters are difficult to calibrate and often violate assumptions of invariance (parametric uncertainty). In addition, calibration and validation data are themselves uncertain and only partially representative of true system behavior (output uncertainty). These uncertainty types influence model behavior in different ways.

Uncertainty may stem from inherent randomness (aleatory uncertainty) or from incomplete knowledge (epistemic uncertainty). Most uncertainties in environmental modelling are epistemic in nature, although both types can be addressed using probability-based methods.

In the MoRE model, input data uncertainty is handled by variants. Variants are typically derived from the input data used for the derivation of the mean or median values that are set for the model as base variants. Best-case and worst-case variants are developed using the 75th and 25th percentile of the input data (concentration database used for the given matrix, for example groundwater concentration).

In the current modelling study, we focus on the assessment of parametric uncertainty, assuming retention parameters to be the sole source of uncertainty in the model. Variations in parameter values propagate through the model structure and lead to corresponding variations in model outputs. By comparing simulated outputs with observations, the likelihood of parameter values can be estimated, enabling the derivation of uncertainty intervals for both parameters and outputs.

A Bayesian parameter inference framework is applied. Due to the absence of strong prior information, prior distributions are used only to constrain parameters to physically plausible ranges. The retention model is encapsulated in an R-based interface and integrated into the generic uncertainty assessment framework developed within the iWaQa project. A model layout file defines the assessment context, including the parameters to be varied and the comparison of model results with observations.

Parametric uncertainty is quantified using Markov chain Monte Carlo sampling of the posterior parameter distribution, yielding a representative sample from the joint posterior distribution. This sample is subsequently used to derive uncertainty intervals for the simulated river loads.

5.3.1 Results of the Inn case study

The uncertainty analysis of the Cd model for the Inn catchment had mixed outcomes for the two involved model parameters. The degradation rate in small streams (k.degr) was clearly identifiable, there is a well-defined yet wider range of values that can be associated with an acceptable fit to the measured loads (Figure 53). While the most likely value was 0.028 [km^{-1}], the uncertainty range spanned from 0.01 to 0.042 [km^{-1}]. On the other hand, the mainstream multiplier (f.main) was less identifiable. There the only certain finding was that an upper limit (0.14) exists for its value, above which the model cannot fit the measurements well. The cross-correlation plot highlighted the complex interaction between the parameters. There exists a well-defined envelope curve in the k.degr - f.main space (Figure 53). Above the diagonal line there were no points in the MCMC sample, which indicates that those parameter combinations cannot produce acceptable fit. The uncertainty range for k.degr means that – according to the emission and monitoring datasets used in this exercise – Cd was retained very efficiently in the system. The Cd flux half-life distance in the Inn would be between 17-69 km, with 23 km being the most likely (This data is not the final data in the model, emissions have been reestimated, therefore a more moderate retention is likely).

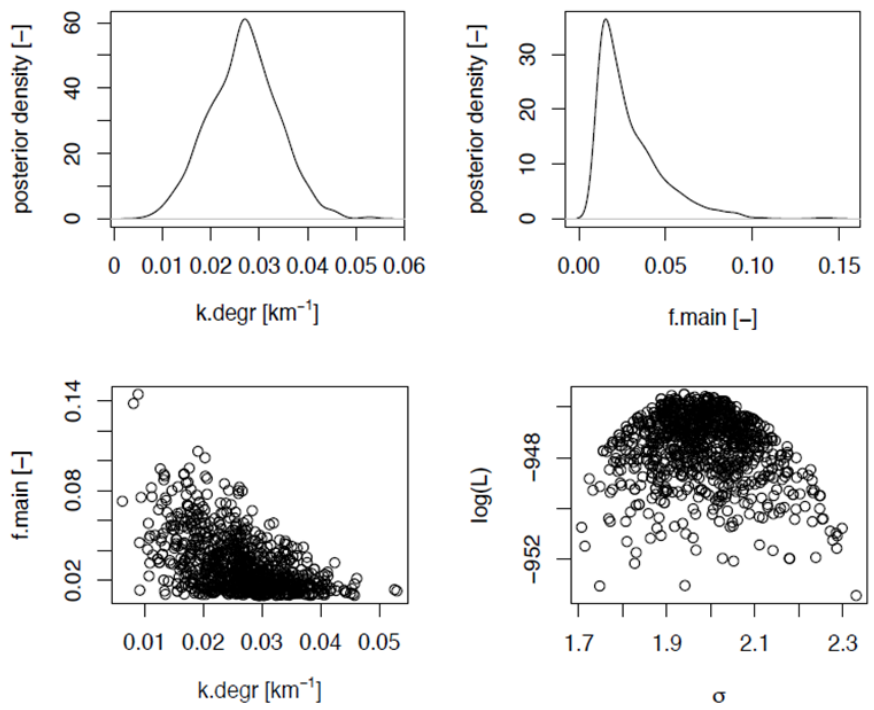

Figure 53 - Posterior probability density for model parameters from the MCMC sample. Top row: marginal distributions, bottom left: cross-correlation, bottom right: model error as the function of log likelihood (L)

5.4 MATHEMATICAL AND TECHNICAL IMPLEMENTATION OF THE SCENARIOS WITHIN THE MODEL

Scenarios can be implemented in multiple ways in the MoRE model.

One of the most convenient ways is through variants, which are a set of input parameters that represent one type of input data set, for example climate scenario, where multiple runoff data sets can be implemented with different variant names: eg. Current condition, future condition.

Scenarios can be also implemented by specific built-in variables, such as variables for measures. In the current model version, there are measure variables present for the CSO and SS pathways of urban systems:

US_cso_E_HM = if(BI_CODE_coun=>100, 0,US_cso_Q * US_cso_CONC_FUN * (1- (**MM_US_cso_EFF_FS** /100 * US_KD_FUN / 1000 * US_cso_CONC_FS / 1000) / (1 + US_KD_FUN / 1000 * US_cso_CONC_FS / 1000)) /1000 /1000)

US_ss_E_HM = if(BI_CODE_coun=>100,0,US_ss_Q * US_ss_CONC_HM * (1- (**MM_US_ss_EFF_FS** /100 * US_KD_HM / 1000 * US_ss_CONC_FS / 1000) / (1 + US_KD_HM / 1000 * US_ss_CONC_FS / 1000))/1000 /1000)

Pollution reduction scenario for atmospheric deposition is also available (MM_AD_RATE_dep_red_HM) that influences multiple pathways, atmospheric deposition (AD), surface runoff (SR) and open road runoff (OR):

SR_E_HM = SR_Q * 86400 * 365 * SR_CONC_HM * (1- **MM_AD_RATE_dep_red_HM** /100) / 1000 / 1000

OR_E_HM = (OR_SFL_HM - MM_AD_RATE_dep_red_HM /100 * BI_RATE_dep_HM) * IM_A_OR_qsr / 10

AD_E_HM = BI_RATE_dep_HM * (1- **MM_AD_RATE_dep_red_HM** / 100) * IM_A_WS / 10

Erosion control can be also implemented this way, unless we implement a different set of C factors in the model.

ER_agrl_E_HM = (ER_agrl_SL_AL * ER_agrl_CONT_SOIL_top_AL_HM * (1 - **MM_ER_EFF_AL_SED** / 100) + ER_agrl_SL_PST * ER_agrl_CONT_SOIL_top_PST_HM) * ER_SDR * 0.01 * ER_ENR_AL / 1000

With this approach, pollution reduction scenarios can be implemented for different pathways.

The last way to implement a scenario is to introduce calculation of alternatives. In the current model, such alternative is present for a lot of pathways, for example, the estimation of wastewater loads, where loads can be calculated based on single concentration values (variant 2), or by wastewater plant treatment stages (variant 3):

variant	formula content
1	if(WWTP_ps_CP >= MM_WWTP_CP_min_TS4, WWTP_CONC_HM * (1 - MM_WWTP_EFF_HM / 100), WWTP_CONC_HM)
2	WWTP_ps_Q * WWTP_CONC_HM/1000000
3	if(WWTP_treatment_type=0,WWTP_ps_Q*WWTP_CONC_NOTREAT_HM*10^-6,if(WWTP_treatment_type = 1,WWTP_ps_Q*WWTP_CONC_PRIM_HM*10^-6,if(WWTP_treatment_type = 2,'

6 APPROACH OF MODEL VALIDATION

The model validation is designed in a way that it can utilize the existing databases. In practical terms this means that the model validation is delivered for different conditions.

a) The JDS based validation focuses on the spatially homogeneous monitoring results of the JDS3 and JDS4 campaigns, which represent a given hydrological condition (summer low flow).

b) Validation for mean discharge condition, using monitoring datasets from the TNMN network, the Danube Hazard m³c project and the Tethys project (Mostly covering TNMN sites). These datasets cover different periods but spatially give enough data to validate the model at smaller hydrological units as well.

Loads are calculated using mean discharge of the given model period and the concentration data of the monitoring stations at baseflow conditions (highflow periods are filtered out).

6.1 CONCENTRATION DATA COLLECTION FOR VALIDATION DATASETS

6.1.1 Hazardous substances inventory created in the Danube Hazard m³c project

In the Danube Hazard m³c project, an inventory of hazardous substances concentrations was created. This product is well described in Kittlaus et al. (2024 ESEU). It was available in the current project without any restrictions. The spatial distribution of the monitoring data is shown in the sections below.

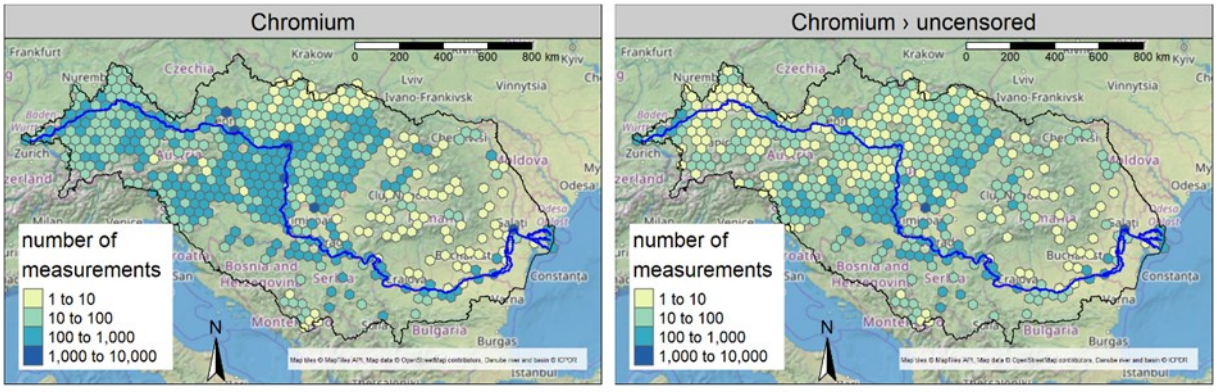

Figure 54 - Cr in surface waters in the DHm3c database. Source: Kittlaus et al. 2024

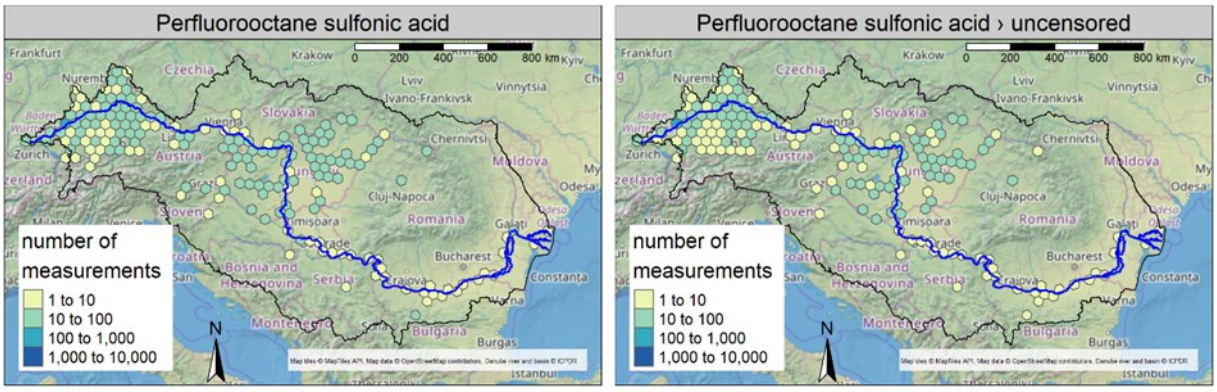

Figure 55 - PFOS in surface waters in the DHm³c database. Source: Kittlaus et al. 2024

Figure 56 - Spatial availability of measurements in surface water (water samples) for carbamazepine. Adapted from Kittlaus et al. 2024 ESEU.

6.1.2 TNMN monitoring data

The spatially and temporally most extensive dataset that covers surface water monitoring data in the DRB is collected in the transnational monitoring network of the Danube countries, managed by ICPDR. It has been operating since 1998, the first yearbook published in 2011 already reports heavy metal measurements from 2008 and 2009. From the three substance groups, this dataset contains mostly heavy metal measurements with monthly sampling frequency. We have downloaded datasets for the modelling period and also for 2021. The stations included in the network are shown in Figure 57. All major tributaries are monitored.

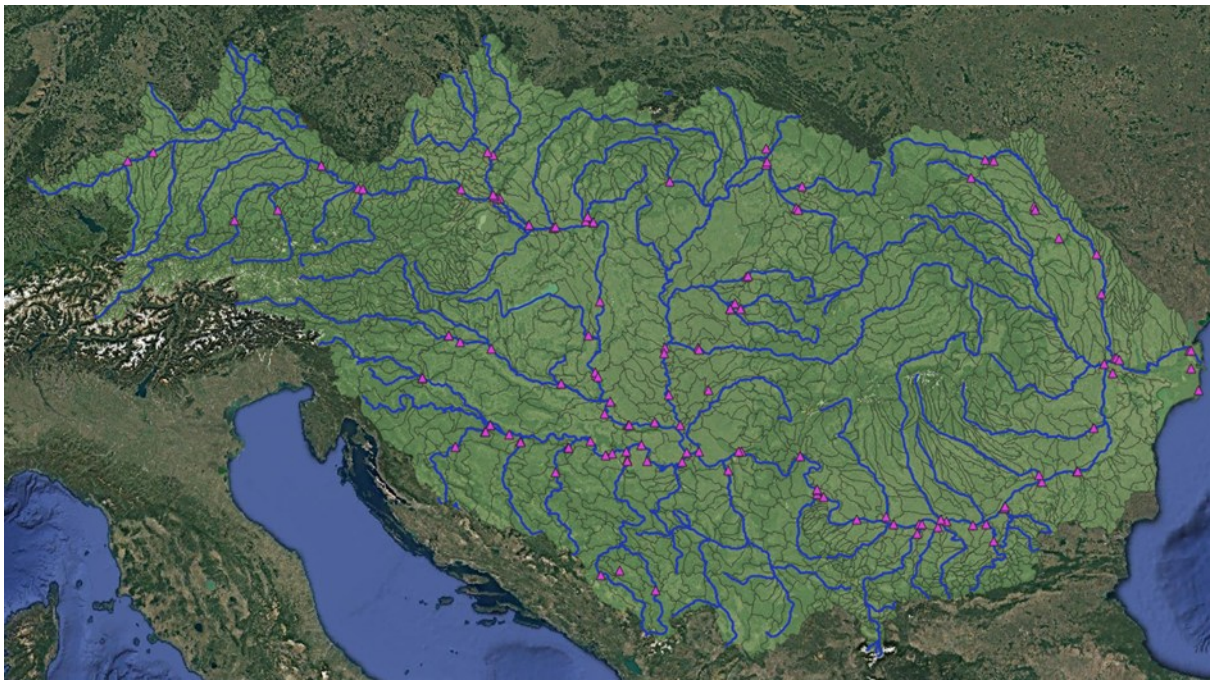

Figure 57 - Sampling points of the TNMN monitoring network

6.1.3 JDS monitoring data

The Joint Danube Survey (JDS) is an international scientific research expedition conducted along the Danube River and its major tributaries. It is one of the largest river monitoring exercises in the world and is organized by the ICPDR. The survey is carried out every six years and aims to collect comprehensive, harmonized data about the river's water quality, biodiversity, and pollution levels.

6.1.3.1 JDS3

Although conducted earlier (2013) than the modelling period, the Joint Danube Survey 3 (JDS3), placed a significant focus on micropollutants due to their growing recognition as threats to aquatic ecosystems and human health. All project substances were covered. Samples were taken along the Danube River and at the confluence of main tributaries but did not cover the tributaries themselves (except for at their most downstream point).

Figure 58 - Sampling locations in Joint Danube Survey 3. Adapted from ICPDR.

6.1.3.2 JDS4

The JDS4 (conducted in 2019) already falls within the modelling period. It covered the Danube River, main tributaries (and even wastewater treatment plants). The findings from JDS4 provide a critical baseline for understanding the distribution and concentration of micropollutants in the Danube River Basin, informing future monitoring and mitigation efforts. However, the fact that the lab analysis was conducted by three independent laboratories, leading to substantially different results, makes the interpretation of the results challenging.

6.1.4 Monitoring data from the Tethys project

The monitoring activity in the Tethys project was completed and produced a spatially extensive dataset that complemented the dataset in the DHm^{3}c inventory. In this program, EU countries collected 25 samples each, focusing on hot spot emissions; Non-EU countries collected ~100 samples each focusing more on riverine concentrations supporting load estimates. The most important outcome is the production of the validation data for PFAS and pharmaceuticals included in the model, as those datasets were very weak.

6.1.4.1 Sampling Design in Non-EU countries

Sampling campaigns in non-EU countries (Montenegro, Bosnia and Herzegovina, Serbia, and Ukraine) focused on river water, wastewater, and groundwater to validate models and identify pollution sources. For rivers, 8–12 sampling sites per country included major tributaries, hotspots, and background locations. Six samples were taken per site throughout the year under varying flow conditions, analyzing metals, PFAS, pharmaceuticals, and suspended solids. In total, countries collected 50–70 river water samples.

6.1.4.2 Sampling Design in EU Countries

The sampling campaign for EU countries, including Austria, Slovakia, Slovenia, Bulgaria, Romania, Croatia, and Hungary, was designed to assess river water, wastewater, and urban stormwater runoff. River sampling targeted three key site categories: outflow points of major tributaries for load validation, hotspots for pollution source identification, and background locations for baseline data. Sampling was consistent with the Transnational Monitoring Network (TNMN) and included major tributaries (e.g., Morava, Tisza, and Mures) and the Danube at country borders. Each country collected 4 samples per site under low-to-mid flow conditions across 2–3 sites, totaling 8–12 samples annually. Parameters measured include metals, PFAS, and pharmaceuticals, with additional considerations for specific pollutants in hotspot areas. Supporting data such as flow, pH, conductivity, and temperature was also recorded. Wastewater sampling focused on effluents from municipal, industrial, and mining-related treatment plants at 3–5 sites per country, collecting 9–15 samples annually. Urban stormwater runoff sampling, conducted where feasible, analyzed road runoff at 3–8 sites per country, yielding 9 24 samples per year. Each matrix's parameter included metals, PFAS, pharmaceuticals, and in situ measurements (e.g., pH, conductivity). In total, the sampling strategy resulted in up to 25 samples per country, targeting representative pollution sources to provide comprehensive data for pollution assessment and management in the Danube region. Each partner country selected sites to best capture specific pollution dynamics.

The applied approach ensured comprehensive coverage of potential pollution sources while providing critical data for model validation. However, the data became available in the second half of the project, at which point the modeling work was already at an advanced stage.

6.1.5 Additional data received from partners

Additional pharmaceutical monitoring data were received from Romanian project partners through the WISE reporting framework. The dataset covered the period 2016–2022 and included measurements for several pharmaceutical substances. Among these, Diclofenac data was incorporated into the validation dataset.

6.2 DISCHARGE DATA COLLECTION FOR LOAD CALCULATION

The following three discharge data collections were used:

- year-independent discharge values from the ICPDR–MONERIS framework,
- measured, year-dependent discharge data from the DWB project,
- modeled discharge values from the CWatM model for 2015–2020, with averages applied for available years as an initial approximation.

6.2.1 MONERIS discharge dataset

Long-term discharge values derived from the ICPDR–MONERIS framework were included as an additional reference, providing year-independent discharge estimates for selected validation scenarios.

6.2.2 Measured datasets from the Danube Water Balance project

Measured discharge data were primarily obtained from the Danube Water Balance (DWB) project, which provides observed river flow information for approximately 1,500 stations across the basin. These stations cover both the Danube main stem and its major tributaries, offering extensive spatial coverage for linking water quality monitoring points to hydrological units. The DWB dataset served as the main source of measured discharge values for load calculations.

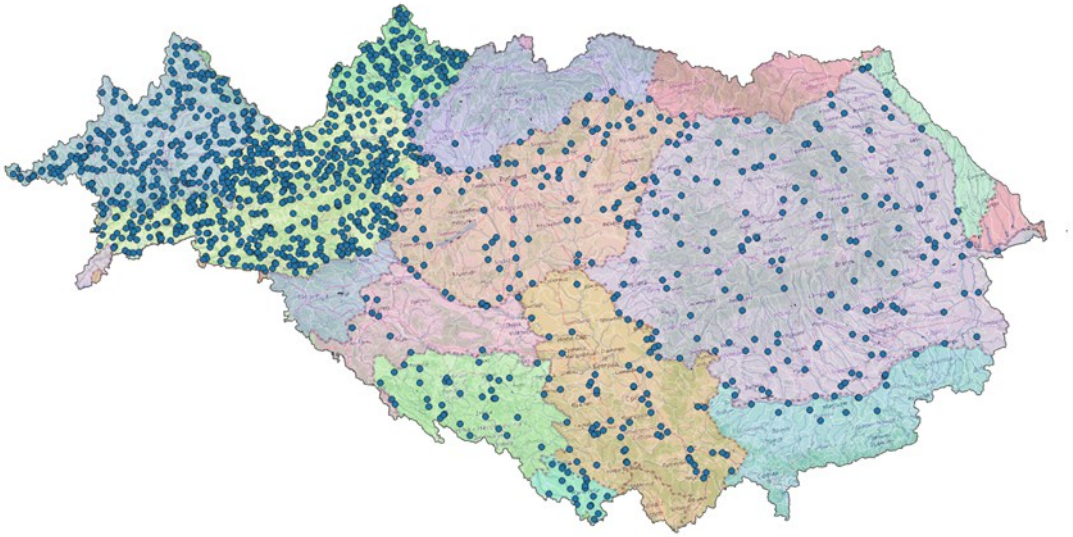

Figure 59 - Municipal sewage discharge points in the Danube River Basin.

6.2.3 Complementary modeled discharge datasets

In addition to measured discharge data, model-based discharge estimates were also considered. Discharge outputs from the CWatM model were available for the period 2015–2020 and were used to complement the measured datasets where appropriate.

6.3 METHODOLOGY FOR CONCENTRATION AND LOAD CALCULATIONS USED IN VALIDATION

Preparing of validation data was performed through a multi-step workflow structured in R programming language, combining water quality monitoring data with discharge information at the scale of the MoRE model analytical units (AUs). The procedure was designed to maximize the use of available observations while ensuring spatial representativeness and methodological consistency across substance groups.

6.3.1 Processing and pre-selection of monitoring data

Existing concentration datasets were processed. High-flow observations were excluded to focus on baseflow conditions, which are more consistent with the assumptions of the modeling framework. Total and dissolved concentration fractions were separated and prepared for independent use in the validation process, as they represent different transport mechanisms.

Concentration data originating from the JDS3 and JDS4 campaigns were handled separately during data preparation and formed a standalone validation output, while all remaining data sources were merged and formed a second validation output.

For metals, monitoring points with fewer than 10 measurements per metal were excluded from further analysis. Only stations with at least 10 observations per metal were retained in the pool from which monitoring stations were assigned to analytical units.

For organic substances (pharmaceuticals and PFAS), no minimum measurement threshold was applied, and all available monitoring points were considered eligible for validation, taking into account the overall scarcity of monitoring data for these substances.

6.3.2 Manual spatial linking of monitoring stations to analytical units

Monitoring stations were manually reviewed and assigned to the MoRE model analytical units separately for each investigated substance group (metals, pharmaceuticals, and PFAS). The assignment aimed to maximize the number of usable monitoring points while ensuring that selected stations were located as close as possible to the outlet of the corresponding catchment. In all cases, the hierarchical structure of the AUs was taken into account.

Where multiple monitoring stations were available within a single AU, only the station located closest to the outlet was selected for validation purposes. For example, in cases (Figure 60) where up to seven metal-monitoring stations (marked with purple dots) were available within an AU, only the most downstream station was retained (marked with a green circle). Year-to-year variations in sampling locations were taken into account to avoid unnecessary data loss.

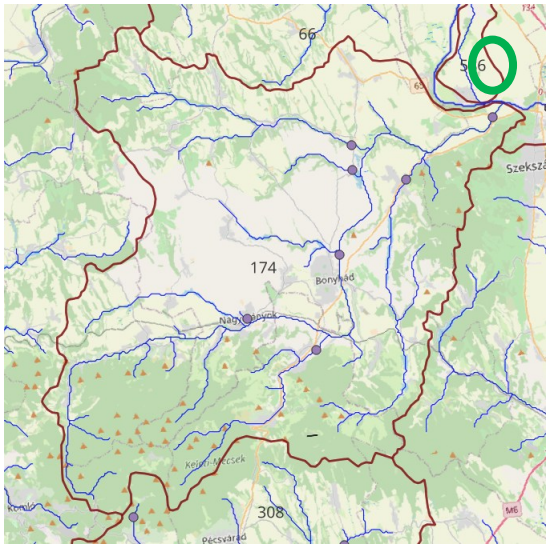

Figure 60 - Available metal monitoring points on AU 174 (marked with purple dots) and the most downstream station selected for validation (marked with a green circle).

The spatial location of monitoring stations proved to be a critical limiting factor. To many AUs, no suitable station could be assigned despite the presence of multiple monitoring points within the catchment. Station selection criteria were applied more strictly in upstream areas, where representativeness is highly sensitive to station placement, while selection rules were gradually relaxed downstream along the Danube main stem with monitoring stations located on side arms in the lower reaches being entirely excluded from the analysis.

An overview of the number of selected monitoring points for validation and the share of stations with long-term average concentrations below the limit of quantification (LOQ) can be seen in Table 30.

Table 30 - The total number of available monitoring stations and the subset selected for validation were quantified for each substance

Substance	Number of monitoring points (total form)	Share of monitoring points with avarage concentrations <LOQ (total form)	Number of monitoing points (dissolved form)	Share of monitoring points with avarage concentrations <LOQ (dissolved form)
Cr	239	25%	324	60%

Substance	Number of monitoring points (total form)	Share of monitoring points with avarage concentrations <LOQ (total form)	Number of monitoing points (dissolved form)	Share of monitoring points with avarage concentrations <LOQ (dissolved form)
Ni	240	13%	326	20%
Cu	236	5%	326	9%
Zn	240	8%	324	14%
As	237	14%	326	14%
Cd	239	68%	318	69%
Pb	240	24%	326	59%
PFOA	165	41%		
PFOS	197	21%		
CBZ	123	6%		
DCF	138	9%		

6.3.3 Manual spatial linking of discharge stations to analytical units

A similar manual spatial linking procedure was applied to discharge stations providing discharge data for load calculations. Measured discharge data from the DWB project (1523 stations) were spatially matched with the MoRE analytical units to determine whether stations represent AU outlet conditions.

589 discharge stations were identified as suitable for validation based on spatial representativeness Figure 61. Of these, data from 396 stations were fully accessible (marked with green dots) and used in the analysis, while data from an additional 193 stations (marked with white dots) could not be applied due to licensing restrictions.

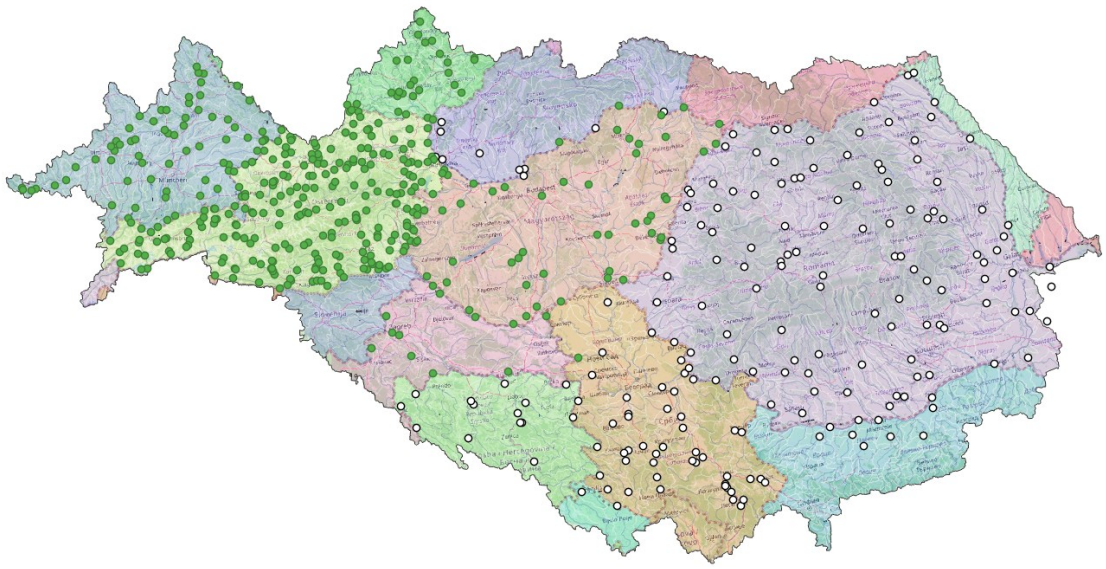

Figure 61 - Annual average discharge values were calculated for the available years and later combined with concentration statistics to derive annual load estimates.

6.3.4 Concentration data processing and handling of censored data

Concentration data for each pollutant were primarily summarized using annual mean and annual median values. In addition, long-term averages and medians were calculated over the full investigated period. The specific years included depended on data availability for each substance and station. Aggregation was performed in two steps: first at the station–year level, followed by aggregation at the AU level.

Given the high proportion of censored data, the Regression on Order Statistics (ROS) (Helsel, 2011) method was applied under defined conditions to maximize data usability while limiting systematic bias. Single and aggregated datasets were treated separately.

For single measurement values for a given monitoring point:

- If all measurements were above LOQ, the arithmetic mean was calculated.
- If the proportion of censored values was above 80% and at least three uncensored observations were available, ROS was applied assuming lognormal distribution of the measurements.
- If more than 80% of values were censored or fewer than three uncensored observations were available, the result was reported as <LOQ.
- Where a limit of detection (LOD) was reported, it was used instead of LOQ.

For aggregated values at a given monitoring point, weighted avarages were calculated using the number of samples.

In order to maximize the use of available but still relevant monitoring data at the outlet of each investigated AU, results from multiple monitoring stations were considered where applicable. If more than one monitoring station was selected for validation within a given AU, AU-level aggregation was applied.

When multiple station-level results (already processed using ROS where applicable) were available for the same year, aggregation was performed as follows:

- all values <LOQ → highest <LOQ value retained
- all values uncensored → arithmetic mean
- mixed censored and uncensored → ROS method applied again

6.3.5 Preparation of measured discharge data

Measured discharge data were available as daily time series for selected monitoring stations. Annual average discharge values were calculated only for years with sufficient data coverage. For stations with complete daily records, annual means were calculated directly. In cases where partial data gaps were present, annual averages were calculated only if the number of missing daily values remained below a predefined threshold (<10%, corresponding to fewer than 36 days). For these cases, missing daily values were filled by linear interpolation prior to annual averaging. Years with insufficient data coverage were excluded from annual averaging. No interpolation of missing discharge values was applied for years with extensive data gaps. Instead, such years were explicitly marked as missing and excluded from further calculations.

For the JDS3 and JDS4 validation datasets, campaign-specific average discharge values were calculated using daily discharge data corresponding to the respective survey periods (August–September 2013 for JDS3 and June–July 2019 for JDS4). Only stations with complete or near-complete coverage during the survey months were retained. These campaign-specific discharge values were used exclusively for JDS-based load calculations.

6.3.6 Calculation of loads

Annual loads were calculated using multiple discharge–concentration combinations. The three discharge inputs described in chapter 6.2 were applied:

Load calculations were performed both on an annual basis and using long-term average concentrations and discharges. Annual loads were derived by combining year-specific concentration values with corresponding year-specific discharge data. In addition, long-term average loads were calculated by pairing multi-year average concentrations with average discharge values, providing a complementary, time-integrated estimate of loads for validation purposes.

In most versions, load values were calculated only when concentration estimates were above the limit of quantification. If a concentration value remained below the LOQ after statistical processing (including ROS where applicable, see Ch. 6.3.4), no load value was calculated for that year. This approach avoids introducing artificial and false load estimates based on censored concentration data.

An additional scenario was implemented in which, if the long-term average concentration assigned to an AU was below LOQ, the load was calculated using the LOQ value ("worst case calculation scenario").

For JDS3 and JDS4 validation datasets, load calculations were performed using campaign-specific average discharge values. JDS3- and JDS4-derived concentrations were combined with the corresponding average discharge estimates representative of the respective survey periods.

6.4 ESTIMATED RESULTS OF CONCENTRATION VALUES AT VALIDATION POINTS

The following figures present average concentration values for selected metals, pharmaceuticals, and PFAS at validation locations, shown before and after the inclusion of Tethys monitoring points in the validation process.

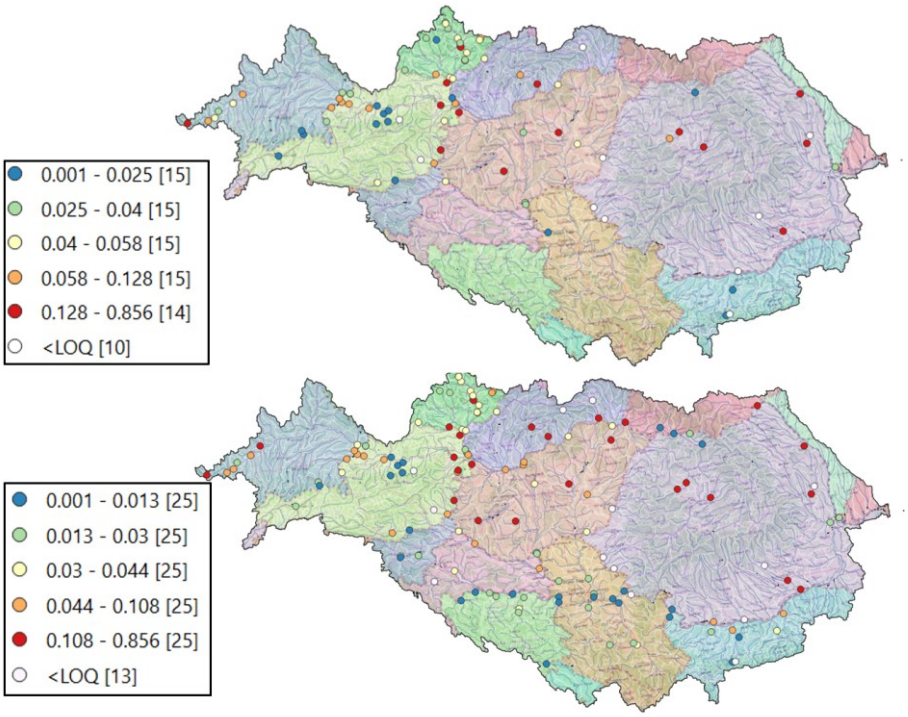

Figure 62 - Average concentration – DCF [µg/L], before and after considering Tethys monitoring points

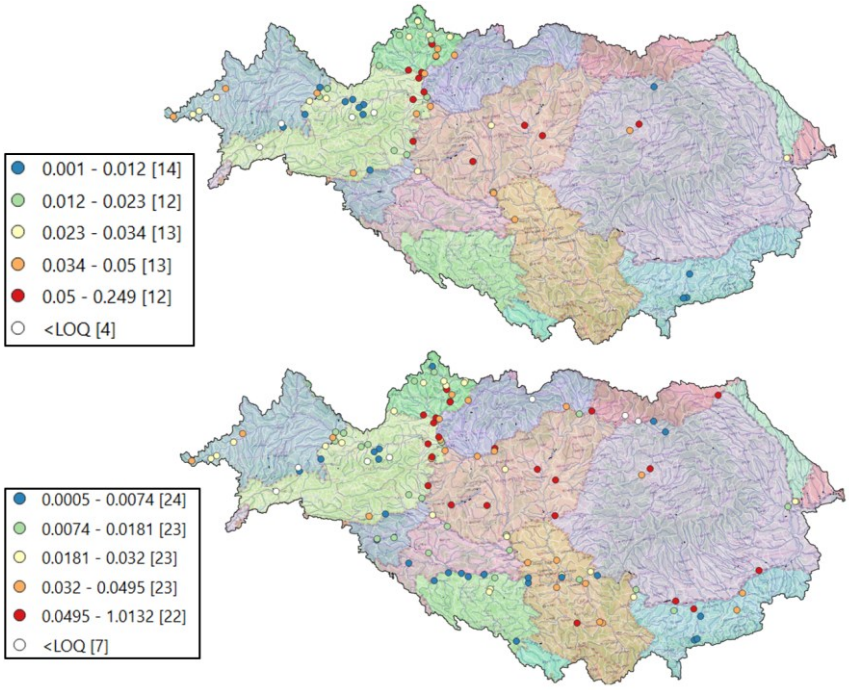

Figure 63 - Average concentration – CBZ [µg/L], before and after considering Tethys monitoring points

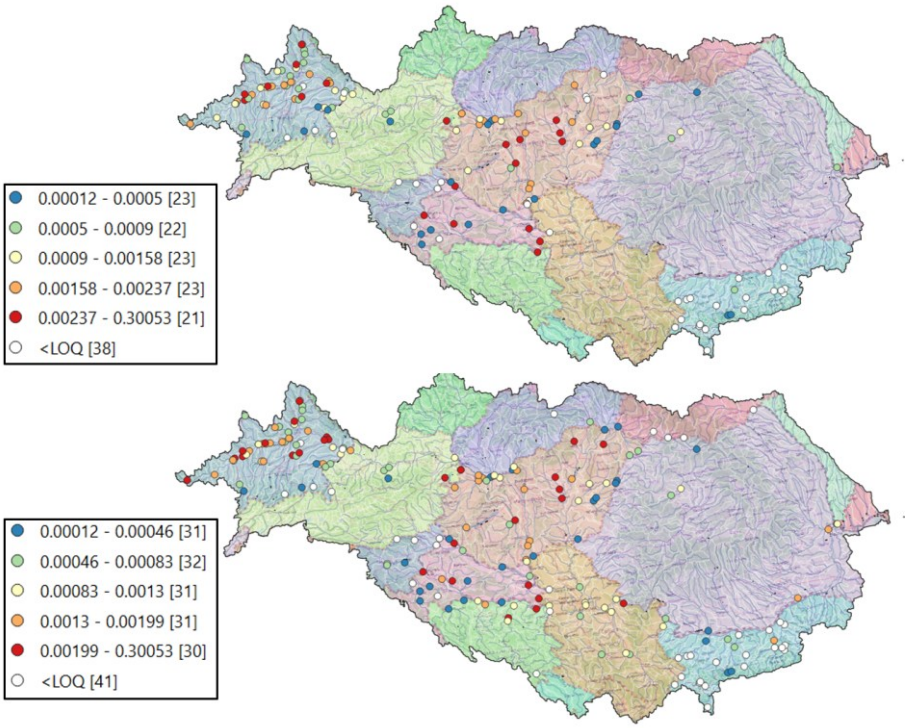

Figure 64 - Average concentration – PFOS [µg/L], before and after considering Tethys monitoring points

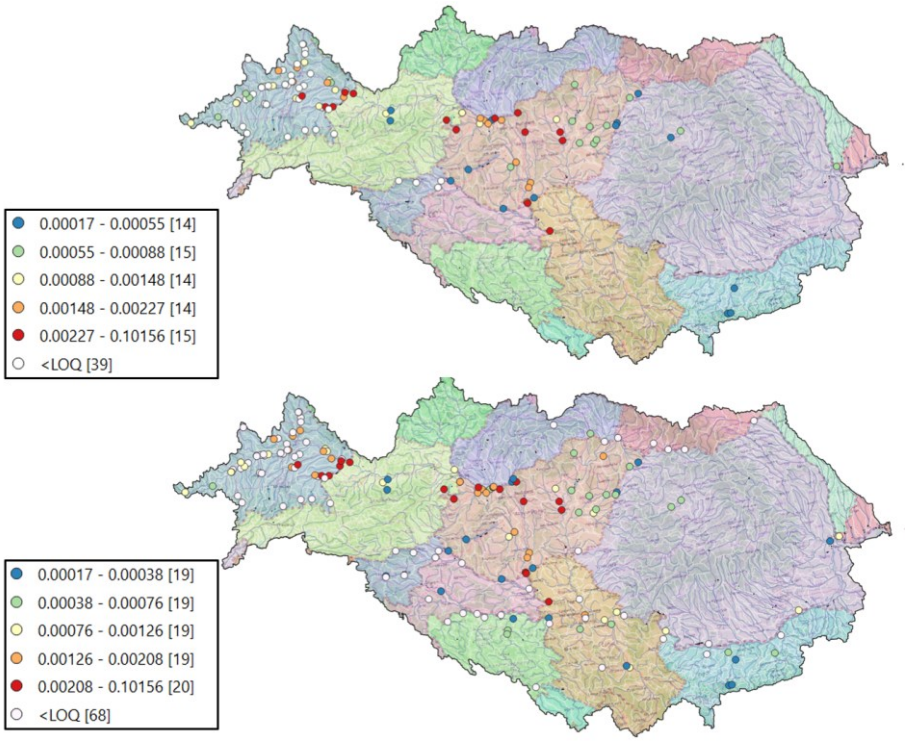

Figure 65 - Average concentration – PFOA [µg/L], before and after considering Tethys monitoring points

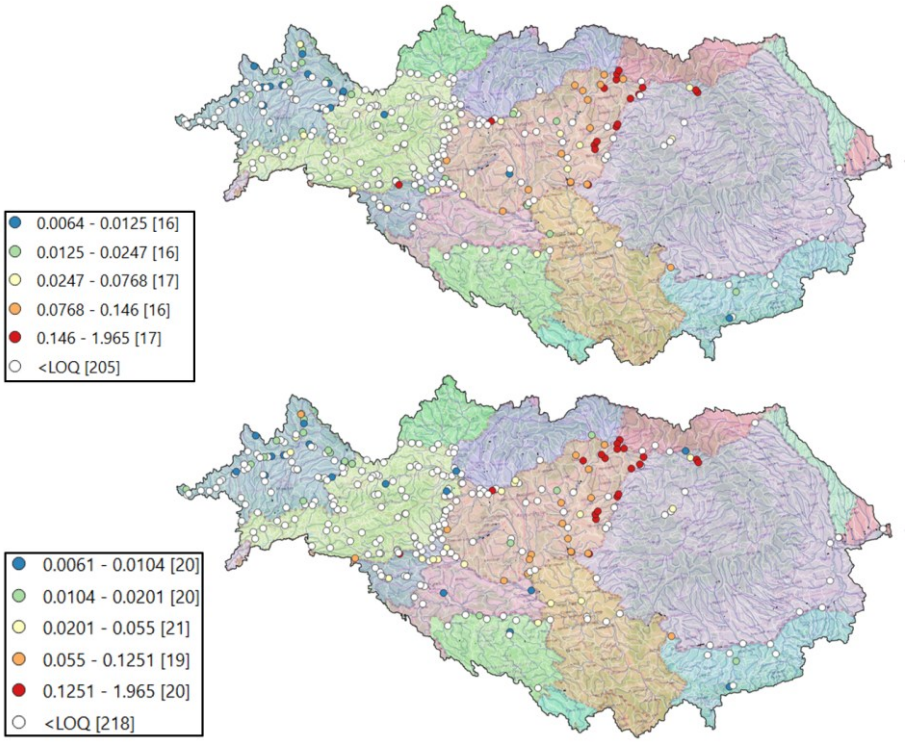

Figure 66 - Average concentration – Cadmium dissolved [µg/L], before and after considering Tethys monitoring points

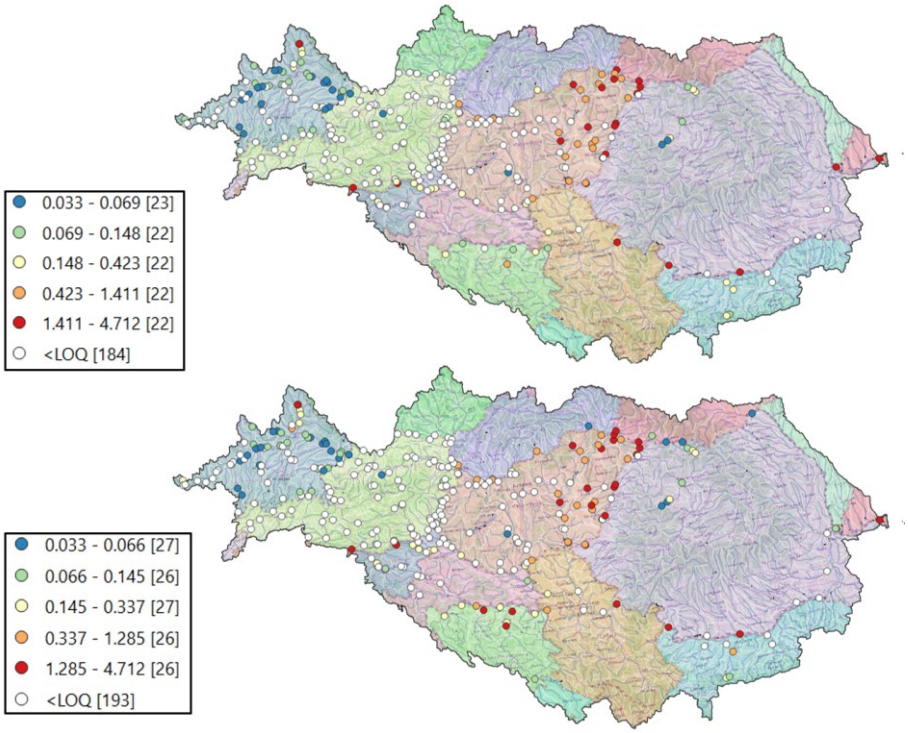

Figure 67 - Average concentration – Lead dissolved [µg/L], before and after considering Tethys monitoring points

7 CONCLUSIONS

The testing and demonstration activities confirmed that the HS emission model, originally developed within the DTP Danube Hazard m³c project, has been successfully advanced into an operative, fit-for-purpose tool for transnational, modelling-based risk and scenario assessment under new and emerging challenges. Through the joint efforts of the project partners, the model was substantially extended to cover the entire Danube River Basin, enabling consistent and comparable assessments across national boundaries.

A key achievement was the alignment of the model with an updated and harmonized input data structure. This included the integration of trans-European databases operated by the European Environment Agency, such as those related to urban wastewater treatment and industrial emissions, as well as datasets provided by the Joint Research Centre, including erosion maps. Most importantly, the model was linked to the transnational database developed under Activity 1.5, ensuring a common and coherent data basis for model application, testing, and demonstration across the basin.

In parallel, new emission calculation algorithms were investigated and implemented. The calculation framework was further strengthened through the enhancement of the retention module, allowing a more realistic representation of substance fate and attenuation processes. Together, these improvements enabled the model to address a wider range of substances, pathways, and pressures, and to support complex scenario analyses relevant for policy and management purposes.

The demonstration exercises showed that the upgraded HS emission model is suitable for performing risk assessments and evaluating scenarios at the transnational scale, even under conditions of increased data complexity and uncertainty. At the same time, remaining challenges were identified, particularly the effort required to integrate input data from non-EU countries and the technical integration of the enhanced retention module into the original software environment. Addressing these aspects will further improve the robustness, usability, and long-term applicability of the model for basin-wide environmental assessments.

8 ANNEXES

The annexes can be found in separate files.

ANNEX I - MORE MODEL ALGORITHM

ANNEX II – COMPARISON OF THE BREFS' EMISSION LEVELS **ASSOCIATED AND THE** AVAILABLE INDUSTRIAL EMISSION NATIONAL DATA

ANNEX III – OVVERVIEW TABLE OF ALL MODEL VARIABLES BY PATHWAY