

Output T4.5

Technical Guidance Manual

of the best practices on hazardous substances pollution management in water for stakeholders

March 2023

PROJECT TITLE: Tackling hazardous substances pollution in the Danube River Basin by Measuring, Modelling-based Management and Capacity building

ACRONYM: Danube Hazard m³c

DATE OF PREPARATION: March 2023

AUTHORS AND CONTRIBUTING PARTNERS

Name co-author	Contributing partner
Elena Țuchiu	National Administration Romanian Waters (NARW), RO
Elvira Marchidan	National Administration Romanian Waters (NARW), RO
Corina Boscornea	National Administration Romanian Waters (NARW), RO
Ioana Nedelea	National Administration Romanian Waters (NARW), RO
Ramona Curelea Marin	National Administration Romanian Waters (NARW), RO
Adrienne Clement	Budapest University of Technology and Economics (BME), HU
Katalin Mária Dudás	Budapest University of Technology and Economics (BME), HU
Máté Krisztián Kardos	Budapest University of Technology and Economics (BME), HU
Zsolt Jolankai	Budapest University of Technology and Economics (BME), HU
Oliver Gabriel	Environment Agency Austria (Umweltbundesamt), AT
Marianne Bertine Broer	Environment Agency Austria (Umweltbundesamt), AT
Radmila Milačič	Jožef Stefan Institute (JSI), SI
Adam Kovacs	ICPDR
Jos van Gils	Deltares, The Netherlands (subcontracted by ICPDR)
Sibren Loos	Deltares, The Netherlands (subcontracted by ICPDR)
Renata Kaps	TU Wien, AT
Steffen Kittlaus	TU Wien, AT
Nikolaus Weber	TU Wien, AT
Matthias Zessner	TU Wien, AT
Ottavia Zoboli	TU Wien, AT

Responsible for the Output: Elvira Marchidan (National Administration Romanian Waters)

Table of contents

Executive summary.....	5
1 Introduction and background.....	6
2 Targeted monitoring as basis for inventorying	7
2.1 General aspects.....	7
2.1.1 Scope and objectives of the inventory-supporting monitoring.....	7
2.1.2 Monitoring performed in the Danube Hazard m3c to support inventory	8
2.1.3 Selection of substances.....	11
2.1.4 Spatial scale and designation of monitoring sites.....	12
2.1.5 Composite sampling approach.....	14
2.2 Sampling of different matrices.....	14
2.2.1 River water.....	14
2.2.1 Stratified composite sampling of rivers to estimate riverine load	18
2.2.2 Continuous online measurements in rivers.....	19
2.2.3 Suspended particulate matter (SPM).....	22
2.2.4 Atmospheric deposition.....	25
2.2.5 Wastewater	28
2.2.6 Soil.....	30
2.3 Sample preparation, storage and transport to the laboratory.....	33
2.3.1 Preservation of samples.....	33
2.3.2 Procedures for cleaning of bottles and vessels	35
2.3.3 Transport of the samples to the lab	36
2.3 Laboratory methods.....	36
3 Inventory of hazardous substances - relational and harmonized database.....	41
3.1 Inventory of HS emissions	41
3.1.1 Setting up the inventory - scope and objectives.....	41
3.1.2 Considered ideas and requirements	42
3.2 Implementation of the inventory	43
3.2.1 Data collection and import.....	43
3.2.2 Data storage in a well-structured SQL database.....	44
3.2.3 Data evaluation	54
3.3 Scope of use of the information in the relational database. Application examples	56
3.3.1 Comparison of HS concentrations in surface water bodies with EQS.....	56

3.3.2	Model input data and improved system understanding.....	63
3.3.3	Model validation data.....	69
3.4	Lessons learned.....	69
4	Modelling and scenarios development for hazardous substances water pollution.....	71
4.1	Introduction.....	71
4.1.1	Why modelling?.....	71
4.1.2	Existing modelling approaches.....	73
4.2	MoRE model.....	74
4.2.1	Pathway-oriented approach.....	74
4.2.2	Scope, spatial and temporal scales.....	76
4.2.3	Main calculation approaches.....	77
4.2.4	Data requirements.....	80
4.2.5	Technical requirements.....	83
4.2.6	Selection of examples of application and of results in pilot regions.....	83
4.3	Danube Hazard Substance Model.....	89
4.3.1	Source-oriented approach.....	89
4.3.2	Scope, spatial and temporal scales.....	91
4.3.3	Main calculation approaches.....	92
4.3.4	Data requirements.....	94
4.3.5	Technical requirements.....	95
4.3.6	Selection of examples of application and of results at DRB scale.....	96
4.3.7	Interpretation of results at DRB scale.....	101
4.4	Implementation of scenarios to mitigate HS pollution.....	101
4.4.1	Pollution control measures.....	101
4.4.2	Other scenarios.....	108
4.4.3	Stakeholder involvement.....	109
5	References.....	111

Executive summary

To achieve the goals of the Water Framework Directive for surface water bodies, Member States must adopt measures to progressively reduce pollution from priority substances and cease or phase out emissions, discharges, and losses of priority hazardous substances. To tackle, reduce or even eliminate emissions, it is first necessary to identify and quantify them and to prioritize which emission pathways are most relevant for critical substances impairing the status of water bodies.

This Technical Guidance Document has been produced by the Danube Hazard m³c project to provide support to managers in the field of water quality across the Danube River Basin in designing and setting up an integrated strategy to efficiently generate exhaustive and harmonized emissions inventories of hazardous substances (HS) at national level, with focus on their compatibility and value for an effective transboundary pollution control.

This document merges recommendations based on the scientific state of the art of knowledge in the field with examples, experiences and lessons learned during the implementation of activities in the project itself.

The technical guidance is structured in three main chapters, which address the three main pillars needed to generate emission inventories:

- Monitoring – strategies, methodologies and devices to gather the necessary information on occurrence and concentration level of HS in rivers for a reliable estimation of riverine loads and in different compartments, which contribute to the most significant emission pathways
- Data base – conceptualization, technical implementation and examples of added value and usages of relational data bases containing the information of HS concentration in rivers and in the most important environmental and engineered compartments
- Emission models – types, scope, technical requirements and usage of emission modelling, with focus on the two models further developed and applied in the project, namely MoRE model (example for pathway-oriented approach) and DHSM model (example for source-oriented approach).

To whom is the technical guidance addressed?

- This Technical Guidance Manual addresses decision makers and technical experts from authorities responsible for water management at the national and Danube River Basin scale, including representatives from ICPDR and its expert groups, interest groups from planners, technology suppliers, industry, agriculture, and associations for wastewater treatment, environmental NGOs, and representatives from higher education and research institutions within the DRB.

1 Introduction and background

The EU Water Framework Directive (WFD 2000/60/EC) has the purpose to establish a framework for the protection and enhancement of waters and to ensure sustainable use of water resources. The main aim of the WFD is to achieve a good ecological and chemical status in all water bodies and to prevent deterioration of the good status of all water bodies. The good chemical status of surface water bodies is achieved by a body of surface water in which concentrations of selected contaminants do not exceed the environmental quality standards (EQSs) established in Annex I of the EQS Directive (Directive 2008/105/EC, as amended by Directive 2013/39/UE). Member States have to identify, control and reduce or eliminate the emissions of priority substances so to achieve a good chemical status in their water bodies. In this context, Art. 5 of the EQS Directive stipulates that Member States are obliged to establish an inventory of emissions, discharges and losses (IEDL) of all priority substances and all pollutants listed in Part A, Annex I of the Directive. The Technical Guidance Document No. 28 “Preparation of an Inventory of Emissions, Discharges and Losses of Priority and Priority Hazardous Substances” recommends a tiered approach for establishing such inventories at the river basin level.

As described in detail in the project output O.T3.1 *Report on existing policies and management plans regarding HS pollution in the DRB*, the status of development of these emission inventories within the Danube River Basin is still at an early and insufficient stage. Only in three countries (Germany, Austria and Hungary), the third tier was applied, by identifying specific diffuse emission pathways via modelling approaches. In most countries only the first and second tiers have been applied so far, which only allow for the identification of major point emission pathways (municipal and industrial wastewater treatment plants), while diffuse emission can only be estimated as black-box via difference between riverine loads and point emissions. A further gap identified all over the Danube River Basin is the lack of targeted monitoring at high-flow events, which leads to the inaccurate calculation of riverine loads for many substances and thus to the likely underestimation of the emission loads via diffuse pathways. Last, a major gap and challenge towards a coordinated control of HS pollution in the Danube River Basin was identified in the lack of harmonized methods and lack of institutional competencies and skills with regard to monitoring and setup of inventories for HS emissions. The criticality of these gaps has recently been exacerbated by the EU proposals for the new Urban Wastewater Treatment Directive, new Water Framework Directive, Groundwater Directive and Environmental Quality Standard Directive, which introduce the necessity of performing risk assessment and complex integrated evaluations at river basin scale, which are not possible without the solid basis of data and system understanding provided by extensive emission inventories.

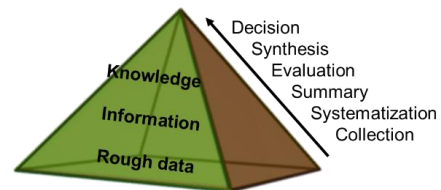
In such context, this technical guidance manual aims at supporting water quality managers in the DRB in designing, selecting and implementing activities aimed at generating solid and thorough emission inventories for hazardous substances at national level, with a focus on important aspects in view of the harmonization of the inventories at transboundary scale and of the optimization of trade-offs between needed resources and expected gain of information. This document is structured in three main chapters, which address the three pillars leading to emission inventories, namely monitoring, set up of data bases and emission modelling. This guidance merges recommendations based on the state of the art of knowledge with examples and lessons learned from the activities developed, implemented and thus tested in the Danube Hazard m³c project.

2 Targeted monitoring as basis for inventorying

2.1 General aspects

Monitoring is a well-organized regular activity which is intended for a specific goal. In general, water quality monitoring aims at evaluating the status quo, studying and predicting its change and finding the causes of its deterioration.

Monitoring is much more than simple data collection. It is a comprehensive activity, which ranges from sampling and lab analysis to the statistical evaluation of the results. Collected data should provide information which can support decisions.



The monitoring system must meet several criteria, the most important of which are:

- representativeness (temporal and spatial variability)
- reproducibility of the applied methodologies
- continuity of the measurements
- flexibility
- cost-effectiveness.

Monitoring is needed to fill knowledge gaps, to know the status quo and to track potential changes in the status of the environment, and the consequences of the measures taken. It is also needed for the development, refinement, calibration and validation of models aiming at understanding the processes describing the release, spreading and transformation of contaminants.

2.1.1 Scope and objectives of the inventory-supporting monitoring

In line with the WFD, Member States have to perform surveillance monitoring of hazardous substances (HS) to assess the chemical status of the water bodies. Specific goals are trend detection and assessing the effect of implementing measures (operational monitoring). The shortage of these programmes is that this type of monitoring approach does not provide information about riverine loads, pollution sources and emission pathways. To fulfil the objectives of the emission inventory, a different approach is needed. Inventory-supporting monitoring is designed to measure concentrations and loads in rivers and in relevant pathways (both point and diffuse).

WFD monitoring

Main goal: status assessment

Specific goals: trend detection (surveillance monitoring), assessing the effect of the implementation of measures (operational monitoring)

Does not provide information about riverine loads and pollution sources and emission pathways!

Inventory-supporting monitoring

Main goal: load calculation and support of modelling

Conceived to measure concentrations and loads in rivers and in different point and diffuse emission pathways

It provides information for substance balances and input for modelling.

Hazardous substances released into the environment as a consequence of human activities can be transported into surface and ground water from numerous sources via different point and diffuse pathways, as shown in Figure 1.

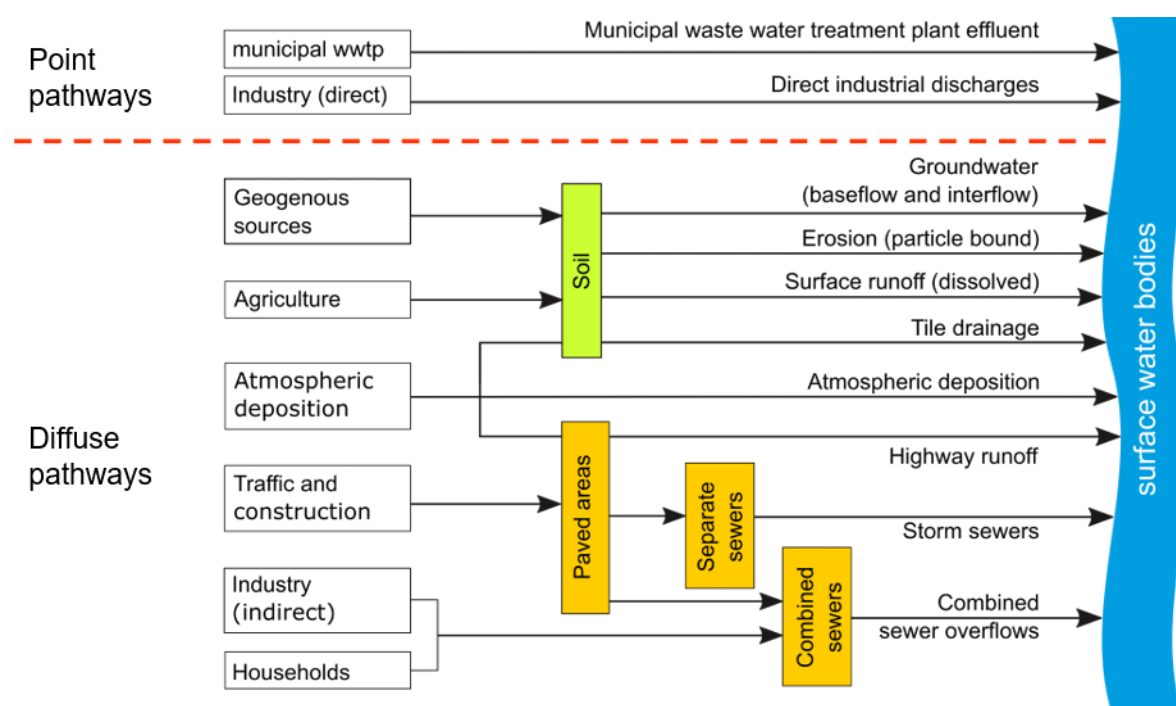


Figure 1: Sources and pathways of HS into the aquatic environment (S. Kittlaus CC-BY-SA 4.0)

2.1.2 Monitoring performed in the Danube Hazard m3c to support inventory

The measurement activities within Danube Hazard m³c (described in detail in the Output O.T1.2 *Demonstration of a harmonized and cost-effective monitoring*) showcase the design and execution over one year of such monitoring aimed at supporting the creation of an emission inventory.

Besides being focused on collecting the necessary information for accurate and representative estimations of river concentrations and load of micropollutants, the project monitoring aimed at gathering the necessary information for the further development and validation of two models (MoRE and DHSM), which estimate the contribution of different point and diffuse emission pathways at pilot catchment and Danube River Basin scale, respectively. Table 1 summarizes the substance-specific input data required by the model MoRE. The substance-specific input required by the DHSM model (reported in

Table 2) does not need data deriving from the monitoring programme, but the validation of its intermediate results does. These are summarized in Table 3. Given the above, a targeted monitoring concept has been designed so to enable the quantification of river concentrations and river loads of selected contaminants across diverse catchments and which considers the data requirements of the MoRE and DHSM models.

Table 1: Requirements of substance-specific data for the MoRE model. AU: analytical unit.

Type of pathway	Pathway	Input data	Spatial scale	Temporal scale
Point	Municipal WWTP effluent	Effluent loads OR water amount and effluent concentration	For each plant or lumped over AU	Annual average
Point	Industrial WWTP effluent or direct industrial discharge	Effluent loads OR water amount and effluent concentration	For each plant or lumped over AU	Annual average
Point	Abandoned mining site	Effluent loads OR water amount and effluent concentration	For each site or lumped over AU	Annual average
Diffuse	Agricultural erosion	Soil content in agricultural land	Lumped over AU	Current conc. level
Diffuse	Erosion from natural soils	Soil content in naturally covered land	Lumped over AU	Current conc. level
Diffuse	Surface runoff from pervious soils	Concentration in surface runoff from pervious land	Lumped over AU	Annual average
Diffuse	Tile drainage	Concentration in tile drainage discharge	Lumped over AU	Annual average
Diffuse	Groundwater	Concentration in groundwater	Lumped over AU	Annual average
Diffuse	Atmospheric deposition	Deposition rate OR Concentration in rainwater and amount of precip.	Lumped over AU	Annual average
Diffuse	Discharge through combined sewer overflows	Concentration in combined sewer overflows	Lumped over AU	Annual average
Diffuse	Discharge through storm sewer outlets	Concentration in storm sewer outlets	Lumped over AU	Annual average
Diffuse	Inland navigation	Emissions loads via steel construction for hydraulic engineering	Lumped over AU	Annual average
Diffuse	Inland navigation	Emissions loads via motor boat exhaust	Lumped over AU	Annual average

In addition, the validation of MoRE requires annual average concentrations or the total annual load of contaminants in rivers at the outlet of each analytical unit employed for the validation.

Table 2: Requirements of substance-specific data for the DHSM model. SC = sub-catchment, basic spatial unit derived from the underlying hydrological model

Type of pathway	Pathway	Input data	Spatial scale	Temporal scale
Point & Diffuse	Wastewater	Use volume and use type of chemical, population density map, wastewater management maps (connection to sewers, treatment level)	Lumped per SC, use volume optionally per country or even on EU level	Annual average
Point & Diffuse	Stormwater	Use volume and use type of chemical, population density map, paved area map, combined-/separated sewers map	Lumped per SC, use volume optionally per country or even on EU level	Annual average
Point	Abandoned mining site	Effluent loads OR water amount and effluent concentration		Annual average
Diffuse	Agricultural emissions (pesticides)	Amount used	Country level or finer if available	Annual average
Diffuse	Atmospheric deposition	Deposition rate	Lumped per SC	Annual average
Diffuse	Inland navigation	Emissions via steel construction for hydraulic engineering	Lumped per SC	Annual average
Diffuse	Inland navigation	Emissions via motor boats	Lumped per SC	Annual average

Table 3: Intermediate results of the DHSM model for which validation data are required. AU: analytical unit.

Type of pathway	Pathway	Validation data	Spatial scale	Temporal scale
Point	Municipal WWTP effluent	Effluent loads OR water amount and effluent concentration	For each plant or lumped over AU	Annual average
Point	Industrial WWTP effluent or direct industrial discharge	Effluent loads OR water amount and effluent concentration	For each plant or lumped over AU	Annual average
Diffuse	Agricultural erosion	Soil content in agricultural land	Lumped over AU	Current conc. level
Diffuse	Erosion from natural soils	Soil content in naturally covered land	Lumped over AU	Current conc. level
Diffuse	Surface runoff from pervious soils	Concentration in surface runoff from pervious land	Lumped over AU	Annual average
Diffuse	Tile drainage	Concentration in tile drainage discharge	Lumped over AU	Annual average
Diffuse	Groundwater	Concentration in groundwater	Lumped over AU	Annual average
Diffuse	Discharge through combined sewer overflows	Concentration in combined sewer overflows	Lumped over AU	Annual average

Type of pathway	Pathway	Validation data	Spatial scale	Temporal scale
Diffuse	Discharge through storm sewer outlets	Concentration in storm sewer outlets	Lumped over AU	Annual average

2.1.3 Selection of substances

Considering the high financial cost of the chemical analyses for micropollutants, it is advisable to select for the inventory-supporting monitoring indicator substances, which can be considered representative of sources and emission pathways and of specific environmental fate patterns. The selection concept applied in the Danube Hazard m³c followed this rationale. Bearing in mind the large number of HS present in the water bodies, the project focused on 46 indicator substances from 5 different substance groups of high relevance in the Danube River Basin, which are representative of different major sources and emission pathways: (a) substances of both natural and anthropogenic origin; (b) industrial chemicals; (c) substances of intensive agricultural use and (d) pharmaceuticals. The final selection is listed in Table 4.

Table 4: Substances measured in the Danube Hazard m3c project.

Group of substances	Substances	Typical emission pathway
Combustion products	16 PAHs: Acenaphthene, Acenaphthylene, Anthracene, Benzo(a)anthracene, Benzo(a)pyrene, Benzo(b)fluoranthene, Benzo(k)fluoranthene, Benzo(g,h,i)perylene, Chrysene, Dibenzo(a,h)anthracene, Fluoranthene, Fluorene, Indeno(1,2,3-cd)pyrene, Naphthalene, Phenanthrene, Pyrene	Atmospheric deposition
Potentially toxic metals	Hg, Cd, Ni, Pb, Cr, Cu, Zn, As	Soil Mining tailings Wastewater
Agricultural fungicides	Tebuconazole	Soil
Agricultural herbicides	Metolachlor, Metolachlor -ESA, Metolachlor – OA	Soil
Pharmaceuticals	Diclofenac Carbamazepine	Municipal wastewater
Industrial chemicals	Perfluorooctane sulfonate (PFOS) Perfluorooctanoic acid (PFOA) Perfluoropentanoic acid (PFPeA) Perfluorohexanoic acid (PFHxA) Perfluoroheptanoic acid (PFHpA) Perfluorononanoic acid (PFNA) Perfluorodecanoic acid (PFDA) Perfluoroundecanoic acid (PFUnA) Perfluorododecanoic acid (PFDoA) Perfluorotridecanoic acid (PFTTrA) Perfluorotetradecanoic acid (PFTeA)	Municipal and industrial wastewater Diffuse emissions

Group of substances	Substances	Typical emission pathway
	Perfluorobutane sulfonic acid (PFBS) Perfluorohexane sulfonic acid (PFHxS)	
Industrial chemicals	Octylphenol Nonylphenol Bisphenol-A	Wastewater

2.1.4 *Spatial scale and designation of monitoring sites*

The emission inventory is based on establishing a mass balance, which is typically set up on (sub-) catchment scale. The riverine load can be measured at the outflow sections based on concentration in water (and in the suspended solids) and discharge data in rivers, considering the basic processes of transport, storage or temporary storage and degradation of substances. The resulting riverine load provides information about the current pollution status and temporal trends in case of long-term information. In combination with the information gained in the inventory of point source emissions, it allows estimating which share of load derives from diffuse emissions (this so-called „riverine load approach” would be the first step towards emission modelling). The inventory-supporting monitoring concept aims to get one step beyond by measuring concentrations and loads in different pathways. In this way, significant transport routes in the catchments as well as the relevance of point and diffuse pathways can be identified.

The inventory approach leads to some important conclusions for the monitoring design (see Figure 2):

- Catchment borders must be considered for the territorial extent of the inventory-supporting monitoring
- River monitoring sites must be located at the catchment outflow section (or outflows of the sub-catchments in case of further subdivision)
- In the river, concentration and discharge should be measured simultaneously to determine the riverine load (sampling sites should be located at or close to river gauges)
- To properly understand and estimate yearly loads of contaminants, high-flow events must be included in the sampling design
- Emissions point sources must be quantified, therefore significant municipal and direct industrial wastewater effluents (volume, concentration) must be measured.
- The measurements should cover other transport pathways that might be relevant in the catchment and for the substances of interest (e.g. atmospheric deposition, soil, groundwater, surface runoff, stormwater overflows, etc.).

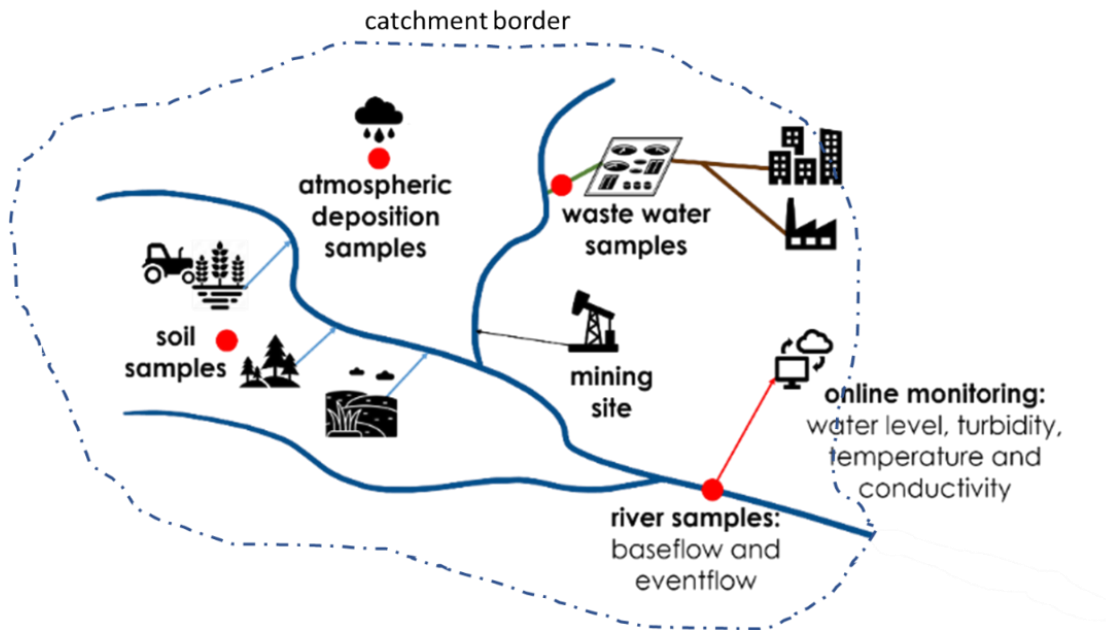


Figure 2: Monitoring concept applied in the project.

Figure 3 demonstrates an example of selecting monitoring sites in a medium-sized catchment (Upper-Zagyva in Hungary, area: 1216 km²). The catchment is divided into four sub-catchments in the applied modelling approach. In the outlets, river monitoring sites were planned at existing flow gauges. Sub-catchments represent different intensities of anthropogenic impacts: (i) the headwater area is covered mainly by forests and natural vegetation, (ii) one tributary receives significant wastewater discharge, (iii) one tributary where arable land dominates land use, and (iv) the lower part representing mixed conditions and the outflow.

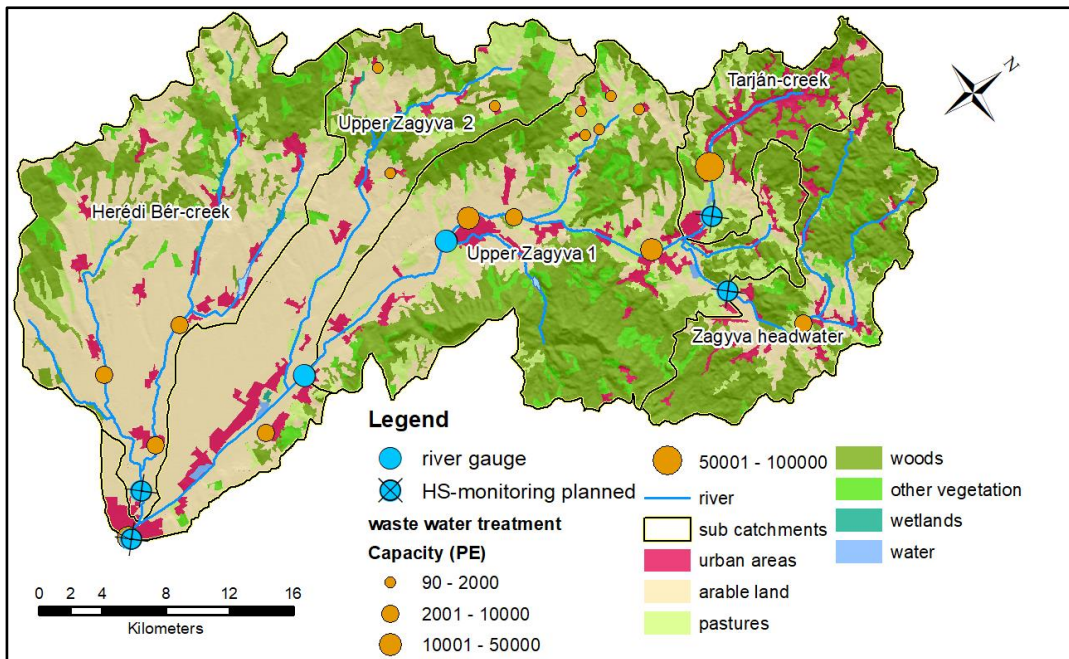


Figure 3: Monitoring sites planned for implementing the inventory support monitoring (Upper-Zagyva catchment, Hungary)

2.1.5 Composite sampling approach

A way to design a cost-efficient measuring concept is to build it on time or space-integrated composite sampling as well as stratified sampling focusing on flow, climatological or land use patterns and considering specific pressures in order to optimize the trade-off between obtained information and number of samples.

The following sections describe adequate sampling strategies and equipment for the monitoring of each compartment.

2.2 Sampling of different matrices

2.2.1 River water

Grab (spot) sampling

The most typical sampling method consists of the lab personnel travelling to the sampling point, taking a limited amount of the matrix in some bottle/case, and analysing it in the lab. The efficiency of this sampling method highly depends on the applied frequency, which has its limits. Traditional monitoring programmes rely predominantly on this approach, which can provide reliable information about the low- and medium-flow conditions in rivers. However, infrequent observations cause errors in the estimation of material flows, because the amount transported during large runoff events is underrepresented in the time series.

Equipment needed for spot sampling (Figure 4):

- A vessel mounted in a telescopic holder, or a simple bucket
- Immerse an open-mouthed bottle of sampling system into a flow stream approximately 30 cm below the river surface
- The inlet of the sampling bottle should face the direction of the river flow
- Bottles should be filled maximum of 85% of the bottle volume.



Figure 4: Spot sampling.

During high-flow events, important processes take place in river catchments, which determine changes in the relevance of emission pathways, such as soil erosion and dilution of point source emissions. Therefore, it is essential to gather information on the concentration of contaminants during such events.

In this context, the aim is to sample rivers when the flow rate is comprised between Q₀ and Q₁₀ (0-10% percentage of exceedance), and turbidity rises significantly above its typical baseflow/mid-flow level. Each event should possibly be sampled flow-proportionally during its whole duration, i.e. both in the rising and descending part of the hydrograph, and the obtained composite samples must be cooled and sent immediately to the laboratories in the provided bottles.

The sampling can either be carried out manually or via autosamplers. In the case of auto sampling, the sampler has to be controlled by the turbidity probe signal in a way that the sample volume is adjusted proportionally to the flow.

Composite samples and automated sampling

This approach means that by the extension of grab sampling, samples are grabbed with a higher frequency, mixed together, and the so-generated mixed sample (composite) is analysed. This is usually done with an automated sampling device.

The concentration values of the composite sample are representative of the whole period in which the samples were taken. Samples can be added to the composite at equal time intervals (time proportional, e.g. every hour) or depending on some hydrological property (e.g. flow proportional). In

the second case, automated flow measurements are needed at the sampling location. Linking sampling frequency to turbidity values (measurable online) is also a possibility (Lewis, 1996).

Requirements for the autosamplers (Figure 5) are as follows:

- Enable time-, volume-, flow-, and event-proportional sampling (the latter requires continuous detection of the water level (or other trigger parameters) and must be programmable – see Figure 8.),
- 12...24 vessels or a composite container are used,
- Temperature control for the sample storage chamber (passive cooling can be applied, too – see Figure 6)
- Power supply required (AC, battery, solar panel),
- Regular maintenance must be provided!



Figure 5: Portable (left) and stationary (right) autosamplers .



Figure 6: Passive cooling of portable sampler .

Autosamplers are widely used for getting series of samples. A proper selection of different automatic samplers is available on the market; however, their price or some of the technical limitations might

pose difficulties for particular research aims. If the goal is to gain information on the intra-event dynamics, a series sampler is needed and many samples will have to be analysed. However, for the quantification of interevent variability, composite samplers are sufficient, on condition that they can adjust the pumping rate to be proportional to the flow during the sampling process (Budai et al, 2020). The concentration of substances in such samples will represent the so-called event mean concentration (EMC) which is a widely accepted parameter used to characterize unique events. (Göbel et al. 2007). As flow-proportional composite samples inherently yield the EMCs (Gasperi et al. 2014) they allow for the calculation of the total event loads as well (by multiplication with the event runoff).

Autonomous flow-proportional water sampler was developed by Budai et al., 2020, for the continuous composite sampling of runoff events (Figure 7). The equipment aims:

- Continuous water level monitoring and recording in user-defined time intervals,
- Sampling driven by a flexible, user-specified program, based on actual measured water levels,
- Inexpensive and simple system parts that are easily replaceable or repairable on-site,
- Low-energy consumption (in order to enable off-grid deployment).

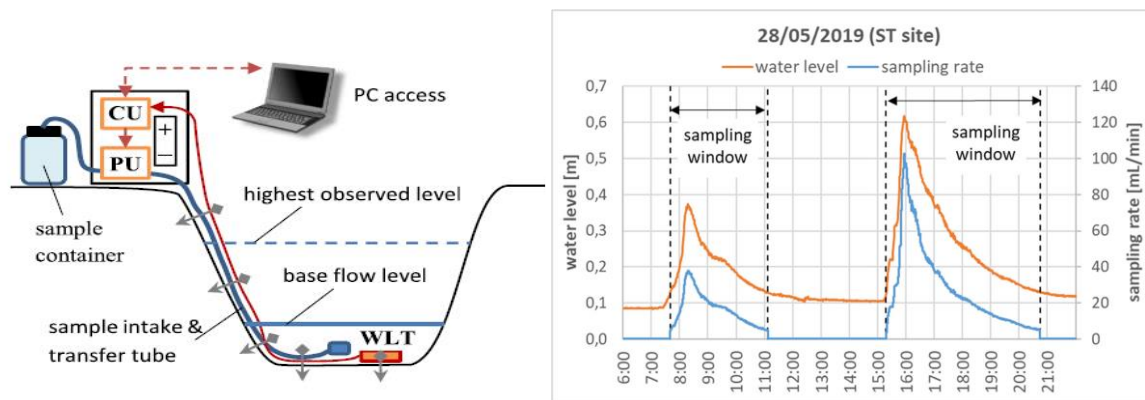


Figure 7: Autonomous flow-proportional water sampler: system layout in a river section (left), an example demonstrating the sampling characteristics for two consecutive runoff events (right).

Flow & Turbidity threshold sampling uses real-time turbidity and river level information to automatically collect targeted water quality samples during high flow events and to estimate suspended sediment loads during a specific time period. The system uses a programmable data logger in conjunction with a water level measurement device, a turbidity sensor, and a pumping sampler. Specialized software enables the user to control the sampling process by setting threshold values for sample collection.

Thresholds are usually chosen so that the square roots of NTU values are evenly spaced to adequately define loads for small storms without oversampling large storms. A programmable data logger, typically recording at 10- or 15-minute intervals, instructs an automatic pumping sampler to collect a sample whenever a threshold is crossed (Lewis and Eads, 2009).

2.2.1 Stratified composite sampling of rivers to estimate riverine load

The general concept for river sampling is to collect information on pollutant levels separately for base flow conditions and high-flow conditions as for some pollutants the concentrations (and therefore the loads) in these conditions are significantly different due to sediment bound transport from the catchment, mobilisation of urban sources or dilution of point source emissions.

This so-called stratified composite sampling approach – which was applied in the DH m³c – is illustrated in Figure 8 and Figure 9). Base flow concentrations were determined by a large set of weekly baseflow grab samples collected during baseflow-midflow conditions. High-flow samples were collected in two ways in the pilot catchments, either (1) by using autosamplers, which either collected pre-programmed series of samples or a composite sample by varying pumping rate and (2) a series of grab samples collected during flood events.

Lowflow sampling has to be continued for a minimum one-year long time at river stations to cover time variability and seasonality. As far as high-flow sampling is concerned, the aim is to sample rivers when the flow rate is comprised between Q_0 and Q_{10} (0-10% percentage of exceedance), and turbidity rises significantly above its typical baseflow/mid-flow level.

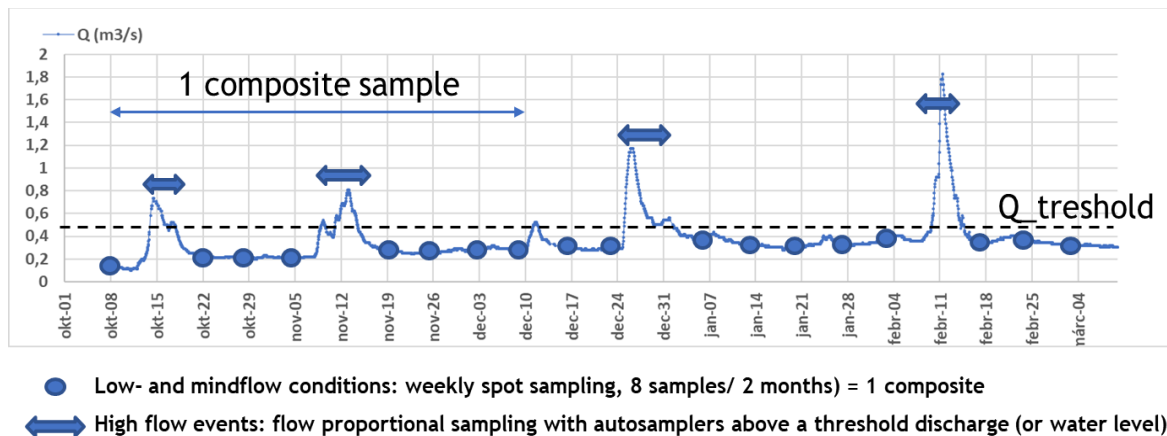


Figure 8: The concept of the stratified composite sampling approach applied in the DH m³c project.

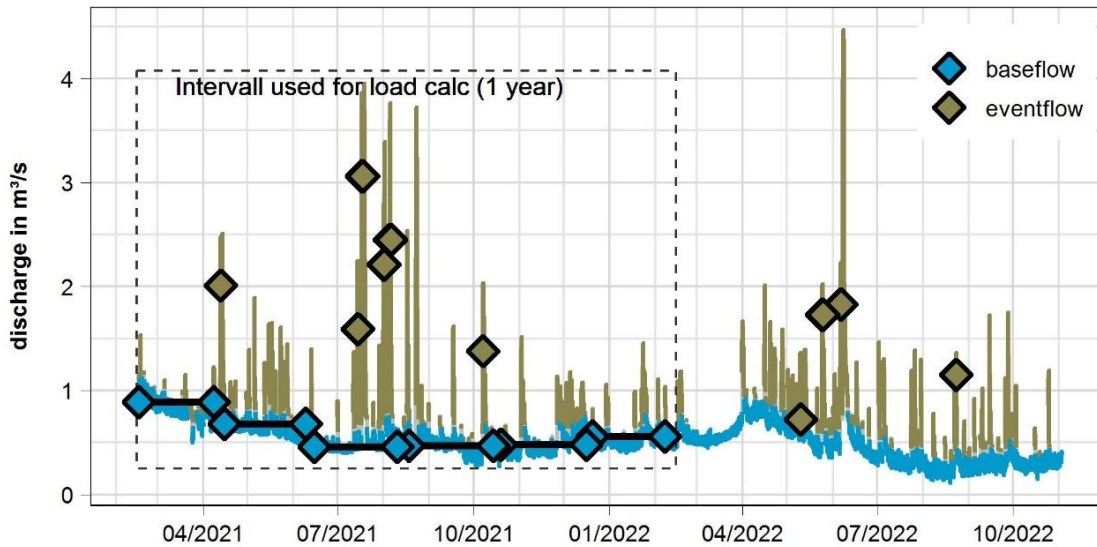


Figure 9: Example for the sampling: discharge timeline of Wulka river (AT) with colour according to flow condition, diamonds show the composite samples used for concentrations data.

Preparing composites samples

The preparation of the composite samples requires extra work in the lab. After each weekly sampling, an aliquot of the fresh sample is poured into the bottle/container in the fridge/freezer (depending on cooling temperature) and if necessary a proportional amount of stabilizing compound is added.

Contrary to the typical protocol, which foresees that chemical analysis must be performed immediately, samples will be processed with a significant time delay. Therefore, proper storage and preservation must be ensured. This issue will be discussed later in a separate chapter.

2.2.2 Continuous online measurements in rivers

Continuous measurements of water level, temperature, conductivity, and turbidity are recommended throughout the whole year. One of the cornerstones of the DH m³c monitoring program was to install the online sensors at the key points of the pilot regions to get continuous information about important parameters for a better understanding of the process dynamics in the catchments and for more accurate riverine load calculations.

Conductivity is an essential water quality indicator, easy to measure, which enables the detection of sudden and unexpected water quality changes. Additionally, conductivity enables the estimation of the share of baseflow or surface flow to total flow. Turbidity is required to estimate the total suspended particulate matter (SPM) load over the monitoring year. Turbidity – SPM correlation can be established based on the SPM concentration measurements in a sufficient number of spot samples.

To gain accurate turbidity measurements, two aspects are very important:

- The placement of the probe in the river needs to be planned and implemented very carefully because the turbidity is not equally distributed over the cross-section (Rode & Suhr, 2007). This also holds true for the sampling spots for establishing the turbidity-SPM-concentration-relationship. The probe should not be placed too close to the channel banks and not too close to the channel bed. Manual turbidity measurements (ideally with the probe to be installed) can help to identify the zone that gives a representative value for SPM. Especially in the case of larger rivers, the cross-section of the sampling spot should be measured in several vertices and several depth/vertices to gain a full picture of the sediment yield distribution. The final location should be in a position that represents the average of the section.
- Drifting of the turbidity-probe caused by the establishment of biofilm on the window of the optical probe has to be avoided by either having an automatic mechanical cleaning or regular manual cleaning. It is preferred to use a probe with an automatic wiper.

In the project, the measured parameters were water level and/or velocity, temperature, electrical conductivity and turbidity. Sensors (and automated samplers) were installed at the river monitoring stations.

The installation of the instruments proved to be challenging in several cases. Common problems were to provide sensor positioning and fixing that is representative and flood resistant at the same time. Another typical issue was to lay cables in a safe way, which was either solved by digging the cables underground or installing the cables in protection pipes. Each location provided different possibilities, construction teams had to adapt to local conditions. Examples from three stations can be seen in Figure 10, Figure 11 and Figure 12).



Figure 10: Autosampler and sampling setup at Eisbach station (Wulka, AT).



Figure 11: Autosampler and probe arrangement at Opponitz station (Upper Ybbs, AT).



Figure 12: Solar panels and wind turbine installed in Hungarian stations to provide electricity supply for sensors and autosamplers (left: Nemti, Zagyva, right: Törökkoppány, Koppány).

Ensuring a continuous and reliable operation of the online monitoring took a considerable effort in maintenance, including sensor cleaning (Figure 13), data collection and solving data transfer problems.

The other major element of operating online monitoring is data storage and processing. In the DH m³c project a large amount of data has been collected (almost 2 years of continuous data, with 1-10 min intervals).



Figure 13: Turbidity sensor before and after cleaning (left) and probe with self-cleaning windows.

2.2.3 Suspended particulate matter (SPM)

Sediments are an important source of information for the level of legacy pollution in a water body and for the internal pool of contaminants that could be mobilized or released back into the water column, whereas SPM reflects the current contamination level. Moreover, in comparison to sediments, SPM contains a higher percentage of fine-grained fraction, in which particulate-bound contaminants mainly accumulate.

The analysis of micropollutants in the solid matrix requires a minimum availability of approximately 200-250 g (dry weight) of particulate matter. This amount of sample is needed as some of the trace elements would be below the levels of detection, therefore contaminants need to be extracted and enriched from a larger mass of sediments. The collection of such an amount of SPM is very challenging and cannot be achieved through small-volume grab samples.

Devices for the collection of SPM

Different devices have been designed for the collection of relevant amounts of SPM, which range from relatively simple traps located in the river itself to relatively expensive devices located on the river banks and activated electronically.

The main advantage of simple trap or decanting devices is that they are relatively inexpensive and they mostly rely on a passive principle, i.e. they do not require electricity (Figure 14). They present, however, major disadvantages. Their performance depends on the type of river and on the grain size of the SPM. Clogging problems occur due to very fine sediment or algal growth. Further, they often do not allow collecting the finest fraction (as they do not settle), which is highly relevant for the adsorption of trace pollutants (Phillips et al., 2000). Last, a severe issue affecting their suitability for the project is that they

do not allow any back-calculation of real particulate concentrations of contaminants in the water column, because it is not possible to know the corresponding volume of water with exactitude and thus the SPM concentration in it.



Figure 14: Time-integrated Philips sampler installation to measure different stage intervals .

A rather expensive alternative, which requires electricity and automatic control, is the use of high-volume samplers. An example of these are the samplers developed and employed within the international INN Project (Kittlaus et al., 2016; Kittlaus and Fuchs, 2015), namely 1 m³ stainless-steel tanks, designed and equipped in a way that optimizes the deposition and the separate collection of SPM. The fact that the total volume of sampled water is known allows the calculation of particulate concentrations of the micropollutants in the water column. The possibility of automatic control allows the targeted sampling at different flow and turbidity conditions.

SPM sampling performed in the DH m³c project

River SPM sampling was performed at only three catchments (Wulka in AT, Koppány and Zagyva in HU) as this was not the main focus of the project, however, the collection of the fine suspended particles carries a lot of interesting information with regard to catchment transport processes.

Sampling was performed using two approaches: (1) a simple time integrated, completely handmade Philips sampler (Figure 15 and Figure 16) (Phillips, Russell and Walling 2000) and (2) flow-proportional automatic sampler (Hungarian catchment: Budai et al., 2020, Wulka: Endress+Hauser, LIQUISTATION CSF48).

The collection methods are different in many ways. The Philips samplers can be used to collect samples over a long period of time, therefore there is a bigger chance to collect sufficient amounts of sediment, especially during low-flow periods. Philips samplers, however, cannot be related to concentrations of a certain flood event, as they are not designed to collect all the SPM in the water phase (part of the fine sediments are released through the outlet hole). In other words, samples from a Philips sampler are suitable to provide a representative concentration of a given substance carried by the suspended matter, but not applicable to provide SPM concentration from a certain flood event. The automatic samplers are suitable for providing exact SPM concentrations of a flood event and therefore link

contaminant levels to it. The downside of this collection method is the limited sample volume and therefore the limited sample amount. The latter was used to create sufficient amounts for metals analysis only, and from some extremely erosive runoff events at Koppány station, for which sufficient solid matter (> 1kg) was collected to measure all the monitored substances.



Figure 15: Philips type sampler for high-flow condition at Nemti station (left) and sample collected from a flood event at Törökkoppány station (right).



Figure 16: Philips sampler at Wulka station and a 20 L glass collection container.

Sampling was successful in both cases. The experiences can be summarized as follows:

- Philips sampler devices are easily built but installation in the riverbed can be challenging. Decanting the Philips sampler on site needs two persons and a large volume sample holder. It is important to mobilize the SPM by swirling to obtain the whole sample.
- Sampling by high-flow events autosampler: a sufficient amount of sample was collected for all analyses (1-2 kg). This is site specific. This approach is recommended for locations with substantial erosive runoff events. At lower SPM concentration this will not work, so multiple flood waves should be sampled together.

2.2.4 Atmospheric deposition

The input of micropollutants via atmospheric deposition (AD) on surfaces (waters, soil, or urban areas) is an essential pathway in many catchments. To quantify this pathway, deposition rates for the substances are necessary.

Three approaches are available for collecting AD:

- Bulk deposition samplers are the simplest and cheapest method to monitor atmospheric deposition. The sampling device is a tray or a bucket. The diameter of the platter or funnel should be chosen depending on the amount of precipitation (~20 – 70 cm). Both settled dust in dry weather and particles bound to precipitation are collected.
- Wet deposition is similar to bulk deposition, but the sampler is equipped with a humidity sensor and a lid, which is closed except during precipitation events (Pekey et al., 2007).
- Wet and dry deposition happen alternatively with the placement of a second bucket: the lid covers either the wet or the dry bucket (Amodio et al., 2014).

Device and sampling strategy

The configuration is similar to devices developed to measure PAH deposition (Foan et al., 2012), where the main elements of the bulk deposition sampler are (a) glass funnel and collection bottle, (b) plate for litter recuperation, (c) collector support and tube for protection against sunlight (examples are shown in Figure 17).

The specific design of the sampler, i.e. the required volume and materials to be utilized, must be defined based on the expected precipitation and on the sampling strategy:

- A. Sampling for 4 months in a year: For this setup, it is required that the samples are collected from the sampler after every storm event during 4 predefined months distributed over one year. As there are two sample bottles, they can be exchanged for sampling. The sample must be immediately transferred into a larger storage container in the freezer. In this option, after four months, after eight months, and at the end of the year, the cumulated bulk deposition collected and frozen in the storage container are thawed and delivered in adequate aliquots to the laboratories for analysis. In this way, three composite samples of atmospheric deposition (each representing four months period) provide information on temporal variability and seasonality of concentrations of HS in atmospheric deposition.

- B. Sampling during a whole year: For this setup, it is required that the samples are collected from the sampler after every storm event for one year. As there are two sample bottles, they can be exchanged for sampling. The sample must be immediately transferred into a larger storage container in the freezer. In this option, each composite sample corresponds to the cumulated deposition collected during each month. The procedure to obtain and to send the samples is the same as indicated for Setup A.

The diameter of the funnel and the volume of the collection bottle has to be calculated based on the expected annual precipitation ranges (delivered from long term statistics) and on the amount of required volume for the chemical analyses.

Material requirement assuming 4 l required sample amount (lab analysis of organic pollutants and metals):

- Funnel (glass, diameter 100 - 300 mm)
- Collection bottle, glass, 2 or 5 or 10 l, 2 pieces
- Storage bottle, glass, 20 l, 2 pieces
- Plug, aluminium foil

It is recommended to set up one AD station in each (sub) catchment. For quality assurance: Duplicate device is to be installed at each sampling location. The reason behind this is that deposition devices might clog, therefore a replacement sample would be available. This would involve the emptying and cleaning of the secondary devices. The sample of the secondary device would only be transported to the composite if the primary sample is lost or not representative.

In DH m³c simple bulk deposition collection was delivered at all locations using large diameter glass or ceramic funnels (Figure 17 and Figure 18) to collect enough water during the sampling periods, which were covering a minimum of one month and was collected at least three times in different seasons. Samples collected during precipitation events were immediately collected after the event and poured into a larger container on site. Following the initial negative experiences with breaking of the samples, the project team decided to change the method of preservation from freezing to cooling. Samples at Wulka and Ybbs catchment were however frozen using specific safety glass containers.



Figure 17: Examples of the applied simple deposition collectors at Viseu (left), Koppány (middle) and Vit (right) catchments.



Figure 18: Atmospheric deposition samplers and transformer at Krenstetten and at Opponitz.

Challenges addressed during the sampling:

- To cover seasonal effects of pollutant distribution and precipitation amounts, a whole year of sampling is deemed necessary. To reduce the associated staff resources, selected months can be sampled to cover to a satisfactory extent the seasonal variations.
- The bigger the sampling devices get, the more complicated the handling and storage become. Thus, the sampled volume should be a compromise aimed to collect the necessary amount without losing part of samples during intense events.

- The samples need to be taken out of the sampling device immediately after the rain events to avoid evaporation and degradation. They need to be stored in a freezer until analysis. The samples should be protected from direct sunlight at all times by wrapping them in aluminium foil.
- The contamination of samples from soil or vegetation must be avoided. Thus, the samplers should be placed in an open space at least 1.5 m above ground.
- Disturbance of the mean spatial rainfall and dry deposition pattern should be avoided by keeping distance to higher objects (buildings, trees...).
- The location of the samplers should ensure their protection from vandalism.

2.2.5 Wastewater

An underlying principle for wastewater effluent sampling is that the greatest possible information concerning household and industry connected to the WWTP is needed. Household effluents are in general relatively homogeneous in time, although the population number can fluctuate seasonally or weekly (due to work and vacation patterns). Industrial discharges are widely diverse and can fluctuate extremely. In general, it is beneficial to do the sampling for at least one week and to repeat it seasonally (Moser et al, 2015).

Two approaches can be applied for the sampling of wastewater:

- Grab samples consist of either a single discrete sample or individual samples collected over a time not exceeding 15 minutes. The grab sample should be representative of the wastewater conditions at the time of sample collection. The sample volume depends on the type and number of analyses to be performed (Simpson, 2017).
- Composite sampling means collecting many samples throughout a longer time into one container and doing the analyses on the mixed liquid. It can be done either time or flow proportionally. A time composite sample consists of equal volume discrete sample aliquots collected at constant time intervals into one container. A time composite sample can be collected either manually or with an automatic sampler. A flow proportional composite sample can be obtained following one of the two following approaches: i) collecting a constant sample volume at varying time intervals proportional to the wastewater flow, ii) collecting the samples by varying the volume of each individual aliquot proportional to the flow, while maintaining a constant time interval between the aliquots (Simpson, 2017).

In addition to effluents of wastewater treatment plants, a subset of untreated influent should also be sampled and analysed. The rationale behind this is that in parts of the Danube River Basin wastewater is not treated yet and therefore, this part of the sampling shall deliver relevant information on how to estimate emission loads into water bodies via untreated municipal wastewater.

Sampling procedure

In general, it is beneficial to do the sampling for at least one week and to repeat it seasonally. Therefore, the influent and effluent of each examined municipal and industrial wastewater treatment plant should be sampled at least 3 times throughout the year at approximately three-four months distance. The aim is to obtain each time a flow-proportional weekly composite sample. During every

campaign, samples have to be collected for seven consecutive days and thereafter merged and homogenized (Figure 19). Depending on the local possibilities, wastewater can be sampled either manually or via autosamplers.

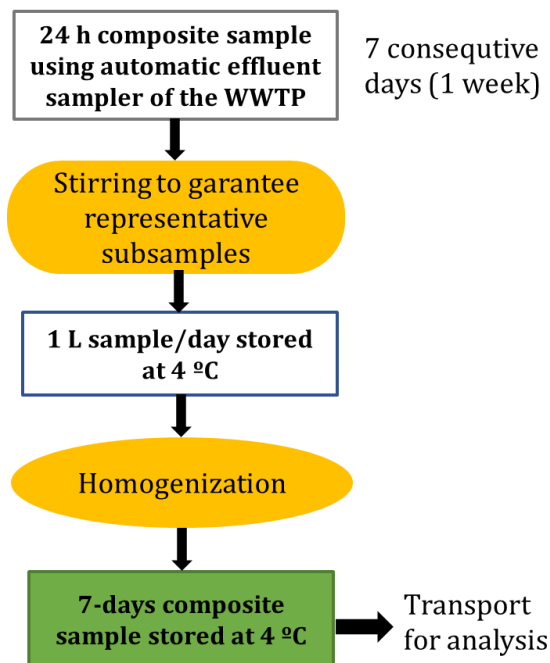


Figure 19: Scheme of sampling procedure for wastewater.

WWTP sampling

The aim is to obtain each time a flow-proportional weekly composite sample

Wastewater discharge must be measured at the time of sampling in order to calculate the load.

Knowledge concerning households and industries connected to the WWTP is needed.

Selection of WWTPs

A different number of plants, both municipal and industrial, shall be examined, depending mostly on the contribution of wastewater plant discharges to total river flow and thus on the relative importance of these point sources for the total transported loads of micropollutants in the different catchments.

In the DH m³c waste water samples were collected at 15 plants (including mining tailings) across the seven pilot regions. Where automated samples were available (Figure 20), samples were collected during one week, each day producing a 24-hour composite sample. Sub-samples were mixed afterwards, proportionally to flow rates. Samples were cooled during collection either by built-in cooling or by passive cooling, using insulated boxes and cooling-packs. Where this was not possible, grab samples were collected manually and then mixed into composite samples.



Figure 20: Sampler of effluent at WWTP A (left) and at WWTP B (middle & right) at Wulka catchment.

2.2.6 Soil

Soils can be a major source of diffuse pollutants in river basins. Particle-bound pollutants are transported via erosion, while dissolved pollutants leaching from soils can reach water bodies by runoff. Soil data (e.g. concentrations of the contaminants by land uses and soil types) are very important inputs to watershed models.

Spatial representativeness of soil sampling

River catchments considered in emission inventories and emission models can be several hundreds of km² in size (or larger) and their soils are very heterogeneous. Therefore, it is difficult to get a spatially explicit sampling from the whole watershed. Nevertheless, it is possible to represent the major soil types and the primary land use types by collecting a high number of samples. To increase the representativeness of soil samples, it is common to use spatially-integrated composite soil samples instead of point samples. In the LUCAS survey (Tóth et al., 2013), a 4x4 m square area was used with five sub-sampling points in an X shape to create a composite sample. Other studies used larger sample grids of 10 m and 9 sample points within (Rocco et al., 2016) and showed that discrete sample concentrations show much higher local concentrations than composite samples. In plot scale studies, it is common to use 1 to 5 ha grids with at least 20 sampling points within (Sarkadi et al., 1986). This study suggests that 20 samples should be the minimum number to create a composite sample with an acceptable error of the real mean concentration.

In the DH m³c project, a cost-effective composite sampling programme was executed in all pilot regions. For the selection of representative points, land use and soil classes should be overlapped by GIS application for each sub-catchment to create land units for composite sampling. The specific sampling points should be distributed randomly within the land units by GIS techniques, and finally, the sample positions should be adjusted manually to position them near the roads where they can be realistically reached. In the project, sampling locations have been selected by random sampling as described above. The presented locations were based on land use data, that is not too accurate, especially at the borders, sometimes outdated, and sometimes biased during the processing of the original remote sensing data. A careful review of each location was thus necessary using up-to-date satellite images. An example is introduced in Figure 21. Even with careful design, the locations were sometimes inaccessible, therefore some of the points necessary to be replaced within the spatial unit.

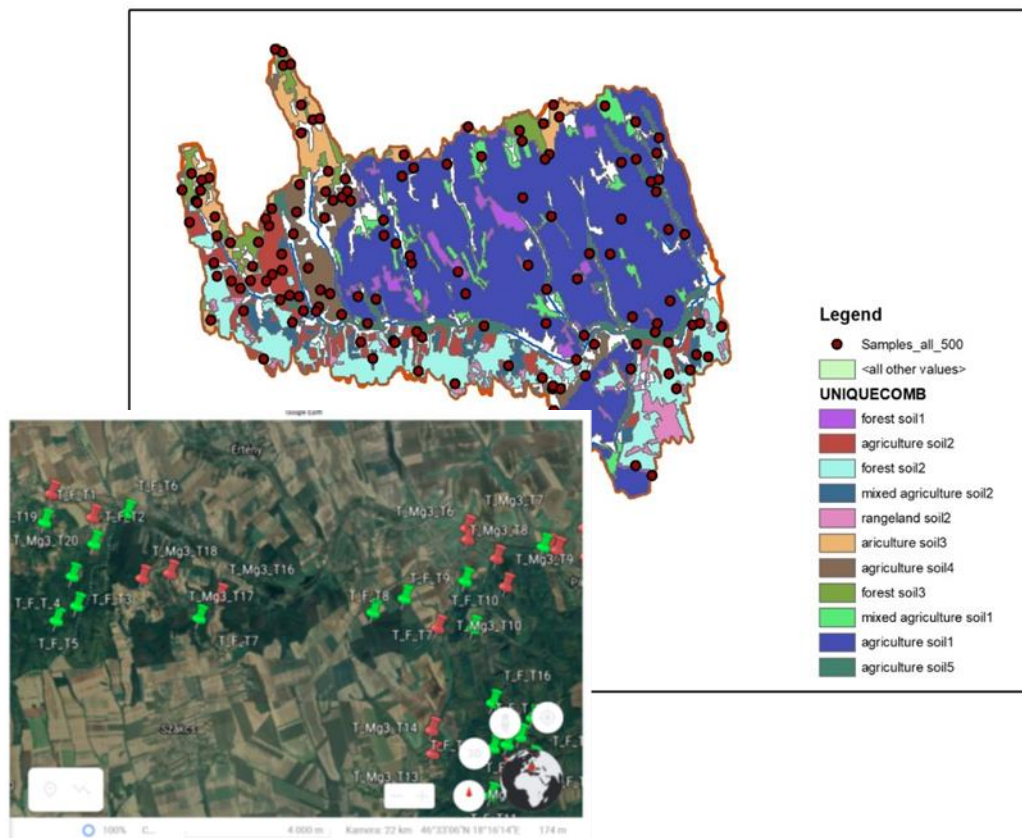


Figure 21: Example of selection of sampling points for the Koppány catchment in Hungary: GIS based planning and fine tuning with satellite images .

Sampling strategy

A reasonable compromise to cover the heterogeneity of soils in the pilot regions was to generate composite samples for each major land-use type, with a total of 10 composite samples in each pilot region. The minimum threshold for the major land uses was set to 5 % of the total area. Each composite sample should be composed of at least 20 sub-samples. Each of the 20 samples is composed in turn of

1-5 subsamples, to be taken close to each other at one location. Soil sampling campaigns should be ideally carried out during spring and summer, when fertilization has occurred but vegetation is not yet dense. This sampling campaign, including several sub-samples to increase the representativity of the composite samples, required a significant effort and was extended over one-two weeks at each pilot region.

Tools and methods

Soil samples can be taken for chemical analysis by simple auger tools, no specific drill is necessary: Pürckhauer ground augers (Figure 22) specifically suitable for dense, hard soils. Edelman augers for softer soils. 1-5 soil sub-samples have to be collected and homogenized by physical defragmentation

The upper soil layer should be sampled because this soil profile is the most important as runoff will play a much bigger part than subsurface flow in the emissions of such chemicals. For grasslands and forest soils, the upper 10 cm shall be sampled, whereas for agricultural soils, the upper 30 cm is relevant due to the tillage mixing of the soil layer. Litter (plant residues) should be removed from the surface prior to sampling. The depth should be similar at each location. Soil samplers should be cleaned between sampling spots. All soil samples are collected in clean and sealable glass jars to prevent contaminant reaction with the container's material. Composite samples are generated by merging equal aliquots of the sub-samples straight after sampling on site.



Figure 22: Pürckhauer type soil sampler used in most catchments.

Lessons learned from the soil sampling campaigns in the pilot regions:

- Soil sampler instrument: In compacted dry soil, the sampling was almost impossible. Rubber hammer was used. In forest and tilled agricultural soils the sampling was easy. Main problems occurred on grasslands and in places where the plantation was already high. In these places the soil was heavily compacted in some cases.

- Auger type soil sampler instrument: Due to the rotation motion for extracting the soil sample, the auger soil sampler met no difficulties even over stiff and compacted land.
- Access to soil sampling points is extremely important. Often geospatial analysis using GIS methods can determine the most accessible locations of these points. However, this analysis does not consider whether the land is fenced or not. Another aspect to be mentioned here is the spatial resolution of the maps used (digital terrain model, road maps, and land use map), a low resolution of the input data can wrongly determine a point with a different land use than the real one. That is why in the maps made for the location of the sampling points the type of land use was also mentioned.
- Sample processing: Pürkhauer sampler with relatively small diameter was used to reduce sample amount. The benefit was that samples could be used in their fully amount and therefore no cross contamination due to mixing and separation had occurred.
- Soil sub-samples were collected in a bucket and homogenized by physical defragmentation. Immediately afterward 100 grams were measured and placed in a glass jar on site. In order for the scale to work properly, a flat and hard surface is needed – the measurements can be done in the trunk of a car.
- Sample processing: soil samples were collected in a ceramic tray and homogenized by physical defragmentation of the samples. Soil was mixed with spoon several times, then adequate amount was measured to the collector glass jar. Very dry samples were hard to defragment.
- Auger type soil sampler instrument: Due to the rotation motion for extracting the soil sample, the auger soil sampler met no difficulties even over stiff and compacted land.

2.3 Sample preparation, storage and transport to the laboratory

As a general rule, the chemical analysis should be carried out as soon as possible after sampling. However, in the case of inventory-supporting monitoring, a non-traditional approach with composite sampling was carried out in order to maximise the information obtained from the measurements at the lowest possible analytical cost.

New methodology for the preparation of composite samples had to be developed to find the best way how to preserve samples for such a long time. Sample preservation was initially determined by the consortium for all sampled matrix and for all samples. The methods were described in the Standard Operations Procedures (SOP) document that was revised a few times during the project when alterations of the methods were needed. Table 5 summarizes the applied preservation in case of the different substances and matrices.

2.3.1 Preservation of samples

River water samples

Samples shall be kept in cold conditions during transport using cooling packs and insulation boxes. Filtration and acidification have to be carried out in the laboratory right after sample transportation (within 4 hours after collection). Preparation: in the case of low-flow composites, the sample have to be stored in a vessel (material and cooling temperature is depending on the analysed substance). Every

week a weekly aliquot + a proportional amount of stabilizing compound is added to the vessel. Storage takes place in cold (<4°C) or frozen conditions. Samples are transported to the analysing laboratory at the end of the collection period. From high-flow events, a flow-proportional composite should be prepared (sampled by flow triggered programming auto samplers or manually from the sample series. After necessary filtering and stabilization (where needed), high-flow samples shall be immediately transported to the analysing laboratories.

Wastewater samples

Have to be collected in cooled instruments. Cooling can be either made by active or passive cooling (insulation boxes and cooling packs). Wastewater samples have to be collected and transported to the lab daily or bi-daily. Samples have to be stored in cold (<4°C) and dark conditions until the total weekly samples are collected. Preservation and transport of the composite samples takes place as described below in the table.

Atmospheric deposition

To preserve the samples during months of collection, it was decided in the consortium that freezing would be the selected procedure for storage, but this required to use large volume (10l) glass bottles to avoid sample loss to plastic. The initial negative experiences of glass breaks forced to change the preservation either to safety glass containers (expensive) or to cooling in dark conditions (<4 °C) instead of freezing.

Soils and sediments

Following sampling, the samples were kept in dark, cold conditions until delivery to the lab, where samples have to be homogenized and lyophilized to preserve contaminants.

Table 5: Preservation and storage information for the liquid phase samples

Substance	Sample matrix	Volume and bottle material	Preservation
Hg and other metals (total and dissolved*)	All matrices	0.5 l Teflon or PE 0.16 L of filtered water	0.16 mL of HCl s.p. (30%) or 0.16 mL of HNO ₃ s.p. (65%) Frozen
16 PAH total and dissolved**	River, Atm.Dep. Wastewater	1 L Amber glass	The inner surface of plastic cups covered with aluminium foil, Cooling (2-4 °C)
16 PAH	SPM, Soil	1 kg, Rex glass	Lyophilisation
PFAs (PFOS, PFOA, PF4C – PF12C, PF4S – PF10S)	All matrices	250 ml, PE	Frozen for composites, otherwise cooling: max 6 days
4-ter Octylphenol, Nonylphenol		1 L dark glass	Cooling (2-4 °C) and max 2 months

Substance	Sample matrix	Volume and bottle material	Preservation
Metolachlor (incl. Metabolites), Tebuconazole		2x40 ml EPA vial	Cooling (2-4 °C) max 2 months
Diclofenac, Carbamazepine, Bisphenol A		1 L dark glass	Cooling (2-4 °C) max 2 months

*Filtering Metals: Due to the carbonate precipitation the dissolved samples should be filtered and acidified immediately after sampling. Only specified pure filters and acids should be used to avoid contamination

**Filtering PASSES: specified glass fibre filters.

2.3.2 Procedures for cleaning of bottles and vessels

Cleaning procedure for glass bottles:

Reagents: Nitric acid HNO₃ (65% pro analysis - p.a.), Ultrapure water e.g. (Milli-Q) (> 18 M Ω cm), or HPLC reagent grade water

- 1) rinse the glass bottles thoroughly with tap water;
- 2) soak glass bottles with 10% HNO₃ (65%) (v/v) for 24 hours and afterward rinse with ultrapure (Milli-Q) water, fill the bottle with 10% HNO₃ and leave for 24 hours;
- 3) rinse the bottles thoroughly with ultrapure (Milli-Q) water (5 times);

Cleaning the PE and teflon bottles

- 1) put the gloves on
- 2) take the bottles out of the zip-locked bags
- 3) empty the Teflon bottles, which are filled with diluted acid
- 4) rinse them thoroughly with the water sample (3 times)
- 5) fill the bottle till the top with the sample (This is only applicable if only one sample is taken and no composite is collected) and acidify them immediately with 1 mL of HCl s.p. (37%) or 1 mL of HNO₃ s.p. (65%) per 1 L of water, so that final concentration is 1% v/v)
- 6) put the sampler to the freezer.
- 7) put the bottles in PE zip-locked bags and store them for further analysis (if the sampling is taking place during the summertime, the samples should not be stored at room temperature) (This is only applicable if only one sample is taken and no composite is collected).

For quality control reasons blank samples (procedure blanks, autosamplers) have to be prepared.

2.3.3 Transport of the samples to the lab

As soon as the composite samples are ready, they are sent to the respective laboratories via express courier transport in boxes of polystyrene foam filled with dry ice, or ice batteries.

The receiving laboratories should measure sample temperatures at delivery. In case of temperatures above 8°C, the laboratory should contact the sender, who should reconsider the shipment procedure.

In the laboratory, careful thawing of the samples should be performed at controlled room temperature (4°C) or at room temperature, in order to prevent losses of volatile PAH (e.g. Naphthalene). This procedure only applies to the atmospheric deposition samples (and maybe wastewater), which are stored frozen and where the distribution to the smaller bottles is arranged at the end of the collection period.

Freezing has a significant impact on the measurements of dissolved metals: white precipitation was observed when thawing of the sample. After chemical parameters showed a strong change after freezing the samples as composites: obvious signs of CaCO₃ precipitation. This is proven indirectly, by increased pH, reduced Electric Conductivity, several fold increase in turbidity and around 40-60% loss of hydrogen carbonate, and strong loss of calcium in water samples. There is a risk of loss of particle bound contaminants from the samples, high risk of measurement errors. Therefore, samples have to be filtered on-site right after sampling delivery to the own lab, using specific pure filter: Sartorius Ministart NML, syringe filter, 28 mm, 0.45 µm pore size.

2.3 Laboratory methods

It is recommended that chemical analysis is performed by the same lab for the same parameters for all samples. Methods and LOD/LOQ concentrations applied in the DH m³c project are summarised in Table 6 for all liquid matrices and Table 7 for soil and SPM.

Table 6: Analytical methods, LOQ and LOD values for liquid matrices.

CAS Substance name	Unit	Method	Standard	LOQ	LOD
Potential toxic elements (PTEs) in AD and RIV_dissolved					
CAS_7440-43-9_Cadmium and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,013	0,004
CAS_7439-92-1_Lead and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,05	0,015
CAS_7440-02-0_Nickel and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,02	0,006
CAS_7440-38-2_Arsenic and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,033	0,01
CAS_7440-50-8_Copper and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,043	0,013
CAS_7440-66-6_Zinc and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,233	0,07

CAS Substance name	Unit	Method	Standard	LOQ	LOD
CAS_7440-47-3_Chromium and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,02	0,006
CAS_7439-97-6_Mercury and its compounds	µg/l	CVAAS	ISO 17294-2:2016	0,00027	0,00008
Potential toxic elements (PTEs) in WW and RIV_total					
CAS_7440-43-9_Cadmium and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,133	0,04
CAS_7439-92-1_Lead and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,5	0,15
CAS_7440-02-0_Nickel and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,2	0,06
CAS_7440-38-2_Arsenic and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,333	0,1
CAS_7440-50-8_Copper and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,433	0,13
CAS_7440-66-6_Zinc and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	2,33	0,7
CAS_7440-47-3_Chromium and its compounds	µg/l	ICP-MS	ISO 17294-2:2016	0,2	0,06
CAS_7439-97-6_Mercury and its compounds	µg/l	CVAAS	ISO 17294-2:2016	0,0027	0,0008
Organic substances in RIV and AD					
CAS_1763-23-1_Perfluorooctane sulfonic acid (PFOS) and its derivatives	µg/l	HPLC-MS/MS	WBSE-121:2019	0,00015	0,00004
CAS_335-67-1_PFOA	µg/l	HPLC-MS/MS	WBSE-121:2019	0,00015	0,00004
CAS_2706-90-3_Perfluoropentanoic acid	µg/l	HPLC-MS/MS	WBSE-121:2019	0,002	0,0006
CAS_307-24-4_Perfluorohexanoic acid (PFHxA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,002	0,0006
CAS_375-85-9_Perfluoroheptanoic acid (PFHpA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,002	0,0006
CAS_375-95-1_Perfluorononanoic acid (PFNA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,002	0,0006
CAS_335-76-2_Perfluorodecanoic acid (PFDA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,002	0,0006
CAS_2058-94-8_Perfluoroundecanoic acid (PFUnA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,002	0,0006
CAS_307-55-1_Perfluorododecanoic acid (PFDoA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,002	0,0006
CAS_72629-94-8_Perfluorotridecanoic acid (PFTrA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,002	0,0006

CAS Substance name	Unit	Method	Standard	LOQ	LOD
CAS_376-06-7_Perfluorotetradecanoic acid (PFTeA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,002	0,0006
CAS_375-73-5_Perfluorobutane sulfonic acid (PFBS)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,002	0,0006
CAS_355-46-4_Perfluorohexane sulfonic acid	µg/l	HPLC-MS/MS	WBSE-121:2019	0,002	0,0006
CAS_15307-86-5_Diclofenac	µg/l	HPLC-MS/MS	WBSE-124:2019	0,001	0,0003
CAS_298-46-4_Carbamazepin	µg/l	HPLC-MS/MS	WBSE-124:2019	0,001	0,0003
CAS_140-66-9-Octylphenol (4-(1,1',3,3'-tetramethylbutyl)- phenol)	µg/l	GC-MS	EN ISO 18857-1:2007	0,01	0,005
CAS_104-40-5_4-nonylphenol	µg/l	GC-MS	EN ISO 18857-1:2007	0,01	0,003
CAS_80-05-7_Bisphenol A	µg/l	HPLC-MS/MS	WBSE-124:2019	0,01	0,003
CAS_107534-96-3_Tebuconazole	µg/l	HPLC-MS/MS	WBSE-93:2020	0,01	0,003
CAS_51218-45-2_Metolachlor	µg/l	HPLC-MS/MS	WBSE-93:2020	0,01	0,003
CAS_171118-09-5_Metolachlor ESA	µg/l	HPLC-MS/MS	WBSE-93:2020	0,01	0,003
CAS_152019-73-3_Metolachlor OA	µg/l	HPLC-MS/MS	WBSE-93:2020	0,01	0,003
Organic substances in WW					
CAS_1763-23-1_Perfluorooctane sulfonic acid (PFOS) and its derivatives	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005
CAS_335-67-1_PFOA	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005
CAS_2706-90-3_Perfluoropentanoic acid	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005
CAS_307-24-4_Perfluorohexanoic acid (PFHxA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005
CAS_375-85-9_Perfluoroheptanoic acid (PFHpA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005
CAS_375-95-1_Perfluorononanoic acid (PFNA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005
CAS_335-76-2_Perfluorodecanoic acid (PFDA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005
CAS_2058-94-8_Perfluoroundecanoic acid (PFUnA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005

CAS Substance name	Unit	Method	Standard	LOQ	LOD
CAS_307-55-1_Perfluorododecanoic acid (PFDoA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005
CAS_72629-94-8_Perfluorotridecanoic acid (PFTrA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005
CAS_376-06-7_Perfluorotetradecanoic acid (PFTeA)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005
CAS_375-73-5_Perfluorobutane sulfonic acid (PFBS)	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005
CAS_355-46-4_Perfluorohexane sulfonic acid	µg/l	HPLC-MS/MS	WBSE-121:2019	0,0016	0,0005
CAS_15307-86-5_Diclofenac	µg/l	HPLC-MS/MS	WBSE-124:2019	0,0016	0,0005
CAS_298-46-4_Carbamazepin	µg/l	HPLC-MS/MS	WBSE-124:2019	0,0016	0,0005
CAS_140-66-9-Octylphenol (4-(1,1',3,3'-tetramethylbutyl)- phenol)	µg/l	GC-MS	EN ISO 18857-1:2007	0,0016	0,0005
CAS_104-40-5_4-nonylphenol	µg/l	GC-MS	EN ISO 18857-1:2007	0,0015	0,0004
CAS_80-05-7_Bisphenol A	µg/l	HPLC-MS/MS	WBSE-124:2019	0,0015	0,0004
CAS_107534-96-3_Tebuconazole	µg/l	HPLC-MS/MS	WBSE-93:2020	0,02	0,006
CAS_51218-45-2_Metolachlor	µg/l	HPLC-MS/MS	WBSE-93:2020	0,02	0,006
CAS_171118-09-5_Metolachlor ESA	µg/l	HPLC-MS/MS	WBSE-93:2020	0,02	0,006
CAS_152019-73-3_Metolachlor OA	µg/l	HPLC-MS/MS	WBSE-93:2020	0,02	0,006

Table 7: LOQ and LOD values for soil and SPM.

CAS Substance name	Unit	LOQ	LOD
Potential toxic elements (PTEs) and rare elements			
CAS_7429-90-5_Aluminium and its compounds	mg kg ⁻¹ DM	99,9	30
CAS_7723-14-0_Total phosphorus	mg kg ⁻¹ DM	72,594	21,8
CAS_7440-47-3_Chromium and its compounds	mg kg ⁻¹ DM	0,41625	0,125
CAS_7439-89-6_Iron and its compounds	mg kg ⁻¹ DM	41,625	12,5
CAS_7440-02-0_Nickel and its compounds	mg kg ⁻¹ DM	1,24875	0,375
CAS_7440-50-8_Copper and its compounds	mg kg ⁻¹ DM	0,624375	0,1875

CAS Substance name	Unit	LOQ	LOD
CAS_7440-66-6_Zinc and its compounds	mg kg ⁻¹ DM	15,8175	4,75
CAS_7440-38-2_Arsenic and its compounds	mg kg ⁻¹ DM	0,8325	0,25
CAS_7440-43-9_Cadmium and its compounds	mg kg ⁻¹ DM	0,208125	0,0625
CAS_7439-91-0_Lanthanum	mg kg ⁻¹ DM	0,041625	0,0125
CAS_7440-45-1_Cerium	mg kg ⁻¹ DM	0,041625	0,0125
CAS_7439-92-1_Lead and its compounds	mg kg ⁻¹ DM	0,208125	0,0625
Organic substances			
CAS_129-00-0_Pyrene	µg/kg DM	1,8	0,88
CAS_85-01-8_Phenanthrene	µg/kg DM	5,6	2,8
CAS_91-20-3_Naphthalene	µg/kg DM	8	4
CAS_193-39-5_Indeno(1,2,3-cd)pyrene	µg/kg DM	0,36	0,1
CAS_86-73-7_Fluorene	µg/kg DM	1,1	0,53
CAS_206-44-0_Fluoranthene	µg/kg DM	2,2	1,1
CAS_53-70-3_Dibenzo(a,h)anthracene	µg/kg DM	0,42	0,12
CAS_218-01-9_Chrysene	µg/kg DM	0,66	0,19
CAS_207-08-9_Benzo(k)fluoranthene	µg/kg DM	0,62	0,2
CAS_191-24-2_Benzo(g,h,i)perylene	µg/kg DM	0,66	0,2
CAS_205-99-2_Benzo(b)fluoranthene	µg/kg DM	0,83	0,25
CAS_50-32-8_Benzo(a)pyrene	µg/kg DM	0,48	0,12
CAS_56-55-3_Benzo(a)anthracene	µg/kg DM	0,75	0,23
CAS_120-12-7_Anthracene	µg/kg DM	0,73	0,37
CAS_208-96-8_Acenaphthylene	µg/kg DM	0,43	0,11
CAS_83-32-9_Acenaphthene	µg/kg DM	0,45	0,13

3 Inventory of hazardous substances - relational and harmonized database

3.1 Inventory of HS emissions

3.1.1 *Setting up the inventory – scope and objectives*

The setup of inventories of emissions of hazardous substances into surface waters is a heavily data driven process. The most important part consists of substance-specific data, which are necessary to quantify the emission loads. This can be concentrations in different environmental or technical matrices (e.g. wastewater or soil), surface specific rates (e.g. atmospheric deposition rates), emissions factors for different activities (emissions per vehicle and driven kilometre) or emission loads (e.g. industrial emission reported to PRTR register). Based on such data, different types of emissions and transport models can be used to quantify the overall emissions. Such models need to be validated, which again needs measured concentration data from surface waters. To collect these data together with all necessary metadata, an inventory of concentrations was created in the Danube Hazard m³c project (DHm³c). This inventory serves as example and guidance for anyone who aims to set up a similar data collection system on a national or transnational scale. Thus, in the following, we present the structure of the database, the method of data collection in detail as well as some important lessons learned.

In the DHm³c project, the aim of this activity was to collect pre-existing data from the Danube basin which were available from different sources and in different formats and combine them together with the monitoring data generated within DHm³c to generate a database as broad as possible for three applications:

1. Generate input data for emission models, especially for the MoRE model (Fuchs et al., 2017) which is mainly based on representative concentrations in different pathways.
2. Supply validation data for emission models such as the source driven Danube Hazard Substance Model (DHSM) (van Gils et al. (2020) and DHm³c Output *Upgraded version of the SOLUTIONS model (now called DHSM) adapted to territorial needs for transnational modelling of HS emissions in the DRB* and the pathway driven model MoRE.
3. Supply harmonized data for surface water status assessment.
4. Create a database for research of drivers behind the concentration patterns, which can lead to a better system understanding and thus to better performing emission models.

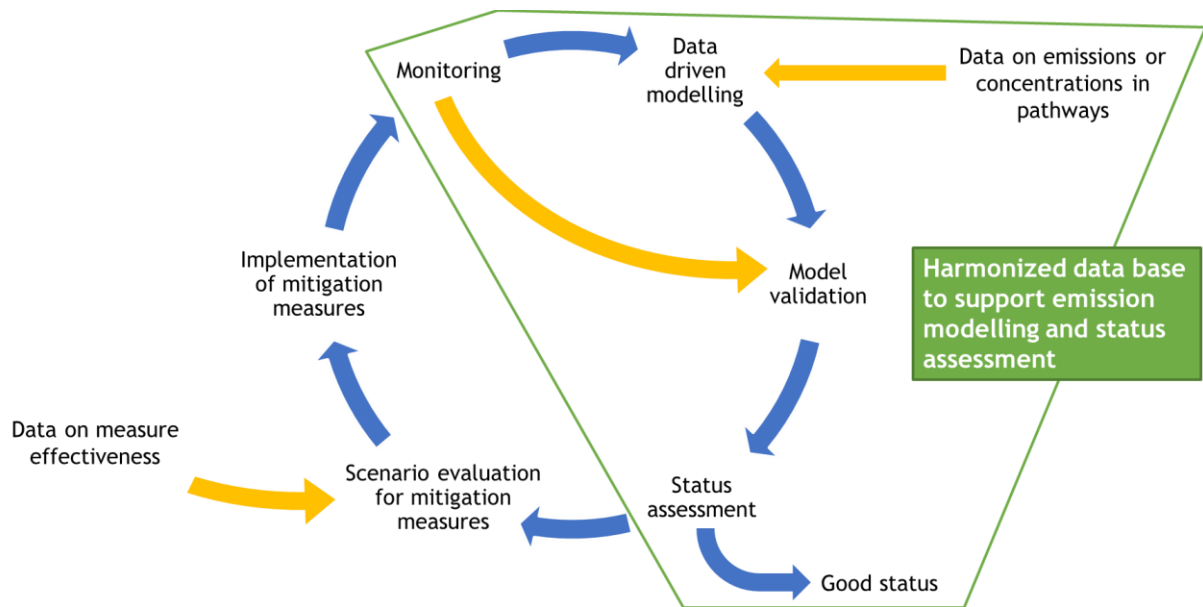


Figure 23: The cycle of hazardous substance management at the river basin scale (blue), the role of data in this cycle (yellow) and which part the database presented here should cover (green) (Kittlaus 2023, CC-BY-SA 4.0)

3.1.2 Considered ideas and requirements

As the DHm³c inventory database had a strong focus on capitalization of pre-existing knowledge, available data sources were investigated regarding their structure and the use of controlled vocabularies. These are the following databases with data available for the Danube Basin:

- Hosted by the ICPDR¹:
 - Transnational Monitoring Network (TMNM)
 - Joint Danube Survey Results (JDS1, JDS2, JDS3)
- Hosted by the NORMAN EMPODAT database²:
 - Joint Danube Survey 4 Results
- Hosted by the EU
 - Reported monitoring data in the WISE framework
 - Reported emissions under the UWWT directive
 - Reported emission under the PRTR directive

Regarding the selection of the technical framework to collect the data, it is essential to have a system which meets the following requirements:

- Collection of data in a format, which is manageable for all contributing institutions.
- Possibility of rigorous quality control of the collected data and metadata during data import.
- Easy handling of large data amounts without any restrictions.

¹ <https://www.icpdr.org/wq-db/>

² <https://www.norman-network.com/nds/empodat/>

- Data workflow to be handled with the available skills of the project team.

In the project, to make data available in the best possible way the principles of FAIR data (Findability, Accessibility, Interoperability, and Reuse) were considered as far as possible:

The DHm³c database has been made available as one main output of the project for further use in the research data repository of TU Wien, accessible via <https://doi.org/10.48436/xwve4-h7v43>.

3.2 Implementation of the inventory

3.2.1 Data collection and import

For a transnational database, the data collection has to be shared between project partners, as national administrations have a better overview and access to national data and e.g. universities better access to scientific publications. Nevertheless, many data sources, especially those which are not directly water related (e.g. soil data) were difficult or impossible for the project team to access. While surface water quality related institutions were included as partners in the project and surface water quality data were available in their department, data regarding other compartments (soil, atmospheric deposition, wastewater) are partly in the responsibility of other institutions or other departments of the same institutions and therefore have been difficult or even impossible to access. Data collection and data harmonization are of prime importance for a well-structured database. Therefore, this document presents, as a good example, the workflow of how the data was collected, checked and uploaded to the database during our project (Figure 24).

1. To facilitate a smooth data collection with the available technical knowledge and setup, it was decided to create data request sheets (DRS) in Microsoft-Excel format, distribute them to the partners and collect the data in these templates in an online storage (cloud) system. Examples of how such DRS were conceived and structured are available as excel files in the Annex “Technical Guidance – Annex – Data request sheets” published on the project homepage³.
2. The filled DRS were collected in the cloud of the TU Wien and then centrally processed by few project members using mainly R programming language (R Core Team, 2022). Questions occurring during data checking were send back to the data collecting partners to clarify the situation and where necessary to update the submitted Excel-files.
3. The supplied data in the format of the DRS was rather easy to import because the data structure was given and fitted to the SQL database structure. However, for various reasons, not all data were made available in the DRS format. The quality of these varied, from easy to process subsets of databases to unformatted .xlsx-Files which took a lot of effort to encrypt. The need for a unified public data collection was clearly evident.
4. The data import was performed with the programming language R (R Core Team, 2022) in reproduceable, easy to verify scripts. Because data import is not always a single task but must sometimes be repeated, e.g. if the data have been imported incorrectly or new metadata become available. It is important that the import-scripts are well structured and documented

³ The DRS included in the Annex do not reflect the last status of the data base structure finally generated in the project, but serve as example of structure, scope and level of complexity.

and follow the guidelines for clean code. Thus, subsequent verification is traceable and later changes are easily possible.

- For the transformation of the raw data into a clean data file a script for each compartment has been prepared. There, the raw data have been processed and cleaned so that they met the criteria of the database. The actual import into the inventory database was performed with an extra script.

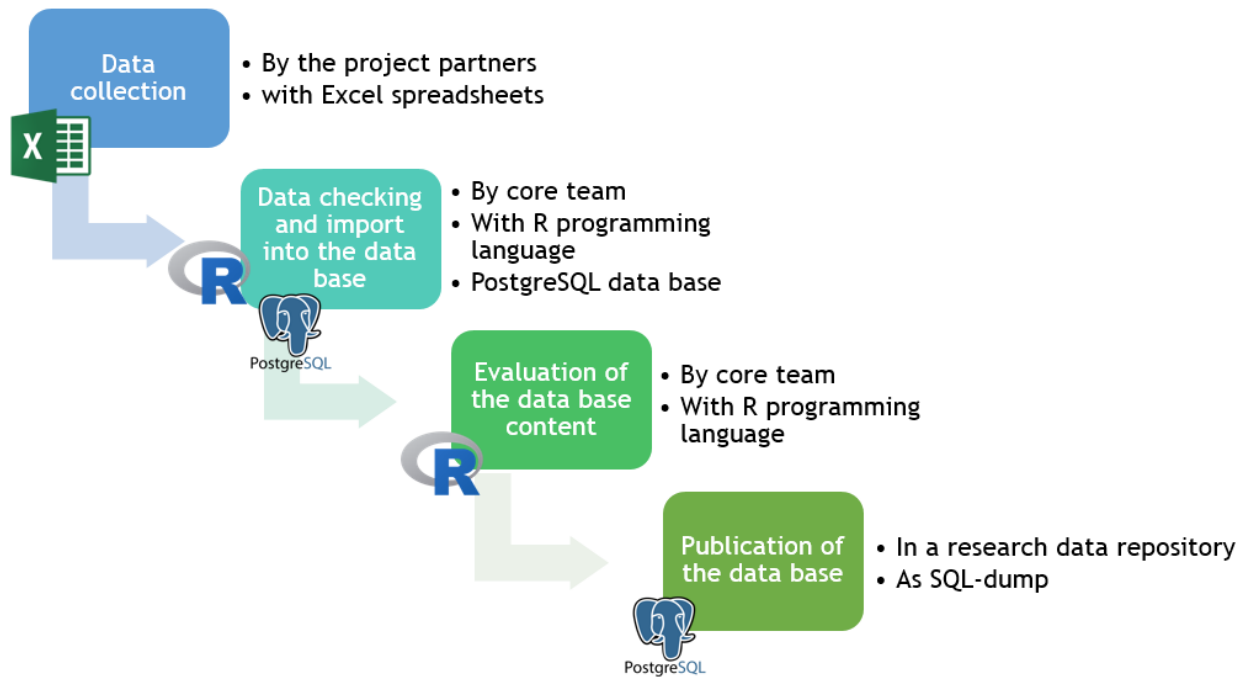


Figure 24: Workflow and software for data collection, checking, evaluation and publication chosen in the Danube Hazard m³c project (Kittlaus 2023, CC-BY-SA).

From the project experience it can be concluded that it is better if only few people are responsible in a centralized way with the whole process of data import, as the understanding of different meta data and their associated controlled vocabularies varies between different persons and a harmonization of this understanding is necessary to finally achieve a harmonized database and to enable an automated and reliable evaluation. This requires either the availability of a detailed documentation or, especially during the development stage of such a data base structure, an intense collaboration of all involved team members.

3.2.2 Data storage in a well-structured SQL database

For storage of the data, it is important to choose a robust, efficient tool which is well-known by the persons importing and centrally managing the data, which allow handling large amounts of data and which is open-source. In the DHm³c project, the relational database was implemented in PostgreSQL.

(Version 9.6)⁴, as this tool was already available, known by some team members and is a powerful opensource data management system which can handle huge amount of data.

The relational database consists of several tables, which are connected by columns, that refer from one table to another. There are main tables which contain the actual data and supporting tables, which in turn list and describe the allowed entries for columns with controlled vocabularies. The use of controlled vocabularies is of outmost importance to harmonize the metadata from different data sources and make the data evaluable as a whole.

Separate tables were created for the following five environmental and engineered matrices:

1. Water bodies:
 - a. surface water including also suspended particulate matter
 - b. groundwater
2. Wastewater treatment plants: municipal and industrial, influent and effluent
3. Stormwater: combined sewer overflows and stormwater outlets in separate sewer systems
4. Atmospheric deposition
5. Soil

In order for the inventory to be able to include both original monitoring data wherever possible, but also data published in an aggregated form, e.g. in scientific publications and other reports, each section was designed to contain tables for single measurements as well as tables for aggregated measurements.

Further tables are needed to contain metadata about the hazardous substances (names, identifiers) and the data sources (license, data owner), which are referenced in every data set in the main data tables (concentrations). The most important tables and columns will be shortly described in the following sections. The full documentation of the DHm³c database tables and columns can be found in the technical documentation which comes along with the database itself.

Tables for general metadata

Determinants

To make the data interoperable and specially to make hazardous substances easily identifiable, several identifiers and - in some cases - several names per substance should be collected. The following substance identifiers and names are particularly relevant and useful:

- CAS registry no. and EEA coded no. (for compatibility with the WISE database)
- EC number used by the European Chemicals Agency (ECHA⁵)
- Norman Susdat ID (used in the NORMAN-network databases)
- name of the substance mainly taken from ECHA website
- abbreviation of the name for presentation in figures

⁴ <https://www.postgresql.org> Current Version of PostgreSQL is 15.2. The delivered database is compatible with this version.

⁵ <https://echa.europa.eu/de/information-on-chemicals>

- further names like IUPAC name and product names

Among all these names and identifiers, CAS registry number and EEA coded number should be used as the primary identifiers (where available), because these are also the primary identifiers of the EU harmonized WISE database.

To support interpretation of HS data, further water quality standard parameters were collected as determinants, e.g. water temperature, pH, suspended solid concentration.

One table classifies the determinants according to different categories, e.g. chemical groups like PAH or PFAS or application area like biocides or pharmaceuticals.

The available environmental quality standards are listed in a further table to facilitate status assessment.

Moreover, relations between determinants can be described in a further table: e.g. whether one determinant is the degradation product of another, substances are chemically similar, or if determinants are mainly used together.

Data sources

To facilitate reuse of the data for all data sets, the source of the data has to be indicated and for each data source details should be given in a dedicated table:

- When the data set was published and where.
- What type of data set it is (e.g. national or regional database, scientific publication)
- Who is the data owner (organisation) and who is the contact person.
- How should the data set be cited if it is reused.
- Which license applies to the data for reuse.

Precipitation data

Data on precipitation amounts are relevant for interpretation of atmospheric deposition measurements and for storm water runoff from impervious surfaces. Therefore, the inventory contains a section to store metadata and data from precipitation gauges.

Single concentration measurements (for all compartments)

Disaggregated concentration measurements are preferable over aggregated data, as more information about data variability and correlation with sampling conditions can be evaluated from single data. All tables for single measurements should contain the following information for each data set:

- identifier of the sample
- identifier of the determinant

- observed value with unit of measure, analytical limit of quantitation (LOQ) and limit of detection (LOD), and the indication of whether the measurement value is below these thresholds or not
- analysed matrix: total (whole sample), dissolved (filtered sample), solid
- analysing laboratory, applied analytical method and norm describing the method
- information if the laboratory is accredited for that analysis method
- identifier of lab analysis as given by the analysing lab
- identifier of the data source
- comments with additional relevant information related to the measurements.

Aggregated concentration measurements (for all compartments)

If single, disaggregated data cannot be made available for the inventory, then two kinds of time-integrated data can still be relevant and informative for the relational database:

- concentration data from grab samples which were analysed separately and the results were thereafter statistically aggregated over a certain time period (usually one year)
- concentration data from single composite samples, which are time-integrative during sampling.

The following information should be collected regarding the concentration measurements:

- identifier of the determinant
- number of aggregated values, number of values below LOQ and LOD
- unit of measure
- highest LOQ and LOD relevant for the aggregated measurements or for the composite sample
- statistical descriptors: minimum, median, mean, maximum
- for each statistical descriptor the information of whether the value is below LOQ
- standard deviation of the aggregated values
- analysed matrix: total (whole sample), dissolved (filtered sample), solid
- analysing laboratory, applied analytical method and norm describing the method
- information if the laboratory is accredited for that analysis method
- comments with additional relevant information related to the measurements.

The information about the samples in these tables are different depending on the sampled compartment, therefore they will be presented in the following separate sections.

Surface and groundwater

This subset of tables should include measurements for surface and groundwater including suspended particulate matter in surface waters, together with information on the monitoring sites.

For the single samples the following information should be collected:

- sample identifier, ideally using the original identifier provided in the source from which the data come from

- International Generic Sample Identifier (IGSN⁶), a persistent identifier for samples
- identifier of the monitoring site where the sample was collected
- information of which kind of water body was sampled, surface water or groundwater
- time of sampling with time zone
- sampling method, e.g. grab sample, large volume sample, and information of whether the method was accredited or not
- river flow at time of sampling in m³/s
- for groundwater: water level below surface in the well during groundwater sampling, negative values for artesian aquifers
- identifier of the data source
- comments with additional relevant information related to the sampling procedure.

For aggregated measurements the information regarding samples should be the following:

- identifier of the monitoring site
- information which kind of water body was sampled, surface water or groundwater
- begin and end of sampling with time zone
- the sampling method, e.g. grab sample, large volume online solid phase extraction, SPM sampling or large volume sampler and information of whether the method was accredited
- mean river flow, TSS concentration and electrical conductivity during sampling, calculated as the mean of the parameter for each sub sample in the aggregated value or composite sample
- for single composite samples the sample identifier should be given
- for single composite samples the IGSN can be given
- identifier of the data source
- comments with additional relevant information related to the sampling procedure.

For the monitoring site, the following information shall be collected:

- identifier of the monitoring site as used in the original data source and based on which schema the identifier was build
- country in which the monitoring site is located, coded by ISO 3166-1 alpha-2
- name of the sampling site
- coordinates including the coordinate reference system coded by the European Petroleum Survey Group (EPSG) code, e. g 4326 for WGS84 -CRS
- identifier of a correlated monitoring site, e.g. nearby river discharge gauge correlated to a quality monitoring station
- area of the monitoring sites upstream catchment
- long-term discharge characteristics of the monitoring site: Mean discharge (MQ), mean base-flow discharge, 10% percentile of discharge
- long-term annual flood discharge with a statistical recurrence rate of one year, one decade and one century
- long-term mean concentration of total suspended solids

⁶ <https://www.igsn.org>

- identifier of the water body the monitoring site is placed in
- ground water monitoring well: depth of water extraction
- ground water monitoring well: long-term water level depth
- ground water monitoring well: main land use surrounding the well and potentially influencing the water quality. The controlled vocabulary from the CORINE land cover classification (Kosztra et al., 2019) is applied
- identifier of the data source
- comments with additional relevant information related to the monitoring site.

A further table gives the characteristics of the water bodies:

- identifier of the water body and scheme from which the identifier was taken
- type of water body, with a classification system based on catchment size and topography for rivers, and type of hydrogeology for ground water
- name of the waterbody
- country in which the water body is situated, coded by ISO 3166-1 alpha-2
- comments with additional relevant information related to the water body.

Wastewater treatment plants (WWTP)

In the following tables, concentration data and metadata from wastewater treatment plants and other wastewater systems are collected. Data can derive from different sampling points in wastewater systems and treatment plants (inflow, outflow, sludge) and from different types of WWTPs (municipal or industrial).

For the single samples, the following information should be collected:

- sample identifier, ideally using the original identifier provided in the source from which the data come from
- International Generic Sample Identifier (IGSN⁷), a persistent identifier for samples
- identifier of the wastewater treatment plant discharge point where the sample was collected. For inflow samples the attribution to the discharge point is not perfectly correct, but this makes it possible to work with the same structure for inflow and outflow data and no further problems should be experienced
- place of the sampling in the WWTP system e.g. inflow, outflow, primary sludge, excess sludge
- time of sampling with time zone
- sampling method, e.g. grab sample, large volume sample, and information of whether the method was accredited
- flow volume of the sampled stream (inflow, outflow) at time of sampling in m³/s
- identifier of the data source
- comments with additional relevant information related to the sampling procedure.

⁷ <https://www.igsn.org>

For aggregated samples the information regarding samples should be the following:

- identifier of the waste water treatment plant discharge point where the sample was collected
- place of the sampling in the WWTP system e.g. inflow (applies also for sampling from sewer outlets in absence of a WWTP), outflow or primary sludge
- begin and end of sampling with time zone
- the sampling method, e.g. grab sample or composite sample and information of whether the method was accredited
- for single composite samples the sample identifier should be given
- for single composite samples the IGSN can be given
- identifier of the data source
- comments with additional relevant information related to the sampling procedure.

For the wastewater treatment plants or other wastewater sampling sites the following information should be collected:

- identifier of the treatment plant or monitoring site as used in the original data source and based on which schema the identifier was build
- country in which the monitoring site is located, coded by ISO 3166-1 alpha-2
- name of the sampling site
- type of the wastewater treatment plant e.g. municipal, industrial, mixed or other
- design capacity of the WWTP in population equivalent (PE) and classification of the capacity in intervals of PE: [0, 2 000, 5 000, 10 000, 100 000, Inf]
- real connected PE and inhabitants
- share of area drained by combined sewer system in the WWTP catchment
- information about the treatment technologies applied at the plant. The system of classification was based on the UWWT-Directive reporting system⁸
- identifier of the data source
- comments with additional relevant information related to the WWTP.

Since a WWTP could have more than one discharge point, a further table gives the characteristics of this aspect of high importance for emission modelling:

- identifier of the discharge point and the source of the discharge point identifier.
- identifier of the WWTP, to which the discharge point belongs
- name of the discharge point
- country in which the discharge point is located, coded by ISO 3166-1 alpha-2
- coordinates including the coordinate reference system coded by the European Petroleum Survey Group (EPSG) code, e. g 4326 for WGS84 -CRS
- NUTS level 3 code to locate the discharge point in cases where coordinates are not available. Nomenclature of Territorial Units for Statistics or NUTS (French: "Nomenclature des unités

⁸ <http://dd.eionet.europa.eu/datasets/latest/UWWTDArt15/tables/UWWTPs>

territoriales statistiques”) is a geocode standard for referencing the administrative divisions of countries for statistical purposes. The standard, adopted in 2003, is developed and regulated by the European Union (European Commission, 2022)

- type of recipient of the WWTP effluent e.g. surface water, ground water, municipal sewer network or soil
- identifier of the receiving water body (if applicable and known), referring to the table with waterbodies
- name of first receiving surface water (might not be a waterbody)
- mean value of the discharge per year in m³/a
- identifier of the data source
- comments with additional relevant information related to the discharge point(s).

Stormwater

This section presents the structure for tables containing measurements from stormwater outflows, both from combined sewers and from separate sewer systems.

For the single samples, the following information should be collected:

- sample identifier, ideally using the original identifier provided in the source from which the data come from
- International Generic Sample Identifier (IGSN⁹), a persistent identifier for samples
- identifier of the storm water monitoring site
- place of sampling in the system e.g. outlet of a sewer (without treatment), after sedimentation tank or after soil filters
- begin and end of sampling with time zone
- sampling method, e.g. grab sample, time proportional sample, flow proportional sample, and if the sampling method is accredited
- mean discharge volume of the sampled sewer during sampling
- number of sub samples in a composite sample
- number of overflow events combined in one composite sample
- percent of event volume which was successfully sampled
- identifier of the data source
- comments with additional relevant information related to the sampling procedure.

For stormwater no data for temporal aggregated grab samples were delivered in the project, thus no table was prepared. However, the information for such cases is similar to the one presented for aggregated samples in previous matrices.

For the stormwater sampling sites, the following information should be collected:

- identifier of the treatment plant or monitoring site as used in the original data source and based on which schema the identifier was build

⁹ <https://www.igsn.org>

- country in which the monitoring site is located, coded by ISO 3166-1 alpha-2
- name of the sampling site
- coordinates including the coordinate reference system coded by the European Petroleum Survey Group (EPSG) code, e. g 4326 for WGS84 -CRS
- type of the sewer system sampled: combined or separate sewer
- characterisation of the sewer catchment upstream the sewer outlet: total catchment area, size of the impervious catchment area, size of the impervious and connected catchment area, population number in the catchment, directly connected catchment area (without upstream overflow facilities), traffic areas in the directly connected catchment area, industrial and commercial areas in the directly connected catchment
- storage volume of the CSO/SSO facility
- throttled flow to the wastewater treatment plant in l/s
- typical discharge at the monitoring site in m³/year
- identifier of a related precipitation gauge, where precipitation data are available
- average annual precipitation
- identifier of the data source
- comments with additional relevant information related to the monitoring site.

Atmospheric deposition

Atmospheric deposition is the process of transferring substances from the atmosphere, where they might be available as gas, as aerosol or as suspended dust, onto surfaces during dry weather (dry deposition) or during rain and snowfall (wet deposition).

Dry deposition is caused by condensation and sedimentation of atmospheric matter, while wet deposition is additionally caused by a wash-out effects.

Atmospheric deposition cannot be easily derived from air concentration measurements; therefore, it requires its own monitoring set up. Nevertheless, bulk deposition (dry + wet deposition) is often reported as concentration, but referring to precipitation volume (or even dry matter in the precipitation), not to air volume. It is important to have the duration of sampling, the area of the sampler surface and the sample volume available to calculate deposition rates (mass/(area·time)), the finally needed information.

For atmospheric deposition samples, the following metadata should be collected:

- identifier of the sample as reported in the data source
- International Generic Sample Identifier (IGSN)
- identifier of the sampling site
- begin and end of the sampling with time zone
- sampling method and if the method is accredited
- type of deposition sample: wet only, dry only, bulk
- solids concentration in the sample.
- value of the precipitation in the sample in mm.

- annual precipitation in the year the sample was taken in mm.
- projected surface area of the sampling device
- collected sample volume
- value of the precipitation during the sampling period from an independent precipitation gauge.
- identifier of the data source
- comments with additional relevant information related to the sampling procedure.

For the deposition sampling sites, the following information were collected:

- identifier of the sampling site, as given by the data source
- country in which the sample was taken, coded by ISO 3166-1 alpha-2
- name of the sampling site
- coordinates including the coordinate reference system coded by the European Petroleum Survey Group (EPSG) code, e. g 4326 for WGS84 -CRS
- classification of the sampling site regarding potential main pollution sources: background, urban, industrial, mixed, rural, agriculture, unknown
- long-term mean annual precipitation at the sampling site in mm/year.
- identifier of a correlated precipitation monitoring site
- NUTS level 1,2 and 3 code. In cases where the exact location of the station can not be disclosed, the NUTS units can be used to give rough localization.
- identifier of the data source
- comments with additional relevant information related to the monitoring site.

Soil

Soil can contribute to surface water pollution via soil erosion. Therefore, concentrations in top soils are very useful to support the estimation of HS emissions via this pathway.

In some studies (including Danube Hazard m³c), soil composite samples are created from a certain number of sub-samples taken from different spots in a catchment based on common properties (mainly land use). To be able to represent these circumstances in the database, the relation between samples and sampling sites can be defined as a many:many relationship, namely many samples can originate from one sampling site and many sampling sites can contribute to one sample.

The following meta data should be collected for soil sampling:

- identifier of the soil sample as given in the data source
- International Generic Sample Identifier (IGSN), a persistent identifier for samples
- date on which the sample was taken
- part of the soil column, which was sampled, e.g. Top soil 0-5 cm or Humus cover
- the method how the sample was taken, e.g. from profile, cutting frame (20x 20 cm) or core or sleeve-type borer (split-tube sampler) and if the method is accredited
- share of dry matter from the total sample weight and organic content in the sample
- identifier of the data source

- comments with additional relevant information related to the sampling procedure.

For the soil sampling sites, the following metadata should be collected:

- identifier of the soil sampling site and its source
- name of the sampling site and country in which the sample was taken, coded by ISO 3166-1 alpha-2
- coordinates including the coordinate reference system coded by the European Petroleum Survey Group (EPSG) code, e. g 4326 for WGS84 -CRS
- texture of the soil at the sampling site, with the controlled vocabulary based on the FAO soil texture classification with 12 classes (Food and Agriculture Organization of the United Nations (FAO, 2006)
- genetic type of the soil, e.g. rendzina, brown forest soil, young raw casting soils. Unfortunately, the controlled vocabulary was not harmonized yet, as genetic soil classification systems differ in the different countries quite a lot
- usage/coverage of the land at the soil sampling site, with the controlled vocabulary from the CLC classification
- identifier of the data source
- comments with additional relevant information related to the sampling sites.

In some cases, it may also be interesting to examine lower layers, in this case the suggested tables below should be supplemented with information on the sampling depth and information about the sampled layer.

3.2.3 Data evaluation

Visualization and pre-processing of the data

The data are imported and stored in different tables for measurements (single and aggregated), samples, sampling sites, determinants and data sources, to have the least possible redundancy in the database and reduce therefore mistakes in the metadata. However, this is not a very suitable structure for data evaluation, as information about the metadata of one measurement (what is the CAS of the determinant, when, how and where was the sample taken?) is distributed over several tables. To overcome this problem, data views shall be defined in the database to join together all the information from different tables into wider view tables. Data retrieval can therefore be done with much less effort from these views, which deliver a directly usable result data set.

The actual data analysis, cleaning, filtering, plotting and statistical testing shall be implemented in a suitable data processing and statistical analysis environment. Within the DHm³c project, these steps were implemented in R (R Core Team, 2022). The data are directly downloaded from the database views into the R programming environment, so the data evaluation is always based on the most current state of the database. In R, further steps of data harmonization, e.g. the recalculation of concentrations from different source units (mg/l, µg/l, ng/l) into one suitable target unit, are implemented in a pre-processing script.

Handling of censored data

When working with HS concentration data, in many cases a significant share of the measurements is below the analytical limit of detection (LOD) or quantification (LOQ). Such data are called censored data, more precisely left-censored data. For these left-censored data it is only known, that they are between 0 and a certain threshold but no better estimate about their value is available. Regrettably, the most common way of treating such censored data is the imputation of a fixed value (between 0 and the censoring threshold, in most cases LOQ/2). It has been widely demonstrated in the literature that this imputation introduces bias in the data and can lead to false conclusions when applying statistical tests and calculating descriptive statistics (Helsel, 2006).

Better methods are available from the field of survival analysis and are introduced for the field of water quality data in Helsel (2012). For the R programming environment two packages are available with tests and functions for working with left-censored data: NADA, NADA2 (Lee, 2020; Julian & Helsel, 2023). If data evaluation is done in Python, these packages can still be utilized by packages which facilitate the utilization of R packages in Python.

The following statistical methods - readily available in the mentioned packages - were used for example in the project:

- Regression on order statistics (ROS) to calculate descriptive statistics from censored data
- Kendall's rank correlation coefficient to investigate correlation between censored (and uncensored) data. Warning: The computing of this coefficient becomes quite resource intensive if the data sets are bigger.
- Peto & Peto modification of the Gehan-Wilcoxon test (Peto & Peto, 1972) which investigates if at least one group of two or more groups of concentrations is significantly different than the others.
- Adapted box-and-whiskers-plots (Figure 25) with clear indication of the (range of) LOQ and notched boxes to display the confidence interval of the median. The notches allow to visually check if the median values are significantly different with a confidence interval of 95%.

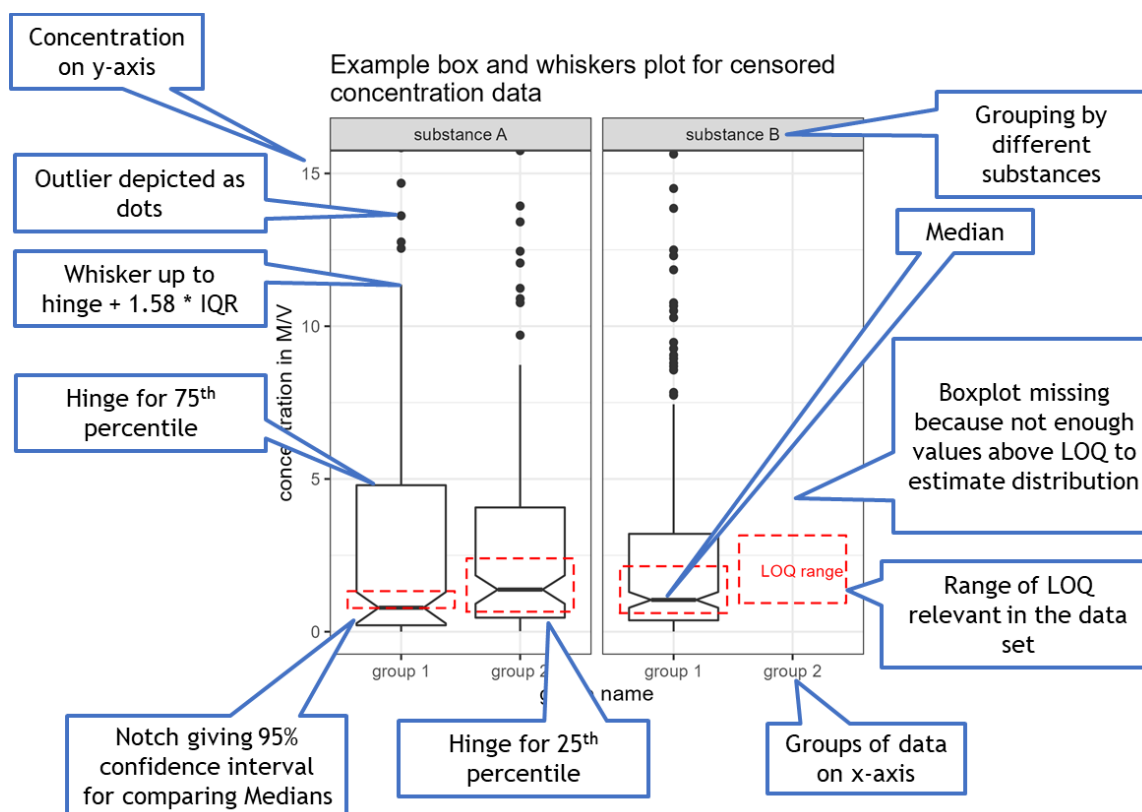


Figure 25: Box and whiskers plots to present and compare groups of HS concentrations censored at multiple levels (Kittlaus 2023, CC-BY-SA 4.0)

3.3 Scope of use of the information in the relational database. Application examples

The inventory of concentrations can be used for several aims. The following sections present three examples based on analyses carried out on the database generated within the DHm³c project: 1) comparison of concentrations in water bodies with environmental quality standards (EQS), 2) derivation of model input data for emission models and 3) derivation of model validation data for emission and water quality models.

3.3.1 Comparison of HS concentrations in surface water bodies with EQS

The harmonized database facilitates the evaluation of water bodies regarding their status of pollution. It is important to highlight that the chemical status assessment is a complex task for the Member States (MS). During the status assessment, MSs can use the load-based grouping of water bodies, consider the correction with bioavailable concentrations, the natural background concentrations, and use the results of biota and/or sediment sampling. Furthermore, the data base stores not only surveillance monitoring sites, but also all available sampling sites, operative, investigative and other measurement results from e.g. research studies, which are not necessarily part of the "official" state assessment of the countries. In the framework of the DHm³c project, the task is not to perform the full official

evaluation of the status, but nevertheless, we would like (1) to provide a picture of the quality of the data available in the data base, and (2) to provide a deeper insight into the information that can be extracted from the data through the example of a substance that can be qualified relatively easily.

Generic assessment, quality of the data available in the database

The simplest way of assessing the status (generic assessment) is to calculate the annual mean concentration for all available monitoring stations and compare it to the water annual average environmental quality standard (AA-EQS) (European Parliament and the Council, 2013).

Figure 26 presents for selected substances at how many stations measurements were performed and which share of the stations have sufficient data to derive a mean annual concentration which can be compared to the AA-EQS to derive a status assessment. The criteria to decide if the available data is of adequate quality to derive a generic assessment are given below. They were selected as a compromise between assurance of a certain quality of the assessment and the aim to derive a status for as many stations as possible:

- At least 12 grab samples or 6 composite samples need to be available (over the whole time period where data are available).
- Less than 80 % of measurements can be below the analytical LOQ, except the case where the LOQ is lower than the EQS.

These criteria can be adapted depending on the evaluation to be carried out and are used here as example to showcase how the inventory can support the status assessment at the Danube basin scale.

In Figure 26, it is visible e.g. that Ni and Pb were measured at almost all sampling locations and that Cd is also a frequently measured element. For these metals, often local or water body specific EQS are applicable, which are considerably higher than the generic EQS, so it may occur that the data frequency appearing in the figure is not of sufficiently high quality. Therefore, it is important to emphasize the importance of the EQS correction with bioavailable concentration. Descending the list of substances in the figure, the “old” priority substances can be seen, which have been routinely measured by laboratories for more than a decade. Starting with the banned pesticides (e.g. DDT), where the number of required samples could be significantly reduced, since several water bodies are in good status and no new loads are expected to be emitted to the water body (WFD annex V. 1.3.3). PAHs are also long-known priority substances, but their EQS values are so low in the monitoring of the water phase that measuring them is still challenging for laboratories. This partly explains why the 2013/39/EU directive introduced EQSs for biota monitoring, moreover, the proposed, revised directive already promote to measure PAHs in sediment too instead than in the water phase. This also shows how important it is to choose the right sampling medium during monitoring. The next group of substances in the list are the chlorinated hydrocarbons, which have also been priority substances for a long time. Given that their effect on the status are well known and their quantification in laboratory does not pose a major challenge, the number of samples can be significantly reduced in many places in order to reduce monitoring costs. Bifenox opens the list for the pollutants introduced by the 2013/39/EU Directive. Their measurement in the latest RBMP planning period was a serious challenge for the laboratories, and the Member States were still mostly groping in the dark about where they should be measured

for the official status assessment. With respect to these substances, a significant quality improvement is expected in the upcoming RBMP planning period.

The substances at the end of the list (DEHP, HBCDD, PFOS, C10-13 chloroalkanes, tributyltin compounds) are those, for which water phase monitoring is hardly feasible and the basis of the status assessment is biota monitoring. This is the main reason behind the small number of results measured with higher LOQs (measurement with a higher LOQ can be useful for emission calculations and to compare the concentrations to the MAC-EQS).

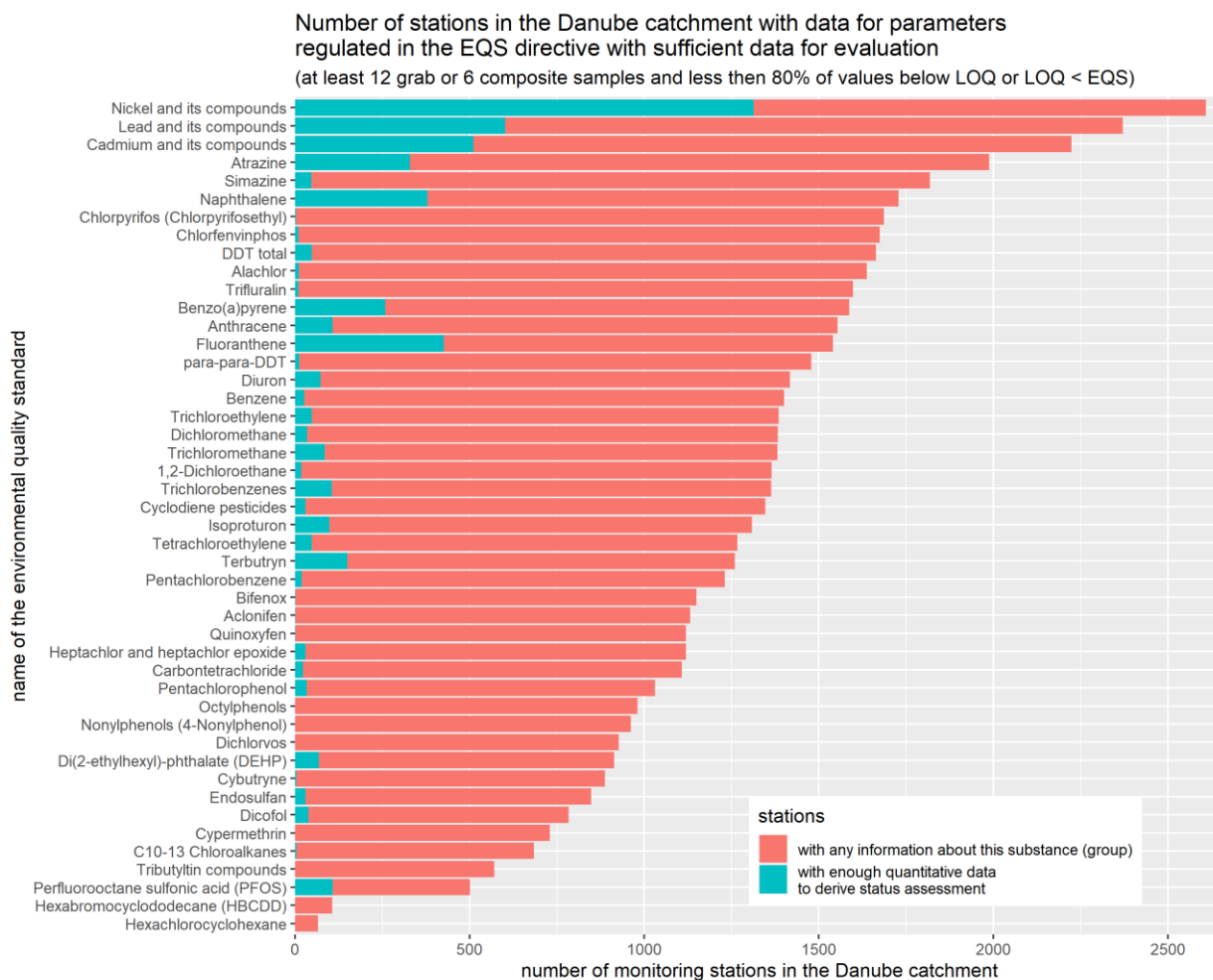


Figure 26: Number of monitoring stations in the Danube catchment with data for status assessment

Finally, the generic status based on monitoring in the water phase for all stations with adequate number of quantitative values can be assessed and it becomes visible which substances are critical in the Danube basin (Figure 27). The generic status assessment can then be visualized on maps (example see Figure 28).

In Figure 27, the high percentile of the grey lines indicates that the assessment based on specific conditions, carried out by experts in the Member States, is very extended. Figure 28 also shows the spatial distribution of the monitoring locations with adequate quality measurement for Ni, which mainly consists of the surveillance monitoring network of the EU countries. Note: The Ni EQS used refers to the bioavailable concentration.

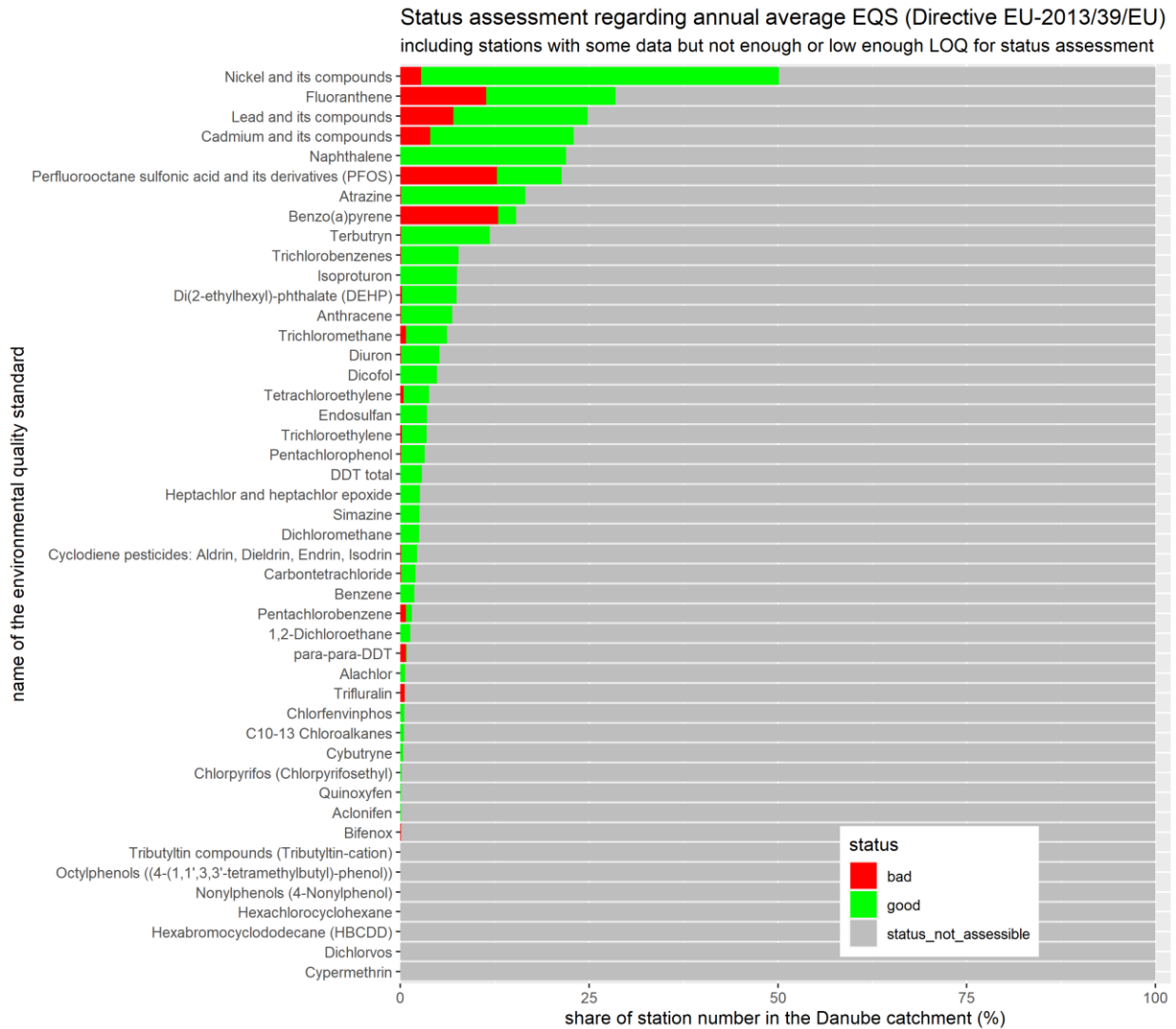


Figure 27: General surface water bodies status assessment regarding AA-EQS

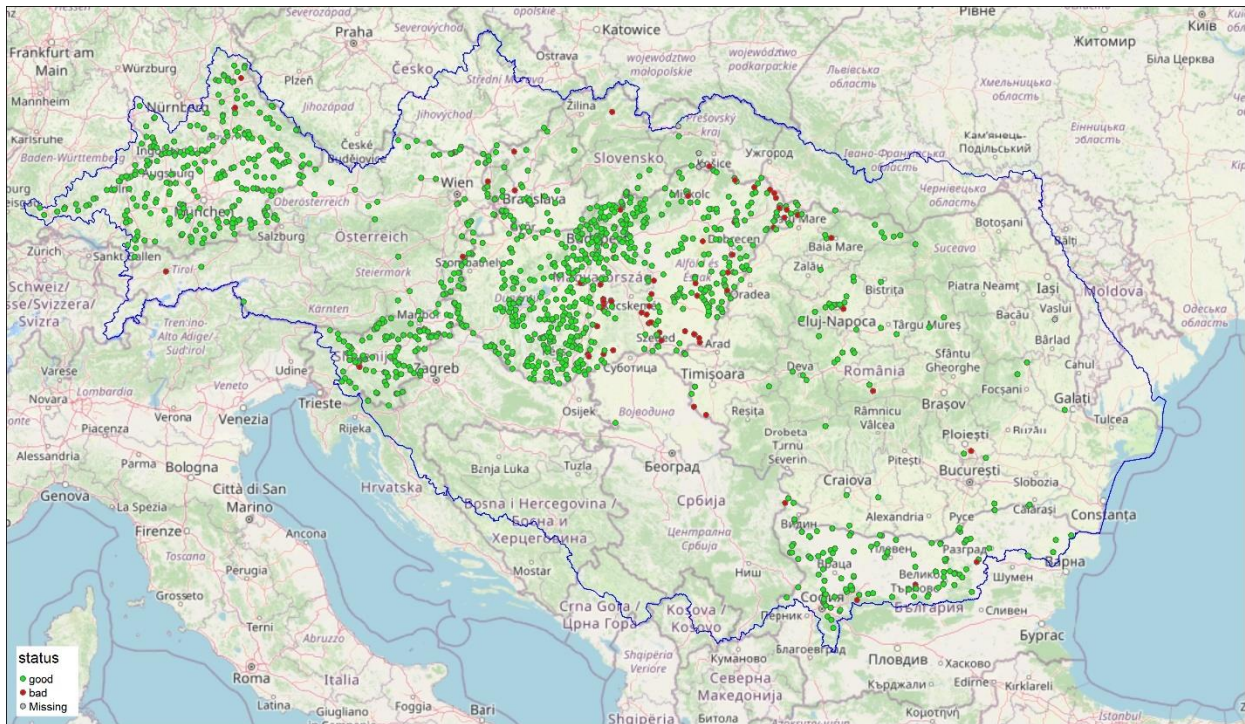


Figure 28: Example for the general status assessment (preliminary results for Nickel). Data sources: background map: OpenStreetMap contributors, Danube river basin: ICPDR

Nicosulfuron: example for evaluation of occurrence and of quality of data in the DRB

Nicosulfuron is a currently permitted pesticide and is a candidate new Priority Substance with following proposed EQS values in surface waters: $MAC-EQS_{\text{freshwater}} = 0.23 \mu\text{g/l}$ and $AA-EQS_{\text{freshwater}} = 0.0087 \mu\text{g/l}$.

Table 8 reports the current data availability in the data base for Nicosulfuron in rivers and in different compartments in the Danube River Basin countries. It can be seen that surface water was measured for Nicosulfuron in only 3 countries (AT, DE and HU), groundwater in 4 countries (AT, DE, SI, SK) and wastewater in only one country (HU). This is complemented by the 2019 campaign of the EU JDS4, which carried out measurements in the Danube and its main tributaries, but only at a few sampling locations.

The 9676 surface water sampling results were measured at 538 locations, with 83.5% of the data being below the LOQ. The 13510 groundwater sampling results derive from 2458 locations, with 99.6% of the data below the LOQ. In the case of wastewater treatment plants, 119 influents and 146 effluents were sampled and 90.2% of the data were below the LOQ.

Table 8: Number of measurements for Nicosulfuron included in the database.

Compartment	Countries and data source type	Number of measurements	Monitored time period	
Surface water (water)	Total	9676	2010	2020
	AT-GZUV-RW	195	2010	2015
	DE-LFU-QUAL	4277	2014	2020
	DE-LUBW	440	2015	2020
	EU-JDS4	153	2019	2019
	HU-FEVISZ/KEHOP	4611	2015	2019
Surface water (SPM)	No data		-	
Groundwater	Total	13510	2009	2020
	AT-GZUV-GW	12142	2009	2020
	DE-LFU-QUAL	1209	2014	2020
	EU-JDS4	21	2019	2019
	SI-SEA	112	2017	2020
	SK-PVM	26	2014	2020
Atmospheric deposition	No data			
Soil	No data			
Stormwater	No data			
Wastewater	Total	265	2019	2019
	EU-JDS4	44	2019	2019
	HU-KEHOP	221	2019	2019
Sewage sludge	No data			

Figure 29 shows the distribution of the data among countries.

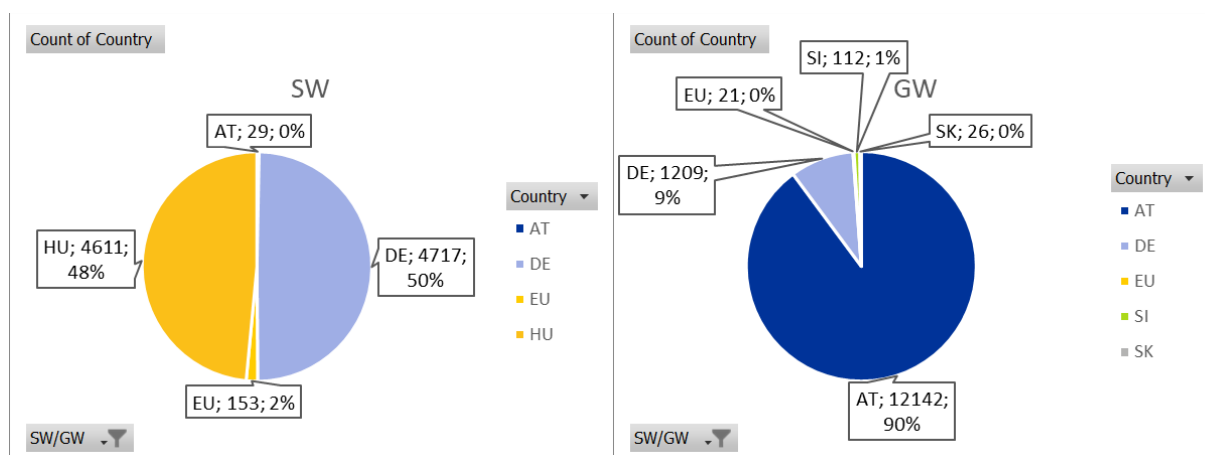


Figure 29: Distribution of surface and groundwater data for Nicosulfuron per country

In both industrial and municipal effluents, the measurement results exceeded the limit of quantification 18 times (with LOQ ranging between 0.004 - 0.1 µg/l). The total available 146 measurement results have poor territorial distribution and cannot thus be considered representative of the Danube River Basin. However, it can be seen that the pollutant can arrive via WWTP, so recommendations of further screening monitoring are justified. Presumably, the pollutant can be present in a relevant quantity in the sewage network. Emission factors cannot be estimated from this information basis, because the LOQ value is not sufficiently low (median is under LOQ) and due to very poor data availability. Further investigation is required by examining the inflow and outflow of WWTPs.

Whether this contaminant can be emitted substantially through stormwater overflows is unknown. For now, there is no information based on the measurement results as to whether the urban runoff could be a potentially relevant source of pollution in the Danube River Basin.

No sample analysis was performed from soil samples, even though this would be very important to be able to model the transport through erosion. It can be seen from the inventory that the pollutant reaches the groundwater as well. Soil erosion, soil infiltration and interflow transport are thus potentially relevant emission pathways which should be further investigated. No sample analysis was performed from atmospheric deposition either; thus, it is not possible to know whether this contaminant can be transported through this route as well in substantial amounts.

With respect to groundwater, only 13% of the measurement values can be used for comparison with the EQS in surface waters, since the LOQ in the applied analytical methods was higher than the EQS. This data basis does not enable a reliable evaluation, but 97% of the high-quality data were measured below EQS, while in 8 cases values above MAC-EQS were identified. Based on this observation, it can be noted that Nicosulfuron can also be occur in relevant quantities in groundwater and therefore groundwater appears to be a relevant pathway, but for a more precise analysis more information from monitoring is needed.

Figure 30 depicts the results of measurements in surface waters, as well as the corresponding EQS values. 39% of the measurements could not be considered in this evaluation, because the applied LOQ exceeded the EQS. This high share of data not valid for further evaluation points at the strong need for

improvement of analytical methods in the laboratories. Within the valid dataset, 73% of values were below LOQ, however, as shown in Figure 30, several measurements often exceeded the EQS. As expected for a pesticide used in agriculture, peaks of highest concentrations and exceedances were measured in summer months. This strengthens the importance of gathering more information and on being able to reliably model its emission via erosion surface runoff and in particular of prioritizing monitoring during the summer months, instead of distributing it regularly over the whole year.

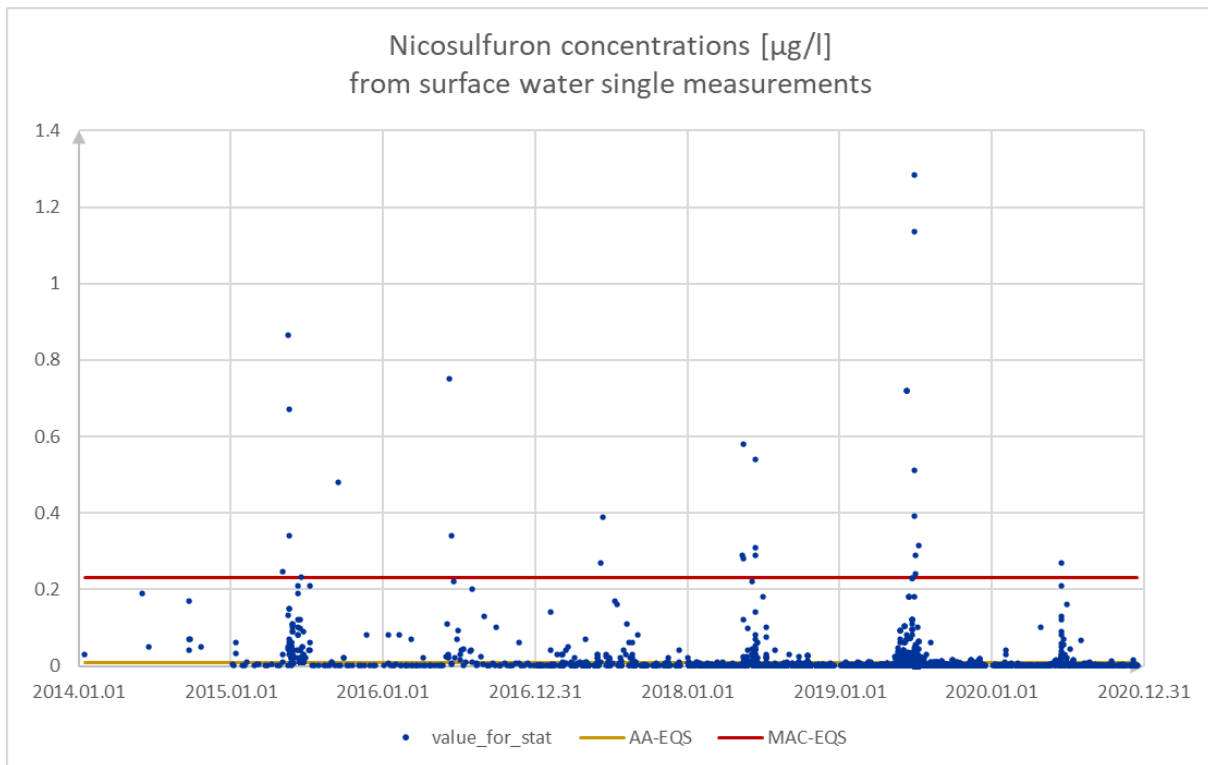


Figure 30: Nicosulfuron measured concentrations in surface waters in the DRB.

3.3.2 Model input data and improved system understanding

The second example for using inventory of concentrations is the derivation of input data for emission models. The first valuable input for the emission modelling is a rough estimation of the relevant pathways. This can be achieved by comparing concentration levels in different environmental compartments, a task which can be greatly facilitated by the database (Figure 31). Such analysis gives a first indication of which emission pathways might play the biggest role for different substances and thus helps to prioritize efforts in emission modelling.

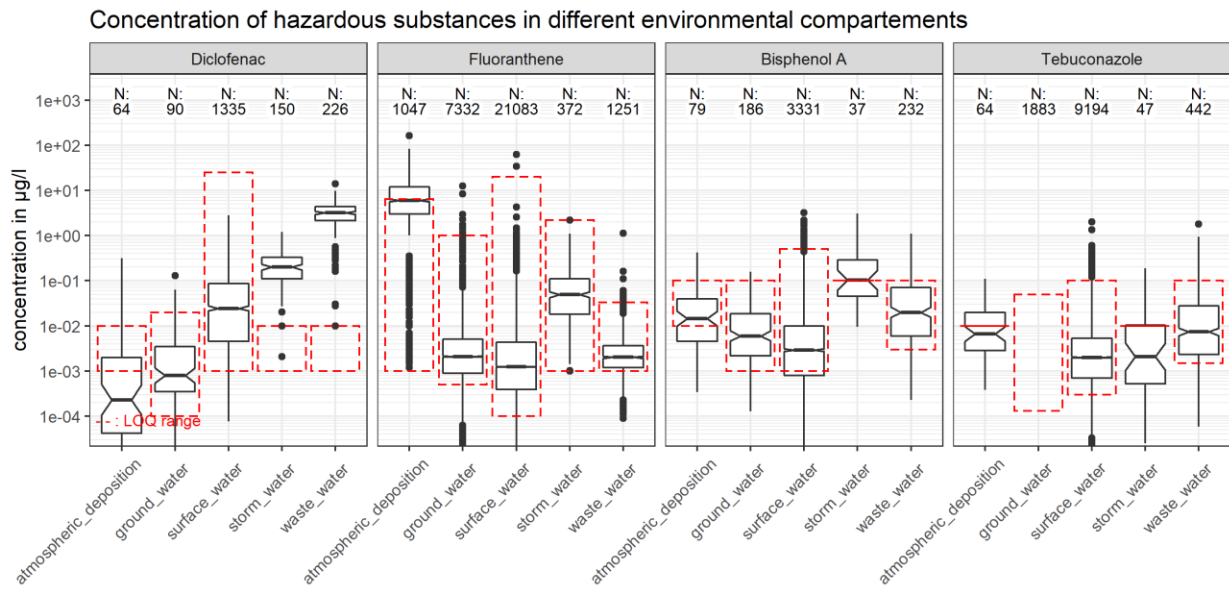


Figure 31: Comparison of concentration levels in different environmental matrices as a result from the inventory of concentrations. 'waste_water' refers here to treated wastewater effluent.

For example, it can be seen in Figure 31 that for fluoranthene (a polycyclic aromatic hydrocarbon) and bisphenol A (an industrial chemical used in production of plastics), stormwater (here not differentiated between combined sewer overflow and storm sewers in the separate system) plays a major role in emissions. While for fluoranthene the even higher concentrations in atmospheric deposition clearly point to atmospheric transport and deposition on surfaces as the main pathway, for bisphenol A other sources seem to contribute to determine the load in stormwater runoff. Contrary to these two substances, for diclofenac (a pharmaceutical compound) the main pathway is the WWTP effluent, followed by stormwater (here mainly due to the share of sewage in combined sewer overflows).

As next step, the concentration levels for different quantiles of the data set can be abstracted from the database to be used as general input data for a pathway in the model. Different quantiles of the data can be used for different model variants (e.g. best case and worst case). Exemplarily, in Table 9, the data for perfluorooctanesulfonic acid (PFOS) are shown; surface water concentration levels are added for comparison. In several cases, the mean values are higher than the 75th or even the 90th percentile values, which indicates the presence of extreme outliers in the data collective.

Table 9: Statistical description of concentration levels in different environmental compartments, as evaluated based on the data base, here for the example of PFOS. P10: 10th percentile of the data, ww: wastewater .

Compartment	Unit	P10	P25	P50	Mean	P75	P90	Share censored data	Number of aggregated values
Surface water	ng/l	0.1	0.2	0.6	1.9	1.7	3.2	0.3	2986
Groundwater	ng/l	0.0	0.0	0.1	4.9	0.7	4.0	0.8	4161
Municipal ww, untreated	ng/l	0.0	0.2	0.7	7.2	3.5	13	0.7	207
Municipal ww, treated	ng/l	0.7	2.0	3.0	7.5	5.6	11	0.2	1257
Combined sewer overflow	ng/l	1.4	2.2	2.6	2.4	2.9	3.1	0.1	8
Storm sewer outlet	ng/l	0.5	1.0	2.9	5.7	7.7	11	0.3	24
Atmospheric deposition	ng/l	0.0	0.0	0.0	0.3	0.1	0.5	0.8	88
Soil	µg/kg	0.0	0.0	0.1	0.4	0.1	0.5	0.2	99

Finally, different groups of data grouped by metadata available in the database can be checked for statistically significant differences and if these are found and possibly hypothesis exist about the mechanisms causing them, they can be used to improve the input data set of the model. This should be shown with two examples.

Example 1: PFAS in municipal wastewater treatment plant effluent

PFAS are a substance group, for which the effluents from municipal wastewater treatment plants are a significant pollution pathway, as conventional WWTPs cannot significantly reduce the PFAS load of wastewater. Therefore, it is important to gain a better understanding of which factors influence the concentration level in WWTP effluents.

In Figure 32, the effluent concentrations are presented depending on the size of the treatment plant. It can be observed that bigger plants in general show higher effluent concentrations, while the variability of the concentrations is very high. This points to the fact that not only the size of the treatment plant might influence the effluent concentrations, but other factors as well. Nevertheless, this finding can already improve the emission modelling by application of different effluent concentrations for WWTPs of different size.

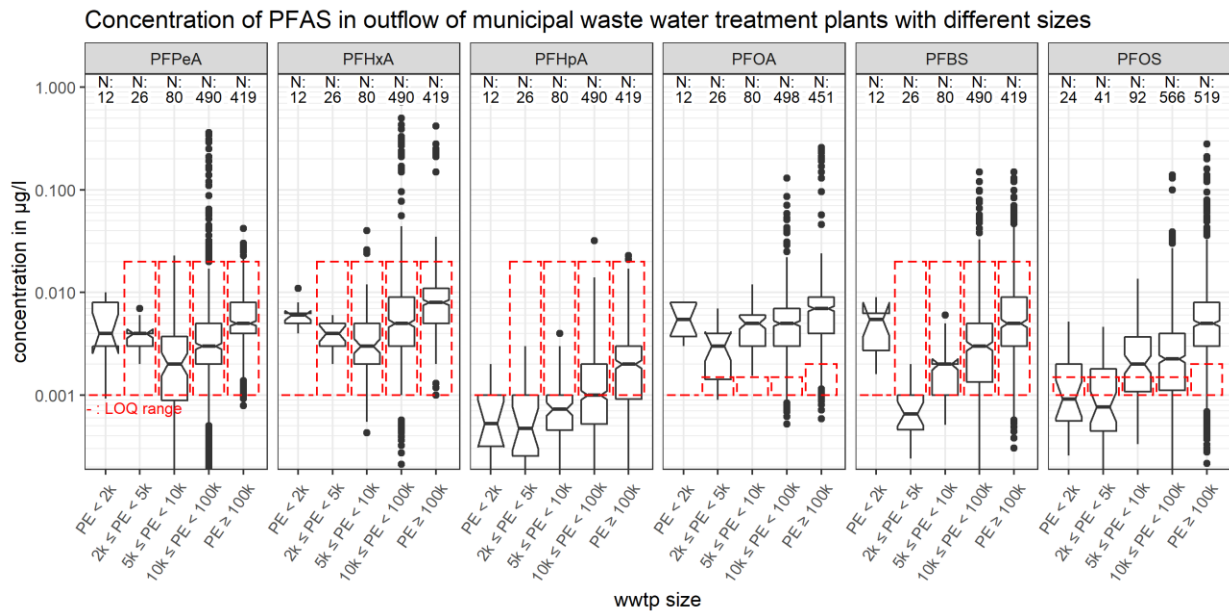


Figure 32: PFAS concentrations in municipal WWTP effluents depending on the size of the treatment plant given as design capacity in population equivalent (PE). N: the number of measurements available.

In Figure 33 and Figure 34 concentrations are shown against country and applied treatment technology, two further factors which might influence the effluent concentrations, while country is only a proxy for regional different application or emission patterns. On the one hand it becomes apparent that differences of PFOA and PFOS concentrations in different countries are statistically significant. On the other hand, it is clear that the data on effluent concentrations in the database currently only cover few countries and types of WWTPs, which does not allow to properly investigate the influence of spatial heterogeneity in the whole basin or dependence of the effluent concentrations from the applied treatment type. Further monitoring data from a broad range of WWTPs would be needed to get a more comprehensive picture.

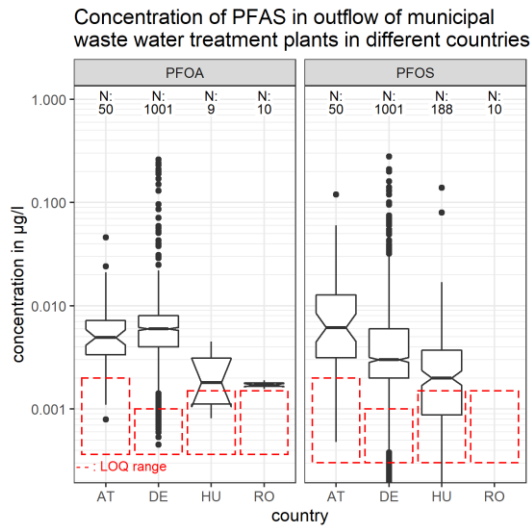


Figure 33: PFAS concentrations in municipal WWTP effluents in different countries. Only for PFOA and PFOS enough quantitative data were available for this comparison.

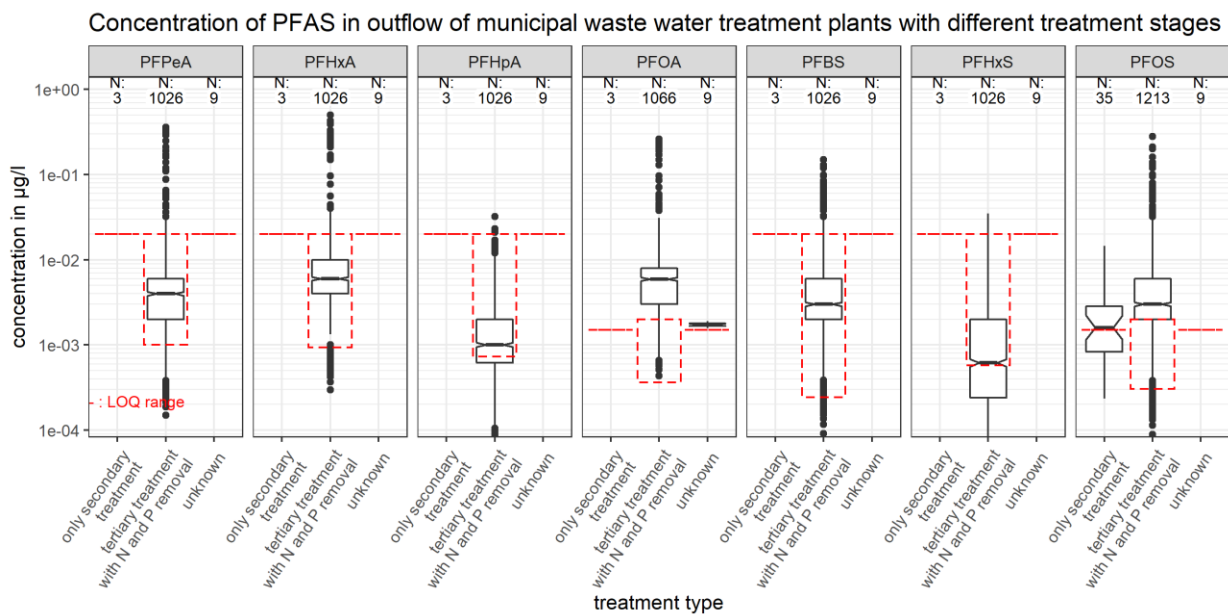


Figure 34: PFAS concentrations in municipal WWTP effluents depending on applied treatment technology.

When implementing further monitoring programs, it would be valuable to collect as much meta data about the treatment plants and their catchment as possible, e.g. the share of industrial waste water in the influent, the type of industries located in the catchment, the connected impervious surface area and the connected inhabitants (not only as Population Equivalent).

Example 2: Heavy metals in top soil

For metals, the pathway of soil erosion is one of the important pathways in areas with significant slopes. For the quantification of emissions via soil erosion, in addition to the soil loss from different

areas or sediment input loads into rivers, the concentration of metals in the mobilized top soil are the important input data for the emission modelling. In Figure 35, the concentrations in topsoil of different land use classes are presented.

While for As, Cd, Ni, Cr, Cu and Zn the concentrations in forest and other soils covered by permanent “natural” vegetation are lower than for other land uses, for Hg and Pb concentrations in these soils are higher. This might be explained by the fact that these two elements are mainly transported via air and reach the soil via atmospheric deposition. In forests, the roughness of the vegetation cover is higher than on pastures and arable land and therefore higher deposition rates may occur.

Furthermore, in forests the soil is not perturbed very often, keeping the enrichment of these elements in the first centimetres of the top layer while on agricultural areas the soil column is perturbed and the inputs from atmospheric deposition or fertilizers are diluted by mixing with deeper layers (up to 30 cm). Another possible explanation for deposition being higher in forests than in other land uses is the large LAI (leaf area index) that characterizes forests. According to this idea, heavy metals transported by air movements deposit on leaves, and fall/settle in autumn and become soil (in this case pollutant-reach soil) throughout the years afterwards. Due to harvest, this process does not happen on agricultural fields. Further investigation is needed to confirm this idea.

A further noteworthy fact are the highest Cu concentrations associated to “diverse agricultural areas” - a land use class which includes vineyards. It is known that copper is used in viticulture as fungicide and thus the elevated concentrations in such soils is reasonable.

The different concentration levels on different land uses can be used in modelling to get spatially better model results. A more detailed classification of land use and further information like soil texture and soil type - as foreseen to be included in the database - might be helpful to further explain the observed concentrations. Unfortunately, these metadata are not yet available for all soil samples.

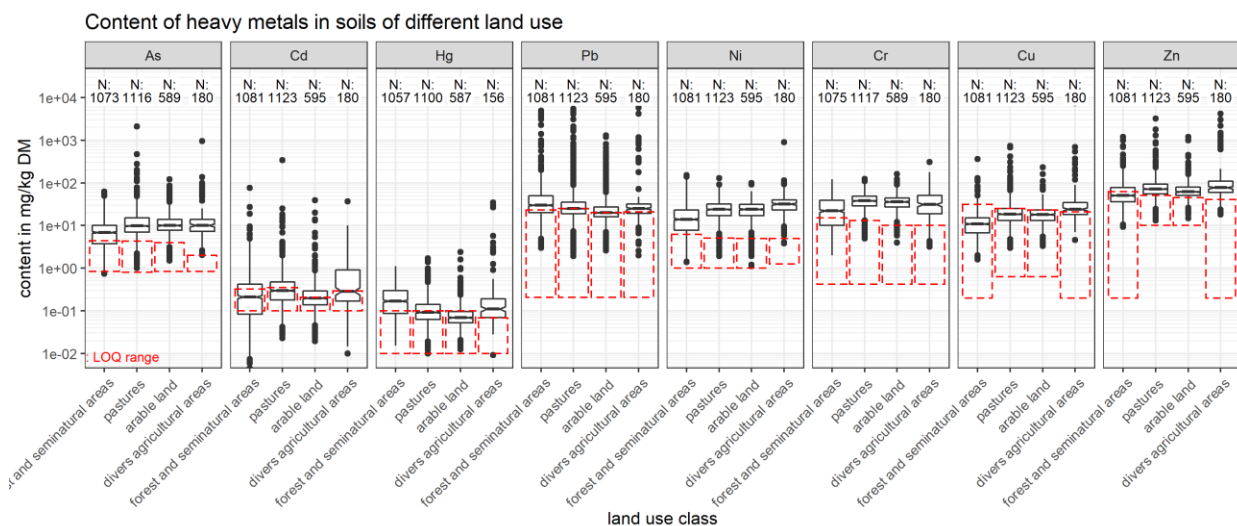


Figure 35: Heavy metal (and arsenic) concentrations in topsoil depending on the land use class of the soil sampling site.

3.3.3 Model validation data

The third examples for using inventory of concentrations is derivation of model validation data for emission and water quality models. Emission and transport models are used to quantify the overall emissions. Such models need to be validated against observed river concentrations and loads, which rely on sufficient data from river monitoring. The MoRE model outputs are for example total annual loads and flow weighted mean annual concentrations. To calculate validation data for the model, monitoring data generated within the Danube Hazard m³c project was used. More precisely, the discharge data of the monitoring stations in the 7 DHm³c pilot regions and their related concentration data was used for the validation of the MORE model. The river loads are calculated with a load calculation method which was specially adapted to the sampling methodology performed in the DH m³c monitoring program. It distinguishes between flow conditions, i.e. baseflow and event-flow, because the observed substance concentrations varied over the flow conditions. Most emissions model outputs consider uncertainties, e.g. the MoRE model produces variants for best-case, average and worst-case evaluations. The validation data were also calculated with an uncertainty range, e.g. the validation loads for the MoRE model with the 25, 50 and 75 percentile concentration. A detailed description of the material and method can be found in the DHm³c project output O.T1.2 *Demonstration of harmonized and cost-effective monitoring*.

3.4 Lessons learned

Inventory Design:

- Consider wishes by users but do not let the users decide about the database structure, as they might not possess the technical background knowledge and might not be fully aware of the full workflow and technical details necessary for the database
- The development of clear definitions for tables, columns and controlled vocabularies is important. This needs time in advance of the data collection to have a clear request what is needed to those, who collect and prepare the data.
- Redesign/Extension of the database schema during data import and evaluation is not avoidable as new data set bring new type of metadata which is considered valuable and should be included.

Data collection:

- Data accessibility is still a major issue. If data are accessible, unclear licenses are the next obstacles. Open data with clear licenses is the solution for the future, awareness needs to be raised regarding the importance and benefits of free data use. Reservation against open data partly arise from the idea, that the use and quality of data can be better controlled, if the data owning institution is the only one offering the data for use and controls the access to the data by requesting a procedure to get a data usage agreement. Two reasons make this a questionable assumption: If data are difficult to access, they will not be used in many cases, where the effort to reach data use agreements for each single data set is too high to be taken. When data are less used, they are less checked for quality. As most data set have data errors, the frequent use and the associated quality checks can help removing

data errors. The only prerequisite is, that errors if found can be reported to the data owning institution. Clear workflows (contact data) should be available to do so. Finally, databases should receive easy to access and standardized automated programming interfaces (API) which make it possible to automatically reload the data from time to time and therefore avoid having different versions of the data around (corrected ones at the data owning institution, other containing known errors in other places).

- National surface water monitoring programs often store only very limited metadata, mainly only those metadata directly needed for the task the data were primarily collected for. E.g. samples are often not stored as own entity and can only be somehow reconstructed from sampling time and place. Information about the applied sampling and laboratory methods, catchment properties and conditions during sampling are often missing but would be useful for further data evaluation.
- Capacity to handle datasets with more than a few hundred lines is still rather limited in many institutions in the Danube basin. Capacity building in this regard is urgently needed.

Data import stage:

- Mapping of substance names and other metadata vocabularies from different sources into one target schema is a time-consuming task during data import. The use of standardized controlled vocabularies should be strongly encouraged. For substances as many different identifiers as possible like CAS number, NORMAN-SusDat-ID, InChIKey should be delivered with the data, to support the mapping process.

Handling of aggregated data:

- Collection of aggregated values creates many problems and it adds no too much value.
- Composite samples should be collected in separate table or together with single measurements, instead of chosen approach to collect them together with aggregated single measurements.

A high level of technical knowledge regarding data management on the one hand and about monitoring data for HS on the other hand is necessary to set up such an inventory. To reach the necessary expertise working with such data for several years is necessary.

4 Modelling and scenarios development for hazardous substances water pollution mitigation

4.1 Introduction

4.1.1 Why modelling?

The Danube countries, united under the umbrella of the ICPDR, agreed to implement the Water Framework Directive (WFD), the main EU regulatory instrument for water management, throughout the entire basin. The WFD requires the preparation of an inventory of emissions, discharges and losses of all PS. Such inventories should give information on the relevance of priority substances (PS) at the spatial scale of the Danube River Basin (DRB) or the national parts thereof, and on the loads discharged to the aquatic environment, thus supporting subsequent river basin management and WFD implementation. In 2012, Guidance Document No. 28 on the Preparation of an Inventory of Emissions, Discharges and Losses of Priority and Priority Hazardous Substances was issued, as part of the Common Implementation Strategy for the WFD. This introduction is based on the information compiled in this guidance.

Among other purposes, an emission inventory should be seen as a tool, which may be used to:

- assist in establishing and implementing targeted reduction of emissions, discharges and losses of HS, eventually leading to the cessation of emissions, discharges and losses of HS (e.g. by identifying the main sources, their relative share with respect to pollution and their pathways);
- support the definition of effective RBMP Programmes of Measures (PoM);
- identify gaps in knowledge and understanding.

The guidance sketches a two-step approach. The first step serves to identify those substances, which are clearly of minor relevance for the DRB, at present and in the foreseeable future, and to concentrate the efforts of subsequent inventory development on the remaining substances. The second step entails a further elaboration of the emission inventory. For this second step, four so-called “tiers” are discussed.

1. Tier 1 comprises an inventory of point sources¹⁰ using statistical data including point source information reported under the European Pollution Release and Transfer Register (E-PRTR).
2. Tier 2 adds a quantification of in-stream loads, based on concentration and discharge data. The resulting riverine load, in combination with the information gained in tier 1 allows the allocation of observed loads to point and diffuse¹¹ sources. A high contribution of diffuse sources is a reason to proceed to tier 3 or 4.

¹⁰ “Point source” - a single localized point of discharge of wastewater containing one or more pollutant(s).

¹¹ “Diffuse sources” - the many smaller or scattered sources whose combined impact may be significant but for which it is impractical to quantify them individually.

3. Tier 3 is a pathway¹²-oriented approach, using more specific information about the land use, hydrology and basic transport processes of pollutants towards the surface water. Regionalised emissions for small catchments (analytical units) are calculated and subsequently aggregated to larger units.
4. Tier 4 is the source¹³-oriented approach, that sets up a mass balance or Substance Flow Analysis (SFA) and addresses the whole system starting from the principal sources of substance release.

Figure 36 provides a schematic overview of the Tier 2, 3 and 4 and their inter-relation (Whalley et al., 2018).

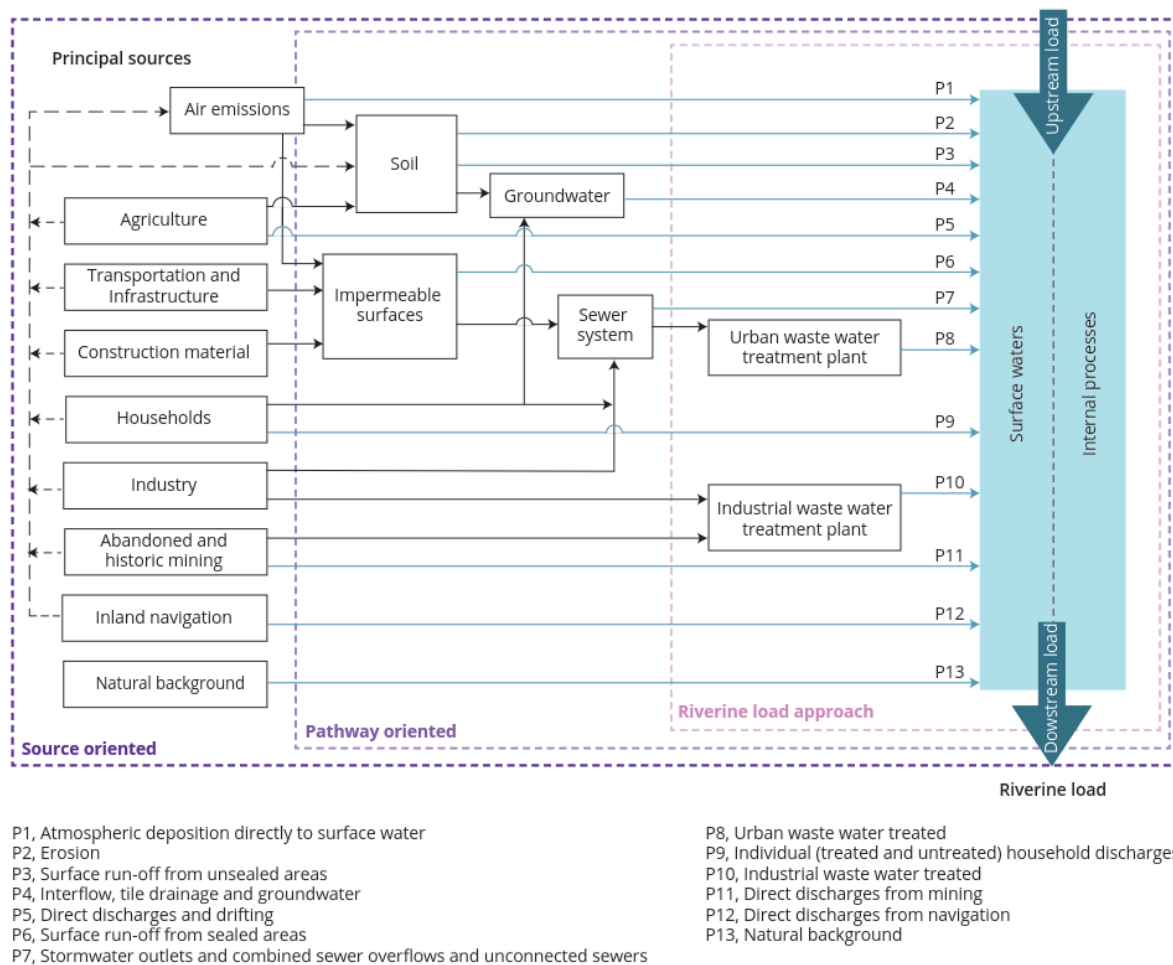


Figure 36: Overview of riverine load (Tier 2), pathway-oriented (Tier 3) and source-oriented (Tier 4) approaches.

The approaches under Tiers 3 and 4 typically involve the use of data-driven models. The use of models in these two tiers supports river basin management practices in three different ways. Firstly, compiling

¹² "Pathways" - the means or routes by which specific substances can migrate or are transported from their various sources to the aquatic environment.

¹³ "Sources" - all processes and activities that are likely to contribute to the input of pollutants into the environment.

an emission inventory provides quantitative understanding of present emissions (diagnosis). This is a prerequisite for being able to define control measures and assess their efficacy. Secondly, it helps understanding what we do not understand (knowledge gaps) and consequently provides focus for future data and information collection. Thirdly, a scenario analysis supported by models helps quantifying the expected effect of measures (prognosis).

The Danube Hazard m³c project applied models in support to both Tier 3 and Tier 4 approaches. This chapter discusses such applications.

The WFD implementation stipulates emission inventories at the spatial scale of the complete basin or the national part of an international basin. Regarding the usefulness of the inventory for RBMP purposes, e.g. for identifying hot spots (areas with high specific inputs of substances) or estimating the effectiveness of measures, the WFD Guidance recommends a significantly higher spatial resolution (~100-~1000 km²). Both model applications discussed here satisfy this requirement. To support water management at a local scale, an even higher spatial resolution is necessary.

The application of any method on these smaller scales heavily relies on the availability, quality and resolution of the required input data. The Danube Hazard m³c project invested in such smaller scale assessments in 7 pilot regions throughout the basin supported by Tier 3 modelling. The lessons learned were then used for a generalisation to the larger DRBD, using the Tier 4 method. It is however obvious that reliable high-resolution input data are not available at the scale of the DRBD. The basin-wide inventory therefore relies on national or even basin-wide statistics combined with regionalisation of emissions using various proxy parameters.

4.1.2 Existing modelling approaches

Models are always an abstraction of the reality. After the selection of processes relevant for the fate of substances (e.g. transport, distribution between environmental compartments, degradation), which can be formalized as mathematical equations, it has to be defined how detailed the modelling approach must and can reproduce this reality. This also highly depends on the necessary spatial and temporal scale and on data availability. It is clear that the goal of modelling always determines the selection of the most suitable modelling approach, but the goal must be necessarily adapted based on the feasibility.

In general, three different classes of models can be distinguished (Figure 37):

- Black-Box models
- Grey-Box models
- White-Box models

Black-Box models represent empiric input-output models, in which all processes are unknown and hidden in a black box. They need the least input data to be setup and allow only a very limited increase in knowledge.

Grey-Box models are semi-empiric, often statistical approaches. Here processes are often reproduced but on a generalized in a highly abstracted way. Equations to describe process behaviour must be evaluated and implemented beforehand and have a statistical character. To run the algorithms to

achieve plausible results which reflect reality, the data needs increase. Typical representatives of such model approaches are MONERIS or MoRE and even the Danube Hazard Substance Model (DHSM) is included in this class. The increase of knowledge is extended but remains at a general level.

White-Box Models represent analytical or physical based models. The processes are based on physical derived equations and to solve them a very extensive amount of input data is necessary. Examples of this class are groundwater models. The model approach demands a deeper knowledge on causal dependencies, but is able to deliver the most detailed increase of knowledge. Because of the huge need of input data, White-Box Models are often restricted to smaller spatial scales. However, due to significantly increased computing power and improvements in data availability these model approaches are increasingly losing their restrictions.

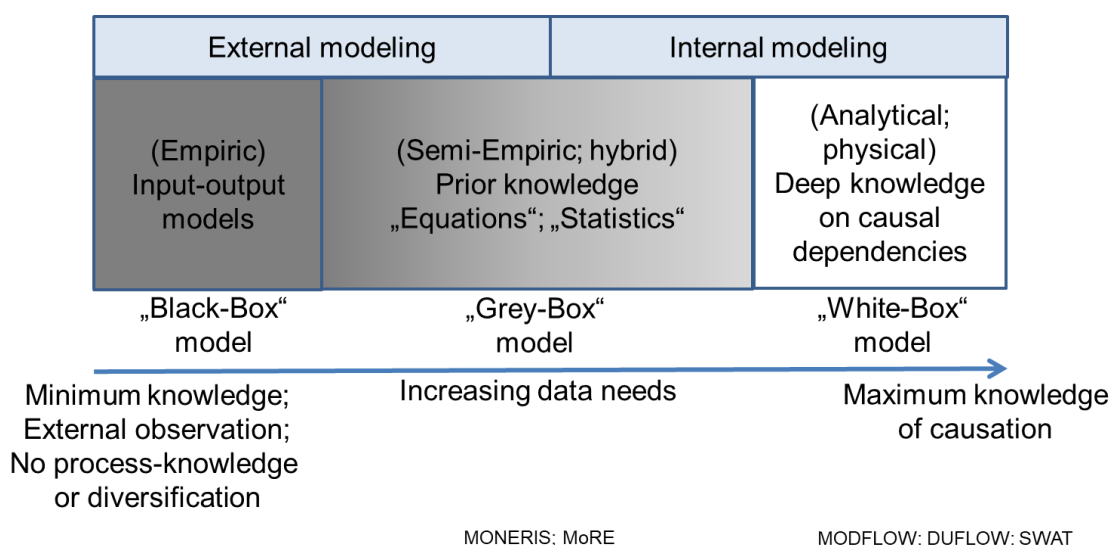


Figure 37: Overview of different classes of models, with different complexity and expressiveness.

4.2 MoRE model

4.2.1 Pathway-oriented approach

In Danube Hazard m³c, a Tier 3 pathway-oriented emission inventory was compiled based on the conceptual model MoRE model (Modelling of Regionalized Emissions) in seven pilot regions in the Danube Basin. The modelled pilots are used as role models for the Danube riparian countries. The approach adopted relies on the Technical Guidance 28 (Figure 36) and reflects the proposals included in a recent report issued by the European Environment Agency (Van den Roovaart et al., 2022).

Pathway-oriented approaches are well established and applied in many European river basins for nutrients and heavy metals and they provide a sound insight in the main transport paths to surface water. However, in the Danube Basin, there is a lack of such approaches as it was stated in the Danube RBMP 2021 update.

The MoRE approach (Fuchs et al., 2017) is a further development of the MONERIS model, which was originally designed and created to calculate nutrient emission (Behrendt et al., 2002). In Europe the MoRE model is currently applied in Germany (Fuchs et al. 2017b) and Austria (Amann et al., 2019) to

quantify emissions for different substances and it was also used in a transboundary project for the Inn river basin (Fuchs et al., 2019).

In Danube Hazard m³c, the complete model environment with numerous variables, formulas, calculation stacks and input data sheets, was translated from the German version (original) to English to enable further applications in the Danube Basin.

Pathways, not yet implemented with relevance for the Danube Basin, were added (e.g. municipal wastewater collected by sewer but not treated in WWTP; significant leakage of old sewer systems; abandoned mining; (Figure 38) and new substances were included in the model structure (phenanthrene, 4-tert-octylphenol, nonylphenol, bisphenol-A) or quantified (chrome, arsenic, tebuconazole, s-metolachlor and its metabolites s-metolachlor ethanesulfonic acids (ESA) and s-metolachlor-oxanilic acid (OA)).

In the Technical Guidance 28, pathways are classified into three blocks:

- Pathways dependent on point sources
- Pathways dependent on diffuse urban sources
- Pathways dependent on diffuse non-urban sources

One benefit and reason for the use of models is the evaluation of diffuse emission, which cannot be simply calculated from data or measurements, but are subject to more or less complex processes that can be reproduced with model approaches in a simplified, abstracted form.

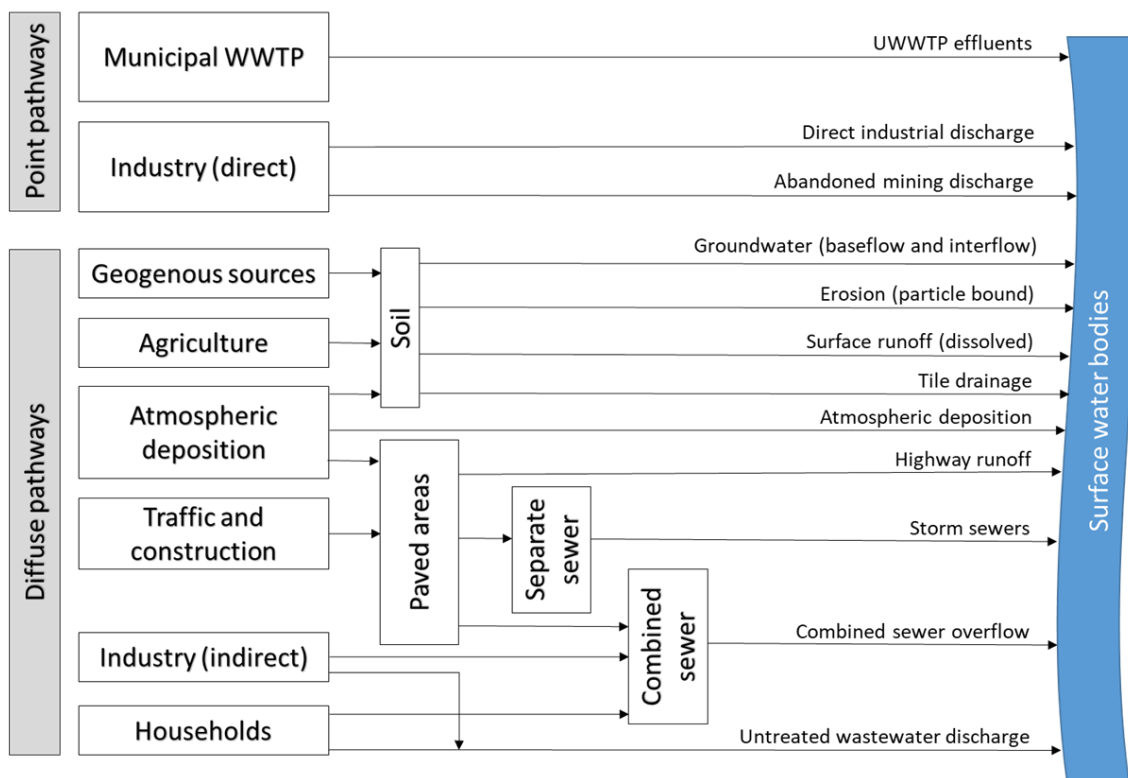


Figure 38: Pathways from point sources and diffuse sources calculated in the adapted DHm³c MoRE model version (adapted from Fuchs et al., 2017a).

4.2.2 *Scope, spatial and temporal scales*

The MoRE model environment for the pilot regions was setup for 27 substances representing different substance groups. The selection reflects the target chemicals of the Danube Hazard m³c project, chosen to fulfil various criteria: (1) substances representing relevant sources and pathways, (2) substances relevant for ICPDR, national and regional authorities in the basin, (3) substances that can be actually detected and measured, so that data can be expected to be available. Because the substances selected for modelling represent relevant sources and pathways, their emission inventories are expected to be representative for a much wider group of substances, and to provide insights for defining effective pollution control measures beyond the modelled substances alone:

- Metals: arsenic, cadmium, chrome, copper, nickel, lead, zinc and mercury
- Polycyclic aromatic hydrocarbons: benzo[a]pyrene, fluoranthene, phenanthrene, naphthalene and the sum of EPA 16-PAH
- Pharmaceuticals: diclofenac, an anti-inflammatory drug and painkiller, and carbamazepine, an anticonvulsant used in the treatment of epilepsy and neuropathic pain
- Industrial chemicals with wide dispersive use: 4-tert-octylphenol, nonylphenol, bisphenol-A.
- Per- and polyfluoroalkyl substances: perfluorooctanesulfonic acid (PFOS), perfluorooctanoic acid (PFOA), perfluorobutanesulfonic acid (PFBS), perfluorohexanoic acid (PFHxA) and perfluoropentanoic acid (PFPeA)
- Pesticides: tebuconazole a fungicide used for wood preservation, and metolachlor and two of his metabolites (metolachlor-ESA & metolachlor-OA), an herbicide in agriculture.

During the execution of the monitoring in the pilot regions (project output O.T1.2 *Demonstration of a harmonized and cost-effective monitoring*), the hot spots of contamination became clear, but other previously selected substances were found to be exclusively, or almost exclusively, below the limit of detection or below the limit of quantification (benzo(a)pyrene, fluoranthene, phenanthrene, naphthalene and the sum of EPA16-PAH, 4-tert-octylphenol, nonylphenol, bisphenol-A). For these substances, no complete modelling was applied. Validation of the model in these cases would exclusively depend on the calculation convention of the values below the limit of detection and the limit of quantification used for the calculation. Consequently, model results would contain a very limited and uncertain gain in knowledge. However, due to the prepared model structure, modelling can be applied easily for these substances, if catchments with higher concentration levels and emission loads are investigated.

The spatial unit of reference for modeling with MoRE is the mesoscale. In 7 pilot regions with 34 sub-catchments the units range from 41 to 667 km² (mean of 232 km²) with a wide range of natural conditions and anthropogenic uses included. The total area under modelling is almost 8000 km², which is almost one percent of the Danube River Basin. The spatial distribution of the pilot areas in the Danube basin are shown in Figure 39.

The delineation of sub-catchments is one important prerequisite for a successful modelling. The delineation is carried out based on DEM (raster data), water network (raster data) and the definition

of outlet points (vector data). The outlet points must be selected in such a way that on the one hand hydrological conditions are considered, and at the same time the possibility of hydrological and chemical model validation based on discharge information and concentration measurements is optimized.

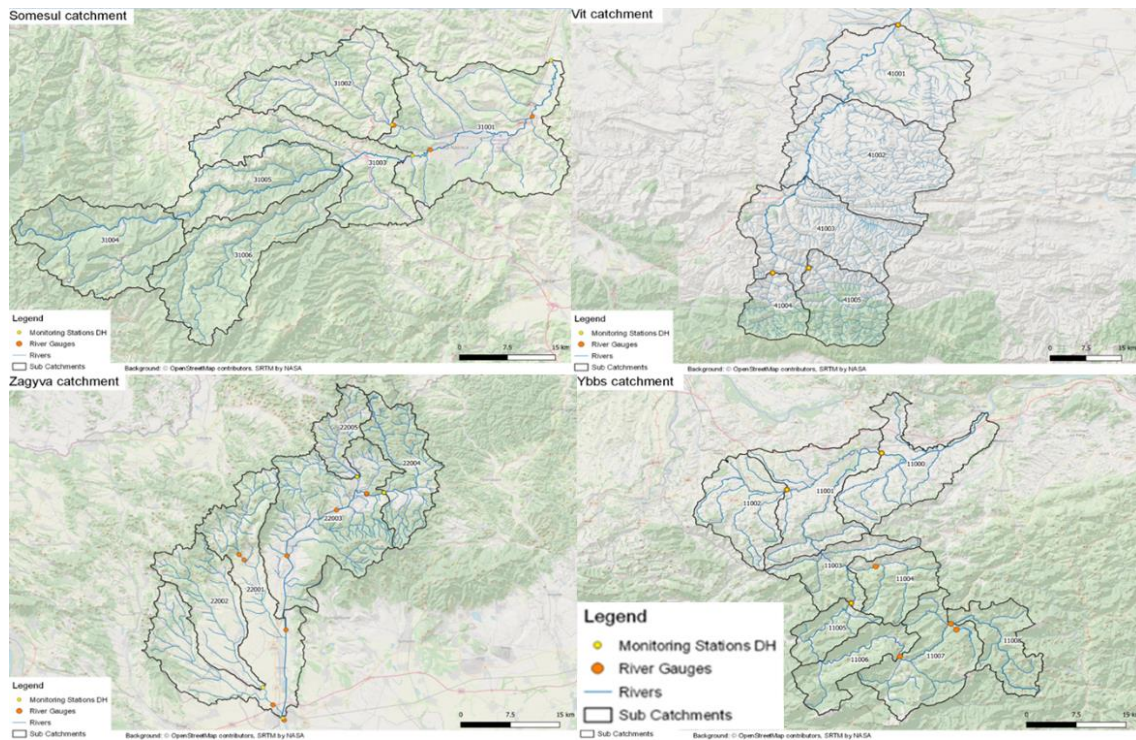


Figure 39: Selected pilot regions in Romania, Bulgaria, Hungary and Austria. Sub-catchments, monitoring stations and river gauges (Data sources: OpenStreetMap Contributors, SRTM by NASA).

MoRE operates on annual time steps. Due to the variability caused by years with different climatic and hydrological conditions, a five to six years period is modelled, assuming balanced conditions over the average of the entire period. The time step makes the inclusion of seasonal fluctuations impossible. However, the comparison of dry or wet years gives insight into climate driven variability of the annual results.

4.2.3 Main calculation approaches

The standardised technical workflow to setup and run the MoRE model can be summarized in four main steps:

- Evaluation and pre-processing of input data (section 4.2.4)
- The creation of variables (constants, time and catchment-related variables, time-related point source variables) and import of data
- Creation and definition of calculation approaches (formulas, calculation paths, calculation stacks)
- Calculation of catchment-related emission and concentration by use of the model quantification kernel for all pathways presented in Figure 38.

- Documentation of master data (e.g. input data, constants, calculation approaches).

The MoRE Developer is the graphical user interface with well-structured main menu and easy-to-use sub-menus, which accesses the SQL database with all variables, constants and formulas created for the model (also in former applications). This has the advantage that information and solution approaches grow with the model, are well documented, can be easily deactivated and replaced by new solution approaches, that better represent the prevailing conditions and data availability. This gives the model approach a great deal of flexibility.

Based on the modelling results further steps are necessary:

- Comparison of model results with data from monitoring (run-off, loads and concentration)
- Check and interpretation of the pathway-related results
- Visualisation of model results (here MoRE also provides a user interface).

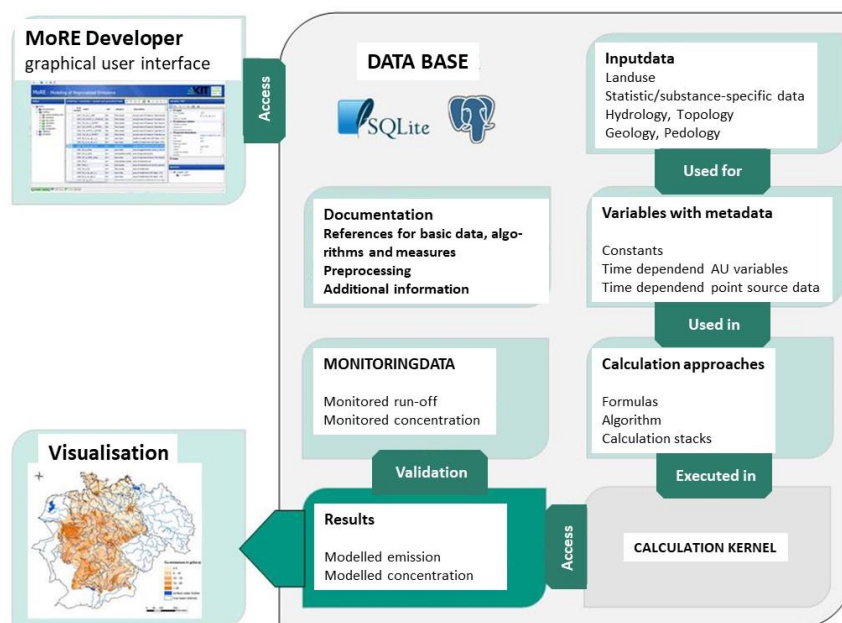


Figure 40: MoRE Model system, with PostgreSQL-Database and calculation engine (adapted from <https://isww.iwg.kit.edu/MoRE.php>).

The calculation procedure per time step and per sub-catchment proceeds in different steps:

1. The water balance is calculated based on information of the total run-off. This may be available e.g. in the form of measurements from gauge stations, provided via interpolation, by precipitation-runoff modelling, or top kriging approaches. The net area runoff of a sub-catchment is divided into different runoff components:
 - a. including wastewater treatment plant run-off (municipal, industrial and defined abandoned mining), combined sewer overflows and storm water runoff from separate sewer systems, collected sewer not transported to WWTP and runoff from non-urban traffic areas and
 - b. run-off components from unsealed surfaces, like precipitation (on surface water areas), surface run-off, run-off from drainages and groundwater (which includes base flow and

- subsurface flow); because up to now no clear (concentration driven) separation of groundwater and drainages is possible, both components were combined and represent the run-off from the underground (flowcharts available in the DHm³c project Output T2.2 *Report on system understanding*).
2. Releases of HS are quantified from point sources, like municipal WWTPs, industrial WWTPs and abandoned mining sites (with defined outlet). Furthermore, from urban diffuse pathways, like combined sewer overflows, storm water runoff, collected and untreated sewer. Finally, from non-urban diffuse pathways, like atmospheric deposition (direct emitting to surface water areas), surface run-off, erosion from arable land and from pastures, erosion from forests and mountainous areas and from the underground (flowcharts available in the DHm³c project Output T2.2 *Report on system understanding*).
 3. HS are traced through various pathways to surface waters; key “pathways” for the evaluated, dissolved substances in areas with increased anthropogenic use are:
 - a. Sewer systems collecting wastewater and storm water washed from impervious areas, which is discharged to surface waters after treatment or without any treatment;
 - b. Rainwater collection systems that discharge to surface waters with or without treatment;

In natural areas without intensive anthropogenic activities “key pathways” are:

 - c. Surface runoff and
 - d. Groundwater (Interflow and base flow).

In mining influenced pilot regions obviously:

 - e. Runoff from abandoned mining sites

Emission from each substance and all pathways are summed up (kg/a) and transported to the surface water as total load. Along a defined discharge tree, the discharges and the loads are summed up and the retention is calculated based on simple approaches, like “hydraulic load” and “specific run-off”, which consider settling of particles and subsequent retention in sediments. Concentrations are calculated from annual mean loads and annual discharge at the outlet of each sub-catchment (flowcharts available in the DHm³c project Output T2.1 *Harmonized MoRE model*).

Results are available for each sub-catchment and each year and can be aggregated e.g. to geographic or use-specific criteria, like upstream-natural regions and downstream-anthropogenic areas or to the complete pilot regions.

The calculation approaches implemented in the model differ with respect to complexity and data needs. While some approaches are reduced to simple load calculations based on multiplication of an “activity rate” (AR) and an “emission factor” (EF) (e.g. point source emission), others need a multitude of input data and calculation steps to achieve a final calculation. By way of example, in Figure 41 the calculation of erosion from arable land transported to surface waters is demonstrated, with a multitude of input data (e.g. land use data, crop specific information, rain intensity) formulas (e.g. the Revised Universal Soil Loss Equation (RUSLE) and Soil Delivery Ratio (SDR) calculation), algorithms (a sequence of formulas) and the final algorithm stack (a sequence of algorithms). The solids, which are finally traced to the surface water, will be multiplied by concentrations of soils from arable land (not shown here), which are specified by the Enrichment Ratio (a factor, which considers the sequencing of finer, heavily loaded particles and their increased share during transport).

A detailed documentation to setup the MoRE model is given in the handbook provided on the homepage of the original MoRE developer (https://more.iwg.kit.edu/wiki-en/index.php?title=MoRE_Developer).

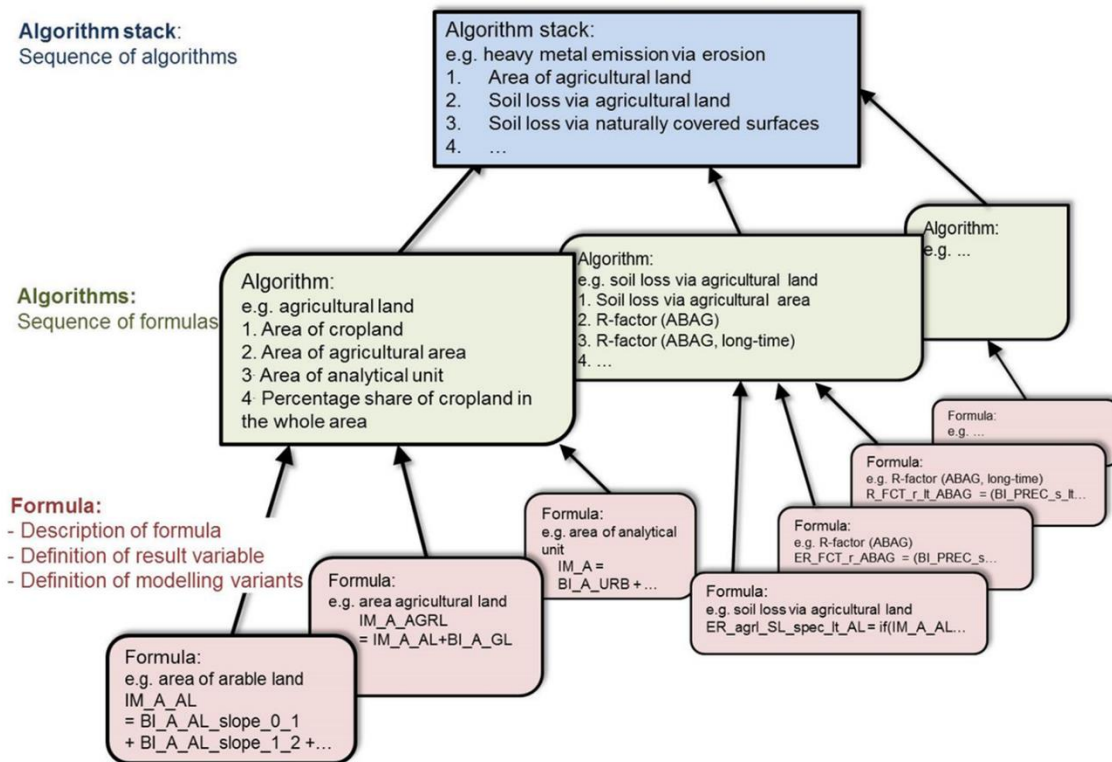


Figure 41: Emission calculation procedure for erosion from agricultural land (from https://more.iwg.kit.edu/wiki-en/index.php?title=MoRE_Developer#Design_of_the_MoRE_Developer_GUI)

4.2.4 Data requirements

Input data can be subdivided in:

- Basic input data.
- Substance-specific input data.

Basic input data (Table 10) includes all kind of background information on the physical characterization of each sub-catchment and are not substance concentrations or substance-specific turnover or removal rates. In general, these data represent GIS data (e.g. specific land use polygons or lines records), but also time series for precipitation or runoff available at specific locations can build these databases. In the latter case, punctual information often has to be interpolated by geo-statistical methods like kriging to produce a valid mean value for the analytical unit of the model: the sub-catchment.

Basic input data, which are immutable, are called “analytical units variables” (e.g. topography). Basic input data may also be subject to annual changes, like precipitation or soil loss and are called “periodical analytical units variables”. Some data are subject to serious changes only over a certain period of time - for example, the share of arable land. In such cases the model developer has to decide

in which category he puts the data. Once the model setup is prepared and an actualisation must be established (e.g. for the next policy cycle) only the “periodical analytical units variables” need to be updated. It should be also tested, if actualized and more precise input data sets were created.

Table 10: Basic input data used in MoRE. (x,t) = function of space and time; (x) = function of space; (c) = function of space using homogeneous values per country.

Actual input data code	Name	Description	Unit	Source
Analitical Unit (AU)	Topography/Area	Delineation of Analytical Units		
BI_A	Area	Area of analytical units	km ²	(x)
BI_ELEVA	Digital Elevation Model	Mean highs of subcatchments	m	(x)
Landuse	Landuse data set	Landuse categories in actual version	km²	
BI_A_AL_slope_0-1	Arable land	5 slope classes: 0-1; 1-2; 2-4; 4-8; >8 % (if available)	km ²	(x)/(x,t)
BI_A_PST	Pastures	Greenland, meadows	km ²	(x)/(x,t)
BI_A_WS_mr	Water surface	Main river (also lakes; reservoirs)	km ²	(x)/(x,t)
BI_A_WS_trib	Water surface	Tributaries (also lakes; reservoirs)	km ²	(x)/(x,t)
BI_A_FOR	Naturally covered areas	Woods; scrubland	km ²	(x)/(x,t)
BI_A_O	Open areas	Mountainous area without vegetation; beaches; dunes	km ²	(x)/(x,t)
BI_A_OPM	Surface mining	Mining areas	km ²	(x)/(x,t)
BI_A_URB	Settlements	Total urban areas	km ²	(x)/(x,t)
BI_A_IMP	Impervious urban area	Paved areas inside urban areas: settlements; industrial estates; car parks....	km ²	(x)/(x,t)
BI_A_WL	Wetlands	Area of Bog; swamp; floodplains	km ²	(x)/(x,t)
BI_A_OR	Country roads	Paved road area; not included in settlements	km ²	(x)/(x,t)
BI_A_REM	Other remaining areas	Other areas not listed above	km ²	(x)/(x,t)
Drainages	Melioration cadastre			
TD_SHR_a_td_agrl	Tile drained areas	From arable land and pastures	km ²	(x)
Meteorological Data	Climatic data			
AD_EVAPO_lt	Evapotranspiration	Longterm mean annual evapotranspiration	mm	(x)/(x,t)
BI_PREC_apr	Precipitation	Monthly values	mm	(x)
Hydrological data	River Discharges			
BI_Q_net	Net runoff	Modelling period; annual data	m ³ /s	(x)
Erosion	Soil loss			
ER_agrl_SL_spec_lt_AL	Soil loss	Soil loss from arable land (optional from 5 slope classes)	t/(ha·a)	(x)/(x,t)
ER_agrl_SL_spect_lt_PST	Soil loss	Soil loss from pastures	t/(ha·a)	(x)/(x,t)
Sewer sytem	Statistical Data about inhabitants and waste water system (partly from UWWTD)			
BI_INH	Number of inhabitants	Populaltion	inh	(x,t)
US_cso_VOL_spec_SOT	Stormwater overflow	Storage volume of stormwater overflow tanks in combined sewer systems, area-specific	m ³ /ha	(x)/(x,t)
US_L_CS	Combined sewers	Length of combined sewers	km	(x)/(x,t)
US_L_SS	Stormwater sewers	Length of stormwater sewers	km	(x)/(x,t)

Actual input data code	Name	Description	Unit	Source
US_SHR_inh_con_tot	Connection rate	Percentage of inhabitants that are connected to sewer systems	%	(x)/(x/t)
US_SHR_inh_conWWTP_tot	Connection rate	Percentage of inhabitants that are connected to sewer systems and waste water treatment plants	%	(x)/(x/t)
US_SHR_inh_nss_tot	Connection rate	Percentage of inhabitants that are not connected to sewer systems	%	(x)/(x/t)
US_INHC_H2O	Water consumption	Inhabitant specific water consumption	l/(inh·d)	
US_nss_SHR_inhl_towwtp_sept		Percentage of inhabitant load that is transported from septic tanks to waste water treatment plants	%	(x)/(x/t)
US_Q_spec_COM		Runoff rate for commercial waste water	l/(ha·s)	
Point source data (one value for each treatment plant)	Urban wastewater (partly from UWWTD)			
WWTP_ps_INH_conWWTP	Connection rate	Number of inhabitants that are connected to sewer systems and waste water treatment plants (point sources)	Inh	(x)/(x/t)
WWTP_ps_CP	Capacity	Capacity of the waste water treatment plant (point sources)	PE	(x)
WWTP_ps_PE	Load	Nominal load of waste water treatment plant (point sources)	PE	(x,t)
WWTP_ps_TS	Treatment type	Current treatment type of waste water treatment plant (point sources)	-	(x)/(x/t)
WWTP_ps_Q	Discharge	Runoff via waste water treatment plant (point sources)	m ³ /a	(x/t)
Industrial wastewater				
ID_ps_Q	Discharge	Runoff via industrial direct dischargers	m ³ /a	(x/t)

Substance specific input data are needed for each pathway (Table 1). For some pathways, concentration data are very sparse or almost non-existent. In such cases data from other pathways are used for a first estimation (e.g. groundwater for drainages or deposition for surface-runoff). Most substance specific data were obtained in the DHm³c project from the relational database (Chapter 3). The data are prepared based of tested datasets and created query routines, which prepare input data with different thematic references, such as:

- All data (data from different countries and different surveys).
- Country-specific data.
- Project specific data (e.g. data from DHm³c monitoring – total, country-specific or pilot region specific).
- Data, which refer to different land use classes.
- Data, which refer to different WWTP size classes or treatment.

To select the “right” data set, different criteria have to be considered, e.g.:

- Representativeness of data set (number)
- Actuality of data
- Spatial resolution

Of course, the best fit of the model results will be achieved using the most regional dataset. However, this approach contradicts the very utility of MoRE, a modelling approach that averages well to a larger spatial unit (e.g., country) and largely avoids specific model calibration using regional or local input

data. Consequently, in this project most often country specific data sets are used. If specific data are not available for a country or too uncertain because of a very low number, e.g. data from other countries are used.

The database additionally provides a statistical evaluation of the data (e.g. c10, c25, c50, c75 and c90 percentiles). The values are used to evaluate a best fit and gain information of possible over or underestimations on base of the available data set. Furthermore, it is used to calculate model variants to address the uncertainties of the approach:

- C50 = basis variant.
- C25 = minimum variant.
- C75 = maximum variant.

4.2.5 *Technical requirements*

The MoRE model is built on a PostgreSQL database in order to store the large datasets required for modelling and is freely available. Input data files and exports of results is prepared as Excel files. A full version is located on the TU Wien (Lead Partner of DHm³c) server and can only be viewed or used by assigning roles and rights. A slimmed-down model (stand-alone version) is provided, which only contains the model approaches used in the pilot regions. This reduced SQLite version is available as DHm³c project output T2.1 *Harmonized MoRE model adapted to specific territorial characteristics within the DRB*.

4.2.6 *Selection of examples of application and of results in pilot regions*

For different substance groups, like industrial chemicals, pharmaceuticals, heavy metals and pesticides the model was applied in seven pilot regions with 34 sub-catchments. The model period runs from 2016-2021 (annual time steps). A comparison of monitored mean annual loads and concentration with model results was performed for 2021 at 20 monitoring stations situated at catchment or sub-catchment outlets.

The model performance shows significant differences. In most cases, good model fits can be achieved (e.g. for PFOS, PFAS and some metals like Cu or Cd). The good performance is a result of combining the improved MoRE model and the new database providing several extensive substance-specific datasets for a number of substances (chapter 3).

For other substances the model performance is weaker (e.g. lead or carbamazepine). Even so, the model reproduces the monitored data to a high extent, it produces some slight systematic underestimations in loads (e.g. Ni, Cu, As, PFOS) and slight overestimations (e.g. Cr, Zn).

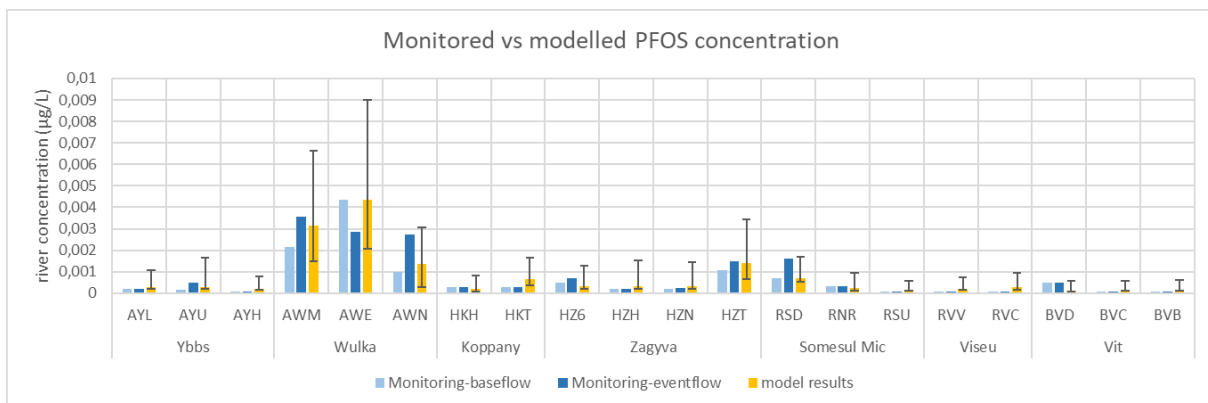
In some cases, one or two catchments out of 20 cannot be modelled sufficiently well. This mostly refers to pilot regions where very specific and extreme conditions prevail. In the Viseu catchment, influenced by abandoned mining, a specific data set needs to be used for several heavy metals to achieve at least approximately reasonable loads and concentration. For some metals, even this dataset developed still leads to serious underestimations. Other problems could be identified especially in pilot regions with

extreme 2021 hydrological conditions, like Ybbs (extreme event in July 2021) where huge amounts of load were transported during three days but also in the Pannonian pilot regions, which faced during the project implementation period an extreme dry weather.

Model results were calculated with base, maximum and minimum variants to address uncertainties. Monitored mean concentration and load calculations are also subject of serious uncertainties. These were addressed by considering monitoring results from base flow conditions and from base flow and event flow results. The first approach was assumed to represent the standardized national monitoring strategies with probing from low flow to well above mean flow conditions but without representative probing of event flows, following a given temporal routine. The second one, including event flow (often starting with the statistical ten percent of the highest discharges from a long-term runoff series) was assumed to be much closer to realistic conditions. The model comparison orientates to the latter. However, if you use model results e.g. for a risk analyses, the comparison to data sets underestimating event flows and which are related to more average conditions are the much more “realistic” ones.

In Figure 42 the comparison of monitored mean annual concentration for PFOS and PFOA calculated from two datasets (baseflow and baseflow including event-flow) and of model results is presented.

The graphs express the model performance and the range of different concentrations with maximum concentration in the Wulka pilot and in Zagyva (upstream) as well as in the outlet catchment of Somesul Mic. All sub-catchments with increased concentrations are subject to high shares of emission from treated urban wastewater and from sewer systems (Figure 44, which lists the percentage shares of the input paths in the pilot regions).



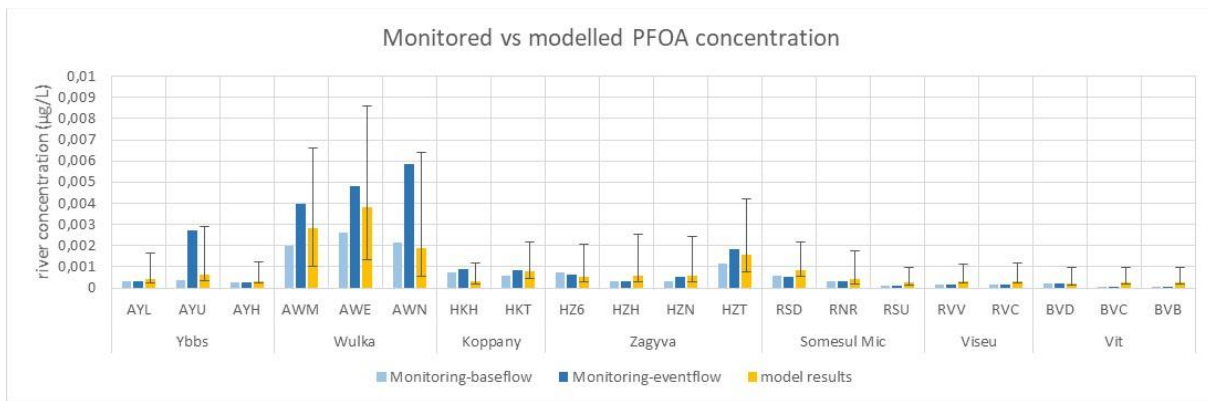


Figure 42: Monitored and modelled concentrations of PFOS and PFOA; monitored concentrations from baseflow conditions and baseflow+eventflow conditions; modelled concentration in three variants: base (column), min. and max. (as bars).

The results demonstrate the relevance of the monitoring approach, which leads to significant differences. The more time representative approach, which considers only “normal” baseflow compared to the more flow representative approach often results in lower concentrations (although a high proportion of load comes from point sources). It becomes clear that the significant higher concentrations calculated under flow representative conditions in some sub-catchments are underestimated by the model (base variant) and can in some cases be only reproduced by the maximum model variant. This becomes even clearer in the following scatter plots (Figure 43).

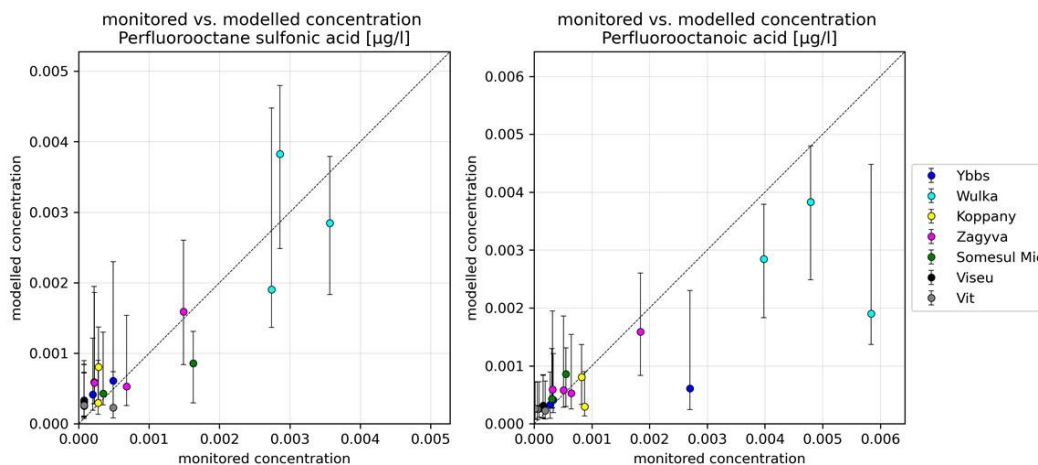


Figure 43: Monitored (baseflow+eventflow) and modelled concentration of PFOS and PFOA [µg/l] at 20 monitoring stations from seven pilot regions.

Pilot regions dominated by often steep natural and mountainous areas (Ybbs, Viseu, Vit) show increased shares from surface runoff (mainly PFOS) and groundwater (mainly PFOA). In pilot regions with higher shares of urban areas emission from municipal wastewater and also from combined sewer systems are of increasing importance, with a clear dominance of wastewater emission in Wulka and Somesul Mic (Figure 44).

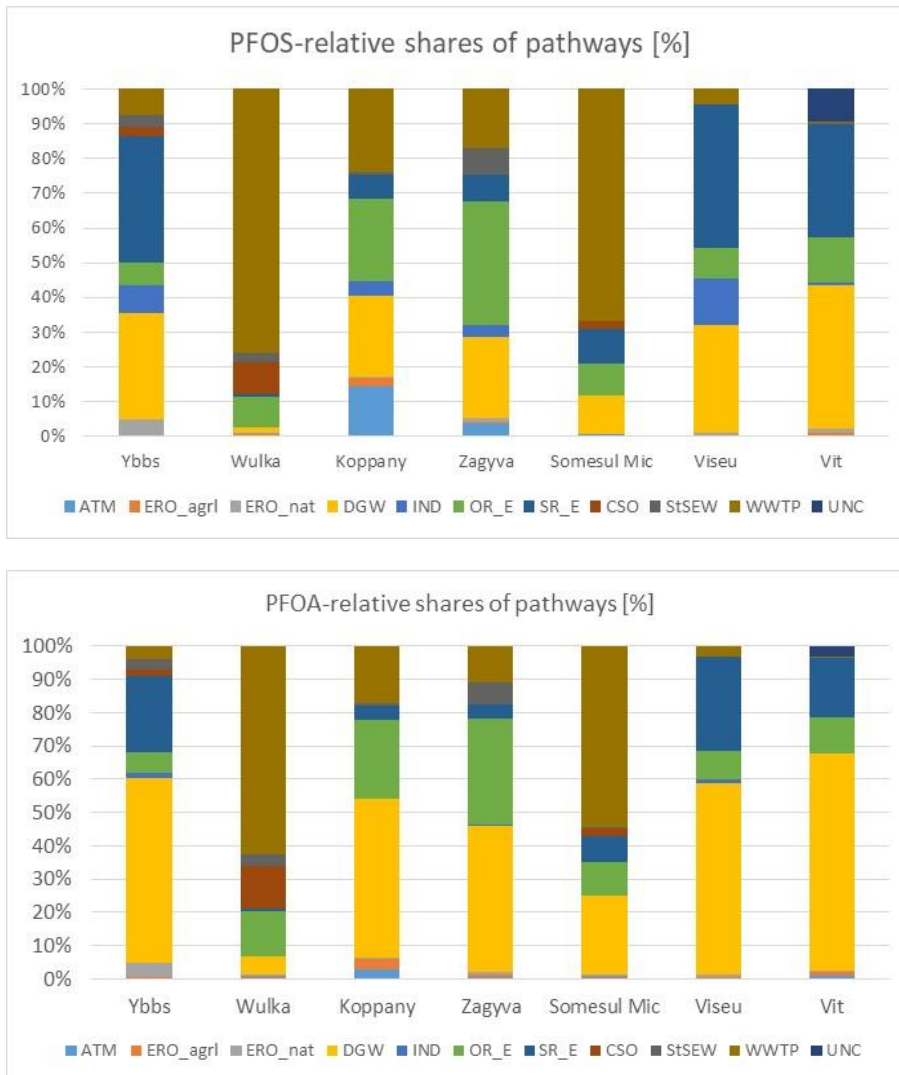


Figure 44: PFOS and PFOA shares of modelled pathways [%] in seven pilot regions; the explanation of abbreviations is provided in Table 13.

Increased area specific rates of PFOS and PFOA [$\text{g}/\text{ha} \cdot \text{y}^{-1}$] can be found in Austrian and Romanian pilot catchments with around 0,003 [$\text{g}/\text{ha} \cdot \text{y}^{-1}$]. Figure 45 shows selected scatter plots illustrating the model performance for the estimation of loads for the year 2021 in the pilot regions. Substances directly influenced by mining in the Viseu catchment (mainly cadmium, copper and zinc) show extreme high loads, and can only be reproduced approximately. Cadmium loads in the direct mining influenced sub-catchment for example are 50-fold higher than the highest loads shown here.

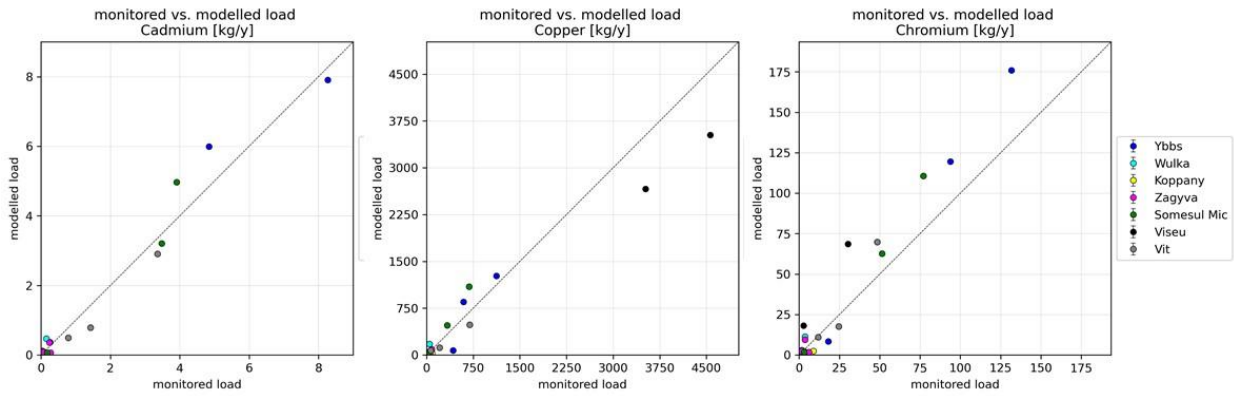
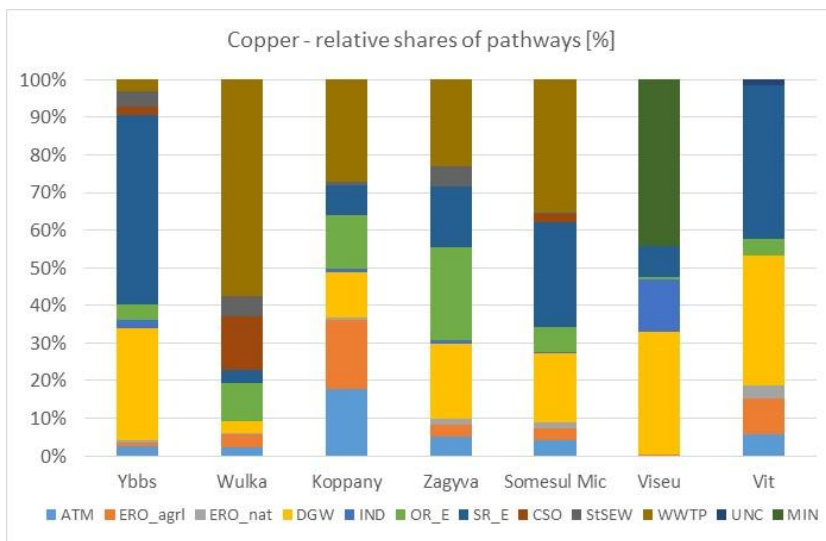
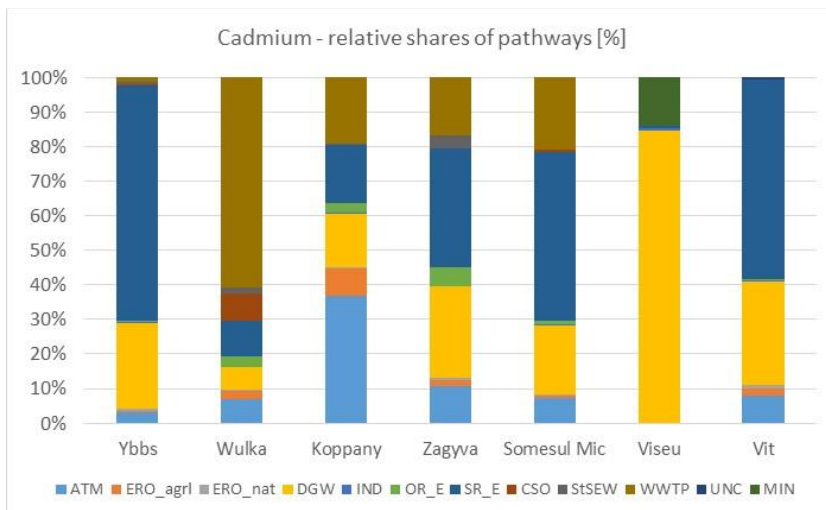


Figure 45: Monitored and modelled loads of dissolved Cd (exclusive results from Viseu), Cu and Cr [kg/y] at 20 monitoring stations from seven pilot regions.

The share of pathways of cadmium, copper and chrome (Figure 46) exemplarily demonstrates the influence of direct discharges of untreated abandoned mining effluent, which is dominant for copper and even for zinc in Viseu pilot (not shown here).



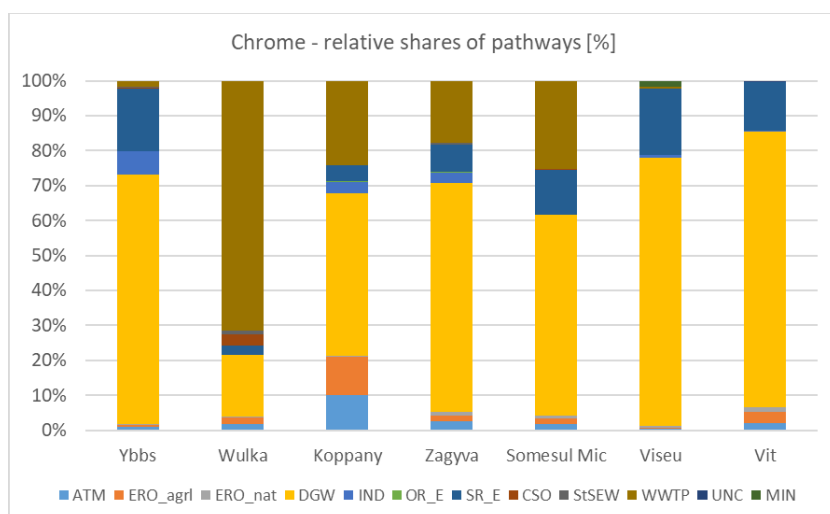


Figure 46: Dissolved cadmium, copper and Chrome shares of modelled pathways [%] in seven pilot regions; the explanation of abbreviations for pathways is given in table Table 13.

The results illustrate the differences of the pilot regions with predominantly natural or with more urban characteristics. In Koppány, a high share of ditches and small lakes leads to a high area of surface waters, which explains the high proportion of discharges from atmospheric depositions. In the Wulka catchment the 2021 year was extremely dry. This is one reason for the very low share of emission from drainages and groundwater and high shares from treated municipal wastewater.

Within the pilot regions, also the spatial differences can be addressed showing results for each sub-catchment. As an example the Ybbs catchment is chosen, characterised by a natural pre-alpine upstream region (11008 – 11004) and increasing urban areas, agriculture and anthropogenic activities in the downstream regions. It is obvious that the dominant inputs for copper stem from natural background, emitted via the groundwater (interflow and baseflow) and surface runoff from the steep areas (Figure 47).

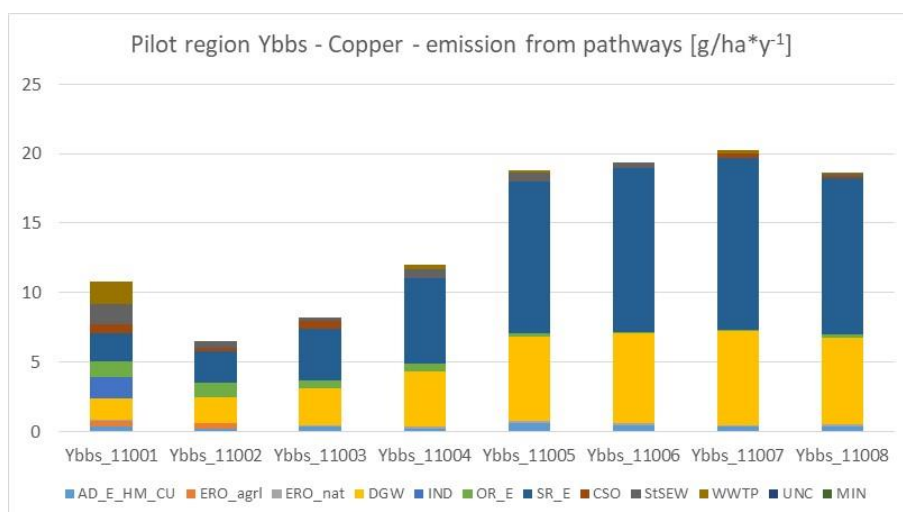


Figure 47: Dissolved copper in [g/ha *y⁻¹] in all sub-catchments of the Ybbs-pilot region; the explanation of abbreviations for pathways is given in Table 13.

From the presentation of all sub-catchments in the Viseu pilot region, the influence of direct emission from abandoned mining for copper becomes obvious (Figure 48). Furthermore, it becomes clear, that even the upstream sub-catchment 32002, which is in rather the same geological formation as 32003 emits high shares of copper via groundwater (baseflow and interflow).

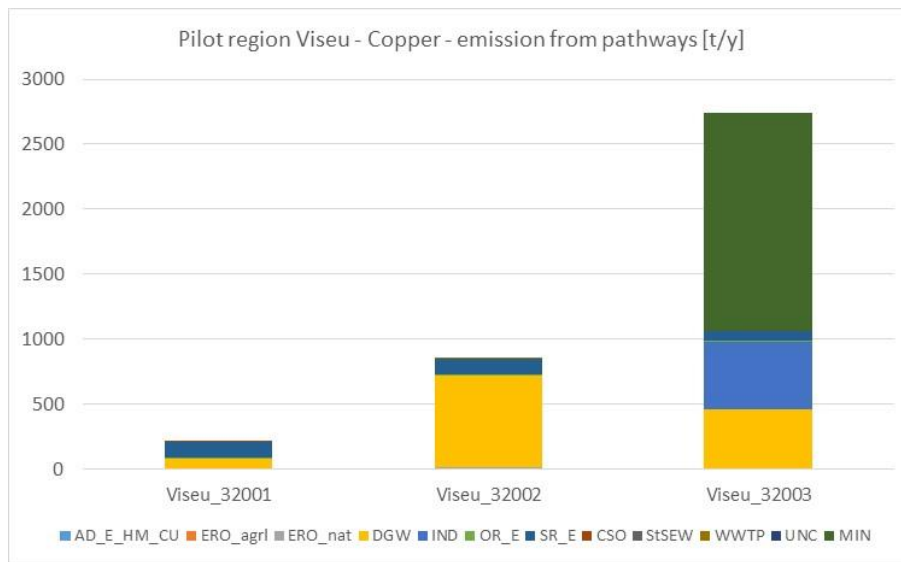


Figure 48: Dissolved copper in [t/y] in all sub-catchments of the Viseu-pilot region; the explanation of abbreviations for pathways is given in Table 13.

Of course, model results can also be presented as GIS maps (for visualizing important messages), e.g. with presenting area specific rates or combine this presentation with statistical evaluations of pathways for each sub catchment; or to visualize the spatial distribution of specific pathways.

More and detailed results can be find in the DHm³c project Outputs *O.T2.2 Report on improved system understanding* and *O.T2.3 Demonstration of management plan development in pilots*.

4.3 Danube Hazard Substance Model

4.3.1 Source-oriented approach

In Danube Hazard m³c, a Tier 4 source-oriented emission inventory was compiled based on the Danube Hazardous Substances Model (DHSM). The approach adopted relies on the Technical Guidance (Figure 36). It further makes use of a recent report issued by the European Environment Agency, which provides supplementary advice and aims to provide practically applicable methods, and to contribute to the harmonisation of the methods used across Europe (Van den Roovaart et al, 2022). The source-oriented approach provides the best possible insight in the origin of the emissions, and therefore produces essential information for the definition of emission reduction measures.

The DHSM is a further development of an emissions, fate and transport model developed in the EU R&D project SOLUTIONS (<https://www.solutions-project.eu/>; <https://enveurope.springeropen.com/articles/10.1186/s12302-019-0248-3>).

The SOLUTIONS model aimed to be applicable to as many chemicals as possible on the scale of the European continent. It therefore had to rely on a relatively simple emission estimation approach and on restricted input data. Within Danube Hazard m³c, the SOLUTIONS model was modified to include additional substances relevant for the DRB (metals). At the same time, the number of modelled substances was drastically reduced. This allowed a more comprehensive emission modelling approach with higher data needs, reflecting deeper system understanding and better data availability than what was feasible within the SOLUTIONS project.

The sources included in DHSM are those listed in Figure 36. For man-made chemicals, the source-oriented method starts from the production of the chemical. Losses to the environment (sources) can be associated to various life-cycle stages, see Figure 49. This includes losses during production and industrial use, from consumptive use, from stocks in products, buildings, infrastructure, and from waste management. Consequently, the sources distinguished in the modelling can be from any of these life cycle stages.

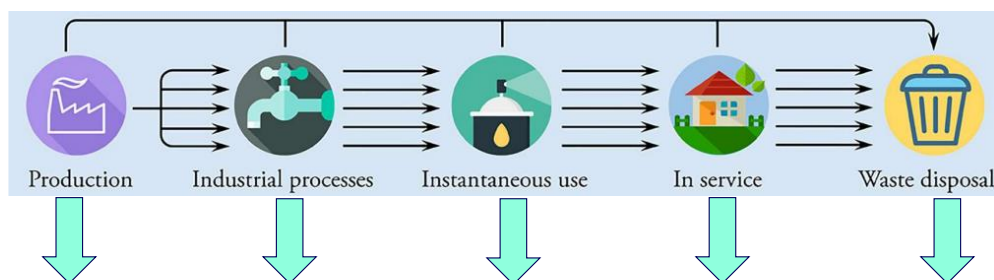


Figure 49: Symbolic representation of sources of HS generated during various life cycle stages.

The Technical Guidance considers atmospheric deposition as a source, though actually it is a pathway. It is the result of sources towards the atmosphere being transported to the aquatic and terrestrial environment. To consider atmospheric deposition a pathway is a practical solution, also because atmospheric transports often involve larger spatial scales that exceed river basin boundaries. As it is a pathway, the interpretation towards the underlying true sources needs to be provided by expert judgement.

The build-up of stocks in the techno-sphere may take a longer period after the start of the use of a chemical. Depending on the service-lifetime of the products or constructions involved, this could span a period of decades or more. To avoid such timescales in the simulation of emissions, the present stocks of HS are seen as the “activity” contributing to the input of HS into the environment. For natural HS (typically metals) a similar consideration holds for their natural presence and/or build-up of stocks in the soils. The current stock of metals in soils is seen as the “activity” contributing to the input of HS into the environment.

4.3.2 *Scope, spatial and temporal scales*

The DHSM was implemented for a selection of hazardous substances. This reflects the target chemicals of the Danube Hazard m³c project, selected to fulfil various criteria: (1) substances representing relevant sources and pathways, (2) substances relevant for ICPDR, national and regional authorities in the basin, (3) substances that can be actually detected and measured, so that data can be expected to be available.

Because the substances selected for modelling with DHSM represent relevant sources and pathways, their emission inventories are expected to be representative for a much wider group of substances, and to provide insights for defining effective pollution control measures beyond the modelled substances alone.

The DHSM was implemented for the following 17 chemicals:

- Metals: arsenic, cadmium, copper, nickel, lead, zinc and mercury.
- Benzo[*a*]pyrene as a representative of the group of Polycyclic Aromatic Hydrocarbons.
- Pharmaceuticals: diclofenac, an anti-inflammatory drug and painkiller, and carbamazepine, an anticonvulsant used in the treatment of epilepsy and neuropathic pain.
- Industrial chemicals with wide dispersive use: 4-tert-octylphenol, nonylphenol, bisphenol-A.
- Per- and polyfluoroalkyl substances: PFOS and PFOA.
- Pesticides: tebuconazole a fungicide used for wood preservation, and metolachlor, a herbicide in agriculture.

Water flows are important for quantifying emissions. The DHSM therefore relies on input from a hydrological model that quantifies runoff, sub-surface flows and river flows. The DHSM uses output from the pan-European E-Hype hydrology model (Hundecha et al., 2016), and therefore adopts the spatial schematization. For the DRB, this schematization exists of 3,523 elements with an average area of 229 km². Figure 50 shows the schematization of the DHSM, together with the DRB delineation, the rivers with a catchment area exceeding 4,000 km² and the lakes with a surface exceeding 100 km², as identified by ICPDR. It also shows the 7 pilot regions, in which targeted measurement campaigns have been carried out to fill critical data gaps needed to provide a robust basis for modelling and management. These pilot regions were also modelled using the MoRE model (Section 4.2). The information derived from these pilot regions has been used to support the implementation of the DHSM.

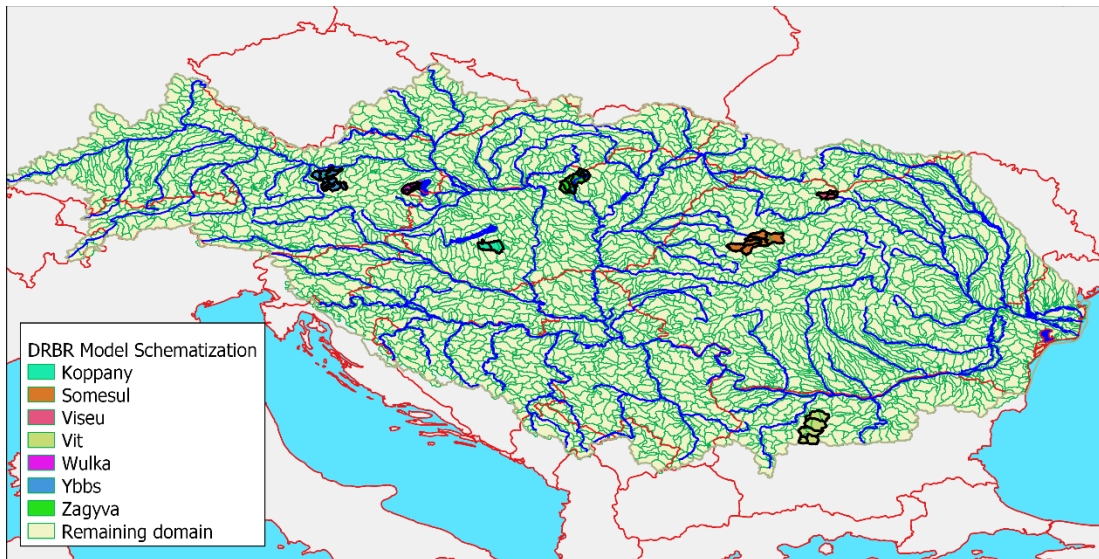


Figure 50: Overview of DHSM schematization, main rivers, lakes and pilot regions.

DHSM operates with a time step of 1 day. It uses a time dependent representation of some sources (e.g. atmospheric deposition) and especially of the pathways affected by rainfall, runoff and subsurface flow. This supports the representation of pathways that respond to weather events. It also supports the representation of interannual variability as controlled by climate variability. The DHSM is therefore set up for a sequence of 10 hydrological years, so that the impact of climate variability can be quantified.

4.3.3 Main calculation approaches

The Danube Hazardous Substances Model uses the flow-chart shown in Figure 51 for its emission calculations.

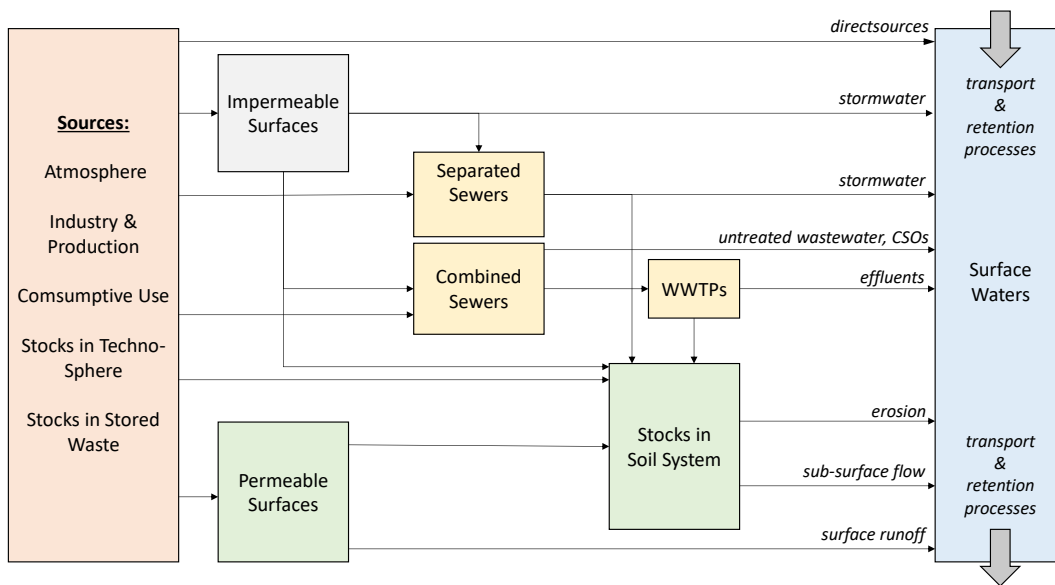


Figure 51: Set-up of the DHSM model system.

The calculation procedure per time step and per schematization element proceeds in steps:

1. Releases of HS (“sources”) are quantified from atmospheric deposition, agriculture practices, road traffic, buildings and constructions, households, industry and navigation. This is done by multiplication of an “activity rate” (AR) and an “emission factor” (EF), that can both optionally be a function of time and space.
2. Releases of HS are allocated to one or more initial receptors: surface waters, impermeable or permeable surfaces, wastewater or the soil system.
3. The soil system is given an initial concentration that represents natural presence or the build-up of stocks from historical releases.
4. HS are traced through various compartments in the techno-sphere and the environment on their way to surface waters; key “pathways” are:
 - a. Sewer systems collecting wastewater and stormwater washed from impervious areas, that is discharged to surface waters after treatment or without any treatment;
 - b. Rainwater collection systems that discharge to surface waters with or without some degree of treatment;
 - c. The pathways through the soil system, controlled by particle erosion and delivery to streams, by surface runoff and by subsurface flow.

The end results of these calculations are emissions in g/day for all schematization elements for 10 consecutive years. These results can be aggregated in space and time to provide long term annual averages and their inter-annual variability, basin-totals, country totals, sub-basin totals, etcetera.

In a second step, the calculated emissions are input to a water quality model. This model uses the in-stream water discharges from the hydrology model and the previously calculated emissions to calculate time- and space-dependent concentrations. The water quality model includes in-stream retention by various degradation processes and by settling and subsequent retention in sediments.

The calculated concentrations were used for model validation. They can be processed into annually averaged concentrations and the inter-annual variability thereof.

4.3.4 Data requirements

The data requirements of DHSM are summarized in Table 11 and Table 12. These tables indicate to what extent these data are space- or time-dependent. This reflects the DHSM implementation available at the end of the Danube Hazard m³c project. Table 11 defines the activity rate and emission factor used to quantify the various sources, whereas Table 12 lists the other input data, which are used to represent the different pathways in the model.

Table 11: Data used in DHSM to quantify sources. (x,t) = function of space and time; (x) = function of space; (c) = function of space using homogeneous values per country; () = constant value.

Sources	Activity Rate (AR)	kind	Emission Factor	kind	Data source
Atmospheric deposition	rainfall	(x,t)	concentration	(x)	EMEP model for Cd, Hg, Pb, BaP
Agriculture	agriculture area	(x)	concentration	(-)	DHm3c database for other HS
			release per unit area	(c)	literature for metals
Road traffic	road length	(x)	release per unit area	(x)	global gridded data for pesticides
			release per unit road length	(c)	global roads database, various literature sources
Buildings & constructions	population	(x)	per capita release	()	literature and DH m3c database
Households	population	(x)	per capita release	() or (x)	DH m3c database
Industry	release	(x)	1	()	E-PRTR database
Navigation	navigable waterways length	(x)	release per km waterway	()	literature

Table 12: Other data used in DHSM. (x,t) = function of space and time; (x) = function of space; (c) = function of space using homogeneous values per country.

Input	kind	source	remarks
soil concentrations	(c)	DH m3c database, Foregs database	

population	(x)	gridded population map	scaled to match reported country population numbers
share of urban population	(x)	gridded settlement type map	
Hydrology (rainfall, surface runoff from impervious and pervious surfaces, infiltration, exfiltration, total runoff)	(x,t)	E-Hype model, daily interval, 2003-2013	
water surface	(x)	E-Hype model	to maintain consistency
impermeable surfaces	(x)	E-Hype model	to maintain consistency
permeable surfaces	(x)	E-Hype model	to maintain consistency
agriculture area	(x)	E-Hype model	to maintain consistency
connection rate to sewers	(x)	ICPDR UWWTD inventory	
connection rate to IAS/LS ¹⁴	(x)	ICPDR UWWTD inventory	
rate not connected to sewers or IAS/LS	(x)		implicitly defined
fraction primary treatment	(x)	ICPDR UWWTD inventory	
fraction secondary/tertiary treatment	(x)	ICPDR UWWTD inventory	
fraction advanced treatment	(x)	ICPDR UWWTD inventory	
fraction not treated	(x)		implicitly defined
fraction of combined sewers	(c)	literature	
fraction of sewage sludge re-used	(c)	literature	
sediment delivered to streams	(x,t)	unpublished data from EU project	distributed in time using hydrology data
rainfall threshold for occurrence of CSOs	(x)	assumed CSO volume 5%	Local threshold derived from hydrology data (rainfall)
share of stormwater centrally collected and discharged to WWTP/surface water	(x)	assumed equal to share of urban population	

In addition, there are some model parameters, often substance specific, including the removal by various treatment types, removal in soils and surface waters and soil and surface water partition coefficients.

4.3.5 Technical requirements

The technical requirements for the DHSM are a Windows based PC (OS Windows 10, 8GB RAM, 20-40 GB storage). The DHSM does not provide any supportive tools to prepare input data or present output data. Input data processing, including GIS operations, is left to the user. Output data can be extracted in comma-separated values text format, for further processing by the user.

¹⁴ IAS = Individual and other Appropriate Systems as defined in the Urban Wastewater Treatment Directive (UWWTD); LS = local systems as defined in the 2021 Danube River Basin Management Plan.

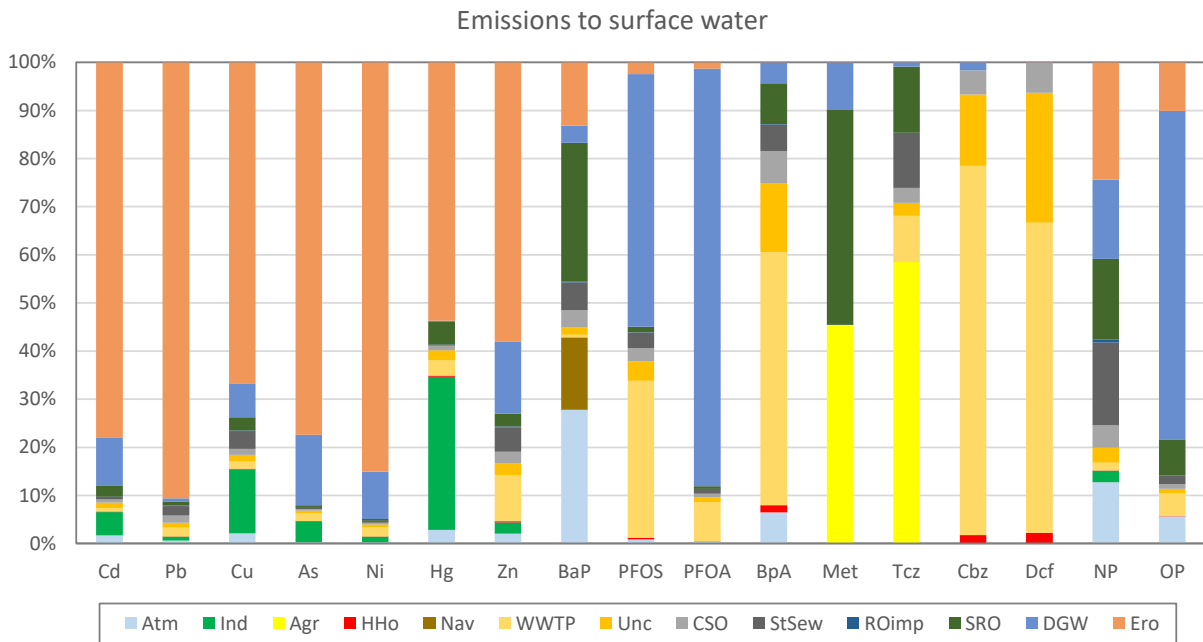
4.3.6 Selection of examples of application and of results at DRB scale

The breakdown of the basin-wide calculated annual emissions to surface waters of 17 HS according to their pathways is shown in Figure 52. The pathways shown in this plot are listed in Table 13. The second column in this table refers to the legend in Figure 52. The third column links to the pathways as they are numbered in the Guidance (Figure 36).

The results show strong differences between the HS. For some metals (Pb, Ni), erosion is the dominant pathway. For the pharmaceuticals (Cbz, Dcf), treated and untreated wastewater are the dominant pathways. For pesticides (Met, Tcz), direct emissions from use in agriculture is a significant pathway. For PFOS, PFOA and OP, subsurface flow is the dominant pathway. Other substances (BaP, BpA, NP) show a mix of different pathways. It is noted that the discharges to surface waters from (abandoned and historic) mining (Pathway P11 in the following table) could not be quantified on the basin scale and therefore had to be left out of our inventory.

Table 13: Overview of pathways included in the emission inventory.

Description	Abbrev.	Pathway	Cluster
Atmospheric deposition directly to surface waters	Atm	P1	Direct releases
Erosion	Ero	P2	Erosion - subsurface flow
Surface runoff from permeable surfaces	SRO	P3	Surface runoff
Drainage + Groundwater (subsurface flow)	DGW	P4	Erosion - subsurface flow
Direct emissions to surface waters from agriculture (spray drift)	Agr	P5	Direct releases
Runoff from impermeable surfaces to surface waters (outside urban areas)	ROimp	P6	Stormwater and CSOs
Emissions from wastewater collected but not connected to a WWTP	Unc	P7	Wastewater
Combined sewer overflows	CSO	P7	Stormwater and CSOs
Emissions from separate storm water collection systems (in urban areas)	StSew	P7	Stormwater and CSOs
Emissions from treated domestic wastewater collected in sewers	WWTP	P8	Wastewater
Direct emissions to surface waters from unconnected households	HHo	P9	Direct releases
Discharges to surface waters from industry	Ind	P10	Direct releases
Discharges to surface waters from (operational and abandoned) mining	Min	P11	Direct releases
Direct emissions to surface waters from navigation	Nav	P12	Direct releases



Cd, Pb, Cu, As, Ni, Hg, Zn are metals; BaP = benzo[*a*]pyrene; Met = metolachlor, Tcz = tebuconazole; Cbz = carbamazepine, Dcf = diclofenac; NP= nonylphenol, OP= octylphenol, BpA = bisphenol A.

Figure 52: Relative distribution of emissions to surface waters over pathways for the simulated HS.

The DHSM makes an inventory of HS flows through the various model compartments. An example is shown in Figure 53. This chart shows for all substances the relative importance of the incoming and outgoing mass flows for the combined sewers and WWTPs. On the inflow side, combined sewer systems receive a mix of domestic wastewater, industrial wastewater and stormwater. The relative contribution of these three contributions varies per substance. On the outflow side, combined sewer overflows, unconnected sewer systems, re-use of sludge and effluents are quantified. The balance is closed by the removal in WWTPs (either true degradation or removal by incineration or isolation of sludge). Similar results are available for the other modelled compartments. Reference is made to the Danube River Basin Scale Assessment Report for more results.

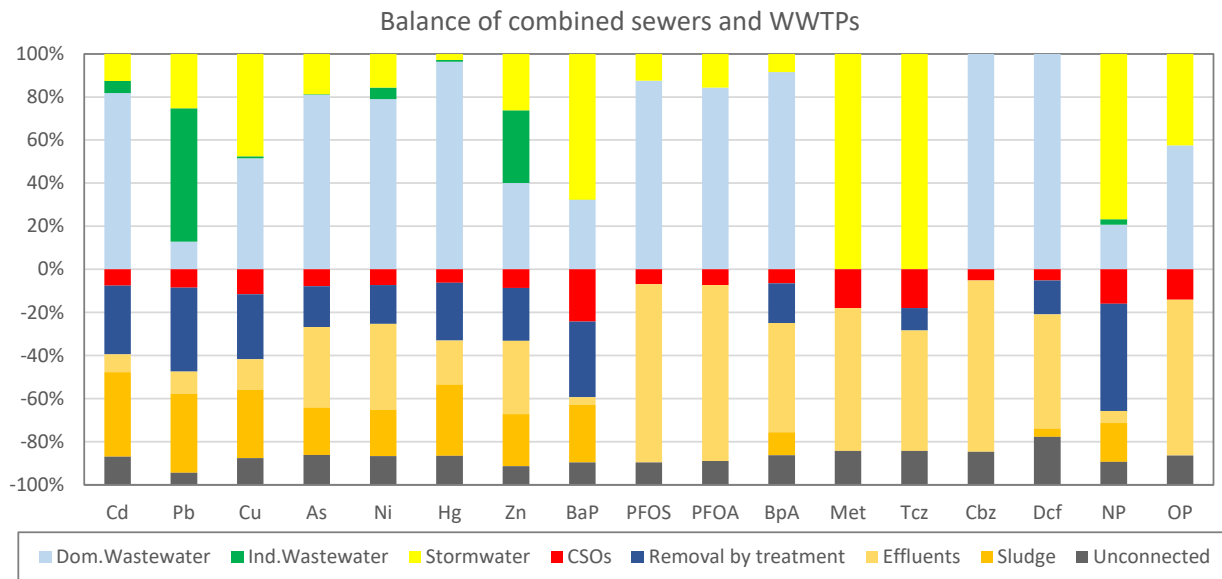
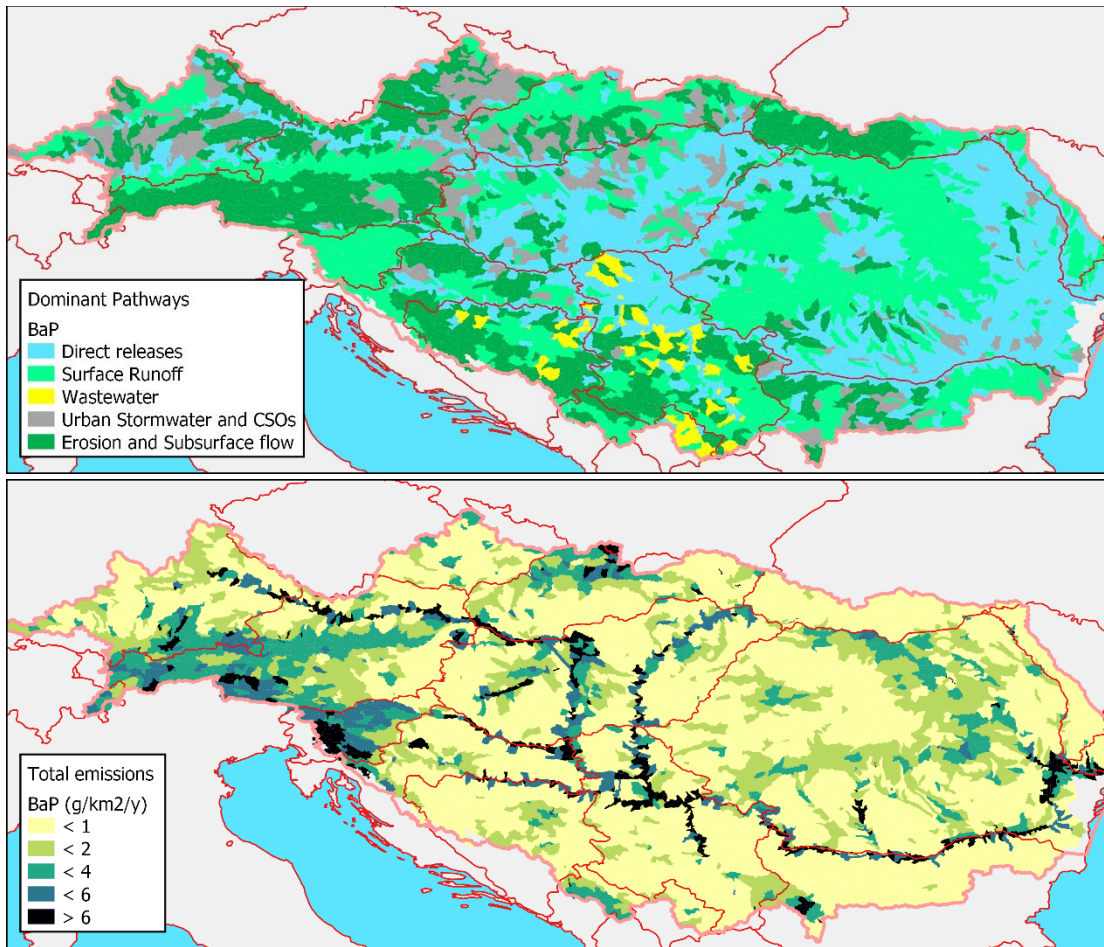


Figure 53: Figure 4 18: Basin-wide mass balance of the combined sewers and WWTPs compartments.

Because of the detailed spatial schematization, all results can also be differentiated in space. This is illustrated in Figure 54 for benzo[*a*]pyrene. The upper panel shows a map of the total area-specific annual emissions. The middle panel shows a map of the dominant pathway cluster (clusters are listed in the fourth column of Table 13). These two maps illustrate the strong spatial inhomogeneity of the total emissions and of the distribution over the pathways. The bottom panel shows a disaggregation of the emissions and their distribution over the pathways per country. The lower panel shows relatively high emissions in Slovenia. This is the result of (a) the locally high atmospheric deposition (not shown here), (b) the local orography and climate conditions that lead to relatively high surface runoff. Similar results are available for the other 16 simulated HS.



Emissions to surface water per country for BaP (g/km²/y)

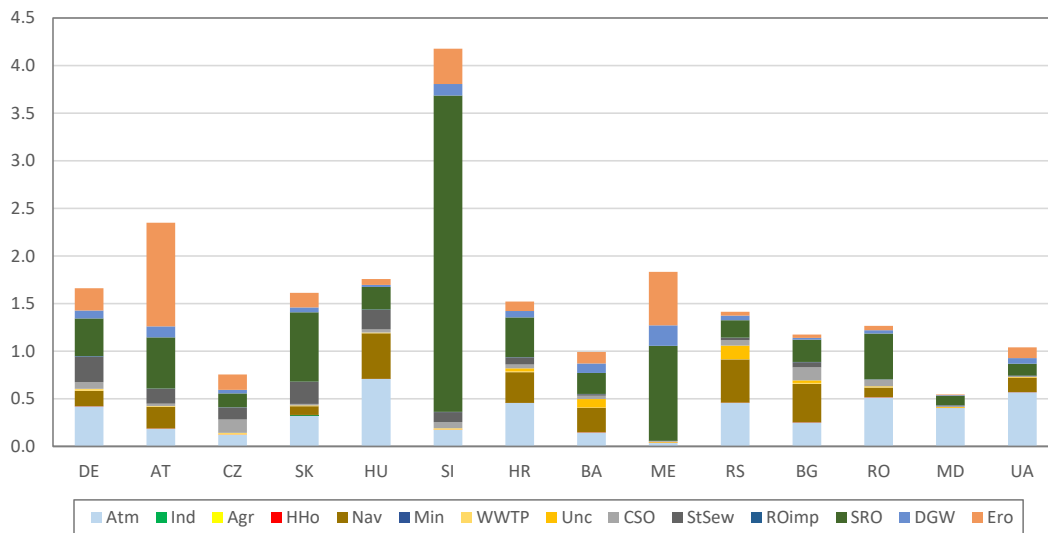


Figure 54: Spatial variability of emissions and dominant pathway: benzo[a]pyrene.

The calculated emissions, together with the river discharges and some retention processes lead to calculated in-stream concentrations. Figure 55 shows the simulated long-term average concentration of PFOS as a result of present emissions. These results were compared to observed concentrations from the Transnational Monitoring Network, from the Joint Danube Survey⁴ and from supplementary monitoring conducted in Danube Hazard m³c at 6 stations along the main branch of the Danube. Figure 56 shows this comparison for PFOS. For the other simulated HS, similar results are presented in the Danube Hazard m3c Danube River Basin Scale Assessment Report.



Figure 55: Simulated long term average concentration of PFOS as a result of present emissions.

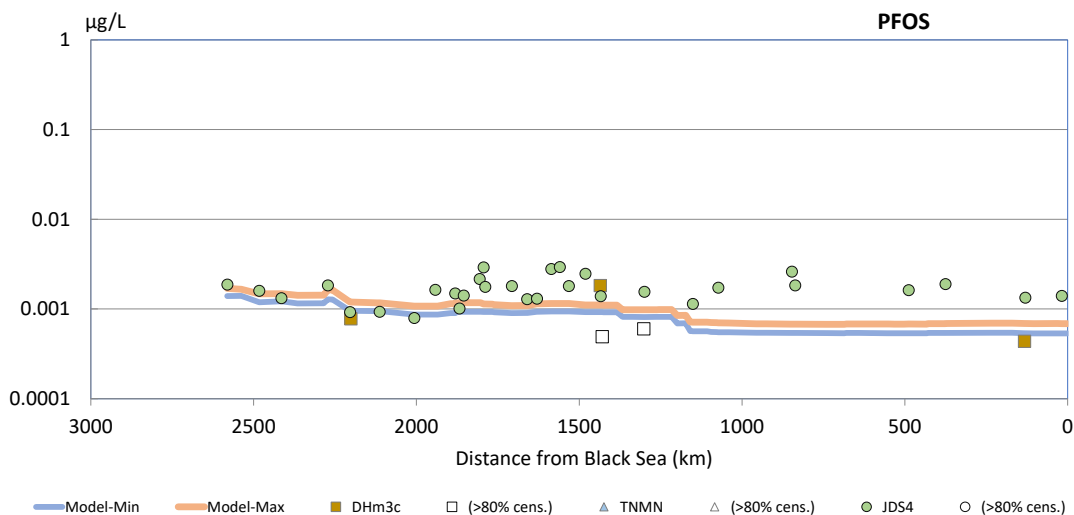


Figure 56: Comparison of simulated concentrations and observations from different datasets, plotted along the length of the Danube River for PFOS.

4.3.7 Interpretation of results at DRB scale

When using and interpreting the results of DHSM (or any similar model for that matter), it is important to realize a few things.

The model can only be as good as its input data are. As the amount of input is significant, there is always a degree of uncertainty. In the Danube River Basin Scale Assessment Report, the order of uncertainty has been provisionally quantified by providing a “low” and a “high” emission inventory next to the most realistic estimate. This uncertainty is significant, and for most substances higher than the year-to-year variability that originates from climate-induced variability caused by different weather patterns that affect rainfall, runoff and river discharges.

The uncertainty increases with decreasing spatial scale. Many spatially variable input quantities could only be quantified on a country-by-country basis. This implies that the results for the main branch of the Danube and for the large international tributaries like Sava and Tisa are probably more robust than those for the smaller tributaries. Very small rivers are not even resolved by the DHSM, while the schematization elements have an average surface area of about 200 km².

The DHSM strictly satisfy mass balance rules and provides full insight in such mass balances. This helps understanding phenomena shown in the output. Based on system knowledge, DHSM users can decide if such phenomena are artefacts of uncertain input data or useful management insights.

Concentrations calculated by DHSM are true time averages and as such they reflect the annual mean concentrations that are regulated by EQS under the WFD. True time averaged concentrations may differ from the quotient of the annual emissions and the mean river discharge. The difference is the result of the time variability of the emissions in the watershed and the river discharges. Present DHSM results in the Arges River (Romania) for example show that the true annual mean concentration is 10-20% higher than the quotient of the annual emissions and the mean river discharge. In the Iskar River (Bulgaria), this difference is 50-100%.

4.4 Implementation of scenarios to mitigate HS pollution

4.4.1 Pollution control measures

MoRE

In the MoRE model different management measures can be implemented after the model setup and validation is finished. Management measures can only be implemented if the selected measure can be related directly to the pathways modelled in MoRE. The following table (Table 14) describes the most common management measures implemented in this model.

Table 14: Implementation of pollution control measures in MoRE.

Measure	Conceptual approach	Data requirements
<i>Wastewater management</i>		
Construction of sewer systems	Increase share of population connected to sewers and decrease share of population connected to IAS/LS	Estimate reduction from case studies, literature research or stakeholder involvement
Construction of conventional WWTPs	Change share of collected wastewater undergoing treatment	Estimate reduction from case studies, literature research or stakeholder involvement
Implementation of advanced wastewater treatment	Change share of collected wastewater undergoing advanced treatment	Estimate reduction from case studies, literature research or stakeholder involvement
<i>Stormwater management</i>		
Decoupling of stormwater collection systems	Decrease of share of combined sewers and increase share of storm sewers	Estimate of changes in each pilot area
Increased storage in combined sewers	Reduction of suspended particulate matter emissions, adapt K_d value for micropollutants	Percentage of reduction of suspended particulate matter emissions (i.e. CSO=30%, storm sewers= 20%), adapted k_d values
<i>Air pollution control</i>		
reduction of atmospheric deposition	Reduction of atmospheric deposition by a certain percentage	Estimate reduction from case studies, literature research or stakeholder involvement
<i>Erosion</i>		
Measures to reduce erosion	Direct reduction of erosion pathway	Estimate reduction from case studies, literature research or stakeholder involvement

Technical implementation: There are two main possibilities to calculate management measures in the MoRE model. The first option is to reduce the emissions from a specific pathway for a substance by a specific percentage. The other option is to calculate the measures outside of the MoRE model and alter the input data directly. As this second option does not require any changes into the model, but only in the input files, this option will not further discussed.

The steps that have to be taken to implement a management measure in the MoRE Model will be shown with the following example measure: Measures to reduce erosion

First, a measure variable for each reduction has to be added to the MoRE Model. In the case of erosion reduction, there is only one reduction, therefore only one measure variable has to be added (MM_ER_EFF_AL_SED), which will reduce the sediment emissions from erosion and the erosion for all substance groups by a certain percentage.

As a second step the measure variables for one variable have to be bundled into a single measure. In the case of erosion reduction, the measure variable MM_ER_EFF_AL_SED will be added to the measure erosion reduction.

The last step is to combine all single measures that are contributing to a management measure in a measure combination. The measure combination erosion reduction contains only the measure erosion reduction. This model structure is very similar to the structure of formulas that form an algorithm and one or more algorithms are combined into an algorithm stack.

DHSM

Table 15 lists a series of measures and indicates how the effect of such measures on the scale of the DRBD can be quantified using DHSM. The list of measures has been derived from information collected in the 2021 Danube River Basin Management Plan. The measures are subdivided in different clusters. In an occasional case, the table indicates that there is no direct way to quantify the effect of a measure, because the specific phenomenon or pathway is not explicitly represented (yet) in the DHSM.

Broadly speaking, there are two ways to represent measures in the DHSM input. For source control measures, often the emission factors need to be modified. For measures involving wastewater and stormwater management infrastructure, the spatial information that describes this infrastructure needs to be modified.

Table 15: Implementation of pollution control measures in DHSM.

Measure	Conceptual approach	Data requirements
Source control		
Use regulation	Reduction of emission factors for one or more sources	Case study / literature to estimate reduction
Industrial discharges control	Reduction of industry source	Sector prognosis to estimate reduction
Reduction of emissions to soil	Reduction of emission factors or activity rates for one or more sources	Case study / literature to estimate reduction
Prevention of accidental discharges and pollution from contaminated sites	No options in present DHSM as this source has not been quantified yet	
Best pesticide application practices	Reduction of emission factor for use in agriculture and/or use outside agriculture	Case study / literature to estimate reduction
Avoidance of tar-based products (PAHs)	Reduction of emission factors for road traffic and navigation	Case study / literature to estimate reduction
Wastewater management		
Construction of sewer systems	Change share of population connected to sewers and connected to IAS/LS	Estimate of which changes where

Measure	Conceptual approach	Data requirements
Construction of conventional WWTPs	Change share of collected wastewater undergoing treatment	Estimate of which changes where
Implementation of advanced wastewater treatment	Change share of collected wastewater undergoing advanced treatment	Estimate of which changes where, optionally based on agglomeration size
Handling of sewage sludge	Change share of sewage sludge re-used	Estimate of which changes where
Stormwater management		
Increased storage in combined sewers	Modify local rainfall threshold that causes CSOs	Estimate of which changes where
Retention and filtration of combined sewer overflows	No option to directly implement, could be mimicked by increasing threshold for CSOs	
Decoupling of stormwater collection systems	Change of share of combined sewers and increase local threshold for CSOs	Estimate of which changes where
Retention and filtration in stormwater collection systems	Change of substance-specific removal in such systems	Case study / literature to estimate removal
Local storage / infiltration of collected stormwater (Sponge Cities)	Change of share of stormwater centrally collected and discharged to WWTP/surface water	Estimate of which changes where
Solid waste management		
Improved solid waste management	No options in present DHSM as this source has not been quantified yet	
Connectivity between rivers and adjacent soils		
Measures to reduce erosion	Direct reduction of erosion pathway	Estimate for reduction factor, spatial distribution thereof
Reconstruction of wetlands and floodplains	Increase of in-stream retention processes	Estimate for (increased) removal and net settling, spatial distribution thereof

Some exploratory scenario simulations have been conducted with DHSM. These were aimed at three types of interventions: (i) the collection and treatment of domestic wastewater, (ii) the management of stormwater, and (iii) the reduction of erosion. The results of these scenarios are illustrated for a pharmaceutical (carbamazepine) and for a metal (zinc).

For exploring the effects of changes in the collection and treatment of domestic wastewater, the following scenarios were simulated:

1. A scenario S01 in which full connection to sewer systems is provided to all DRB inhabitants, without providing wastewater treatment.
2. A scenario S02 as the previous, with tertiary treatment applied for agglomerations with more than 100,000 population equivalents, and secondary treatment applied for agglomerations with more than 10,000 population equivalents.
3. A scenario S03 as the previous, with advanced treatment applied for agglomerations with more than 100,000 population equivalents.
4. A scenario S04 as the previous, with advanced treatment applied for agglomerations with more than 10,000 population equivalents.

Figure 57 shows how the basin-wide emission change in these four scenarios, compared to the present situation. In S01, the small contribution from direct emissions to surface water by unconnected households disappears (“HHo”). The total emissions increase strongly, because of the full connection without treatment of all currently unconnected households (“Unc”). In S02, where secondary/tertiary treatment is provided in agglomerations >10,000 PE, the emissions do not change, as carbamazepine is assumed not to be removed by conventional treatment. In the pathway distribution, part of “Unc” shifts to “WWTP”. In S03 and S04 an increasing share of the collected wastewater is undergoing advanced treatment with an assumed 90% removal of carbamazepine. This implies that the “WWTP” pathway decreases.

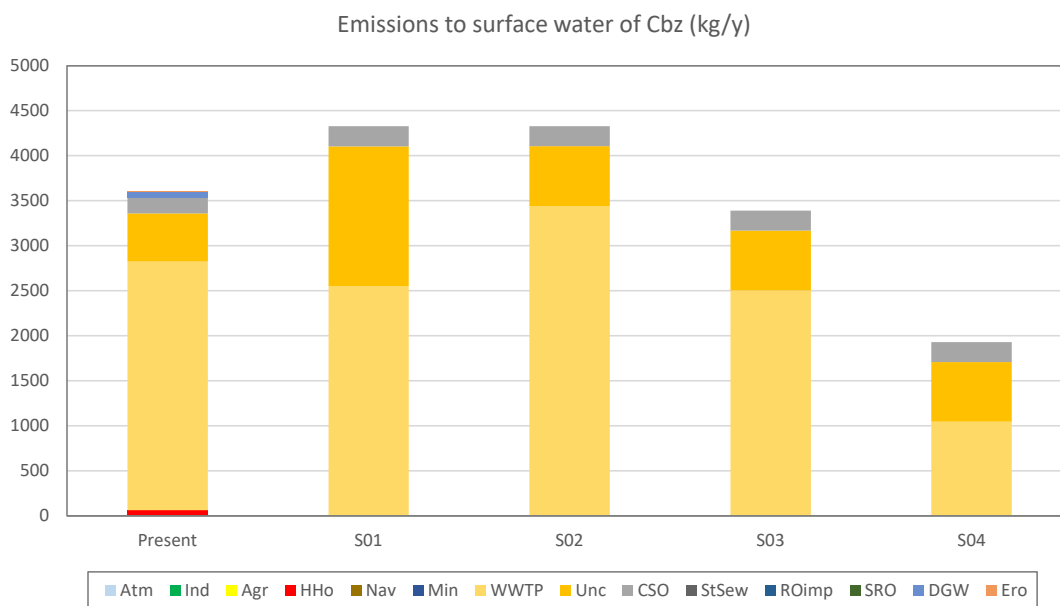


Figure 57: Changes in basin-wide emissions of carbamazepine in four exploratory scenarios.

Figure 58 shows the simulated effect in in-stream concentrations along the Danube River of these four scenarios. For context, this figure also shows the inter-annual variability of the annually averaged concentrations with current emissions. This variability is caused by the variability of river flows. The increased emissions in scenarios S01 and S02 lead to higher concentrations in the downstream section

of the river, as most of the presently unconnected households are in the central and east part of the basin. The reduction by advanced treatment in agglomerations > 100,000 PE (S03) is significant, but still within the range of natural variability. The reduction by advanced treatment in agglomerations > 10,000 PE (S04) is a lot stronger, and exceeds the range of natural variability.

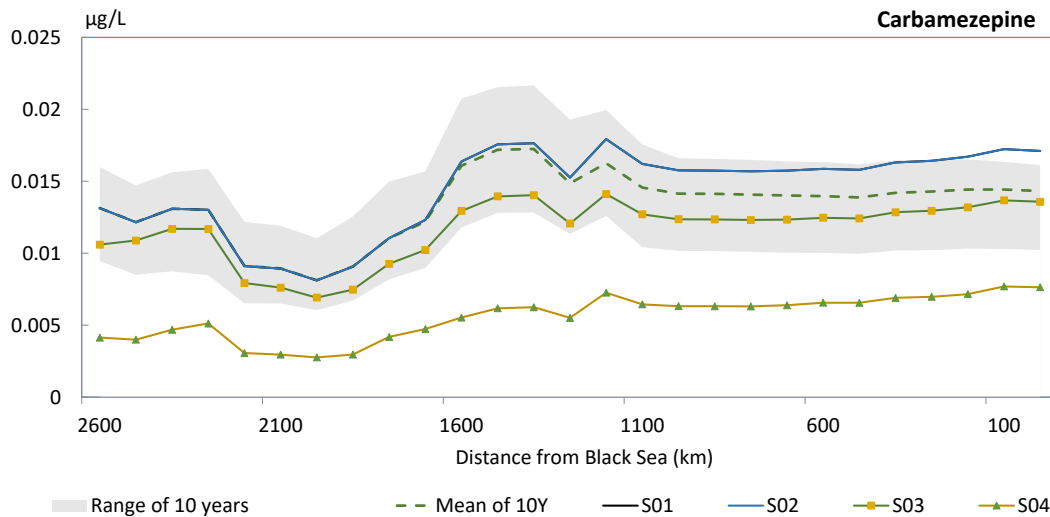


Figure 58: Concentrations along the Danube River of carbamazepine in the various scenarios.

For exploring the effects of changes in the management of stormwater and of erosion control measures, the following scenarios were simulated:

1. A scenario S05 that is equal to S02 discussed above, with 100% collection of stormwater, 100% treatment of the collected stormwater and no combined sewer overflows.
2. A scenario S06 that is equal to S02 discussed above, with 100% local storage and infiltration of stormwater and no combined sewer overflows.
3. A scenario S07 that starts from the present situation, in which all erosion is reduced by 50%.

Figure 59 shows how the basin-wide emission change in these four scenarios for zinc, compared to the present situation. In S02, the same phenomena emerge as those described above for carbamazepine. The overall change is however much smaller, as the contribution of domestic wastewater to zinc emissions is a lot smaller. In S05, the total emissions of zinc increase. This is the result of the assumption that all stormwater is now collected and discharged to surface waters, where in the present situation only urban stormwater was collected. The applied treatment cannot make up for this. In S06, the total emissions decrease, because most stormwater is now infiltrated and only a small fraction is able to reach surface waters by natural processes. Despite the quite extreme assumptions in S05 and S06, no dramatic changes to the emissions are calculated. This is evidently the consequence of erosion and subsurface flow being the main contributing pathways to zinc emissions. These pathways are controlled by concentrations in soil, which are part natural and part legacy pollution. As a consequence, scenario S07 has the strongest effect on zinc emissions.

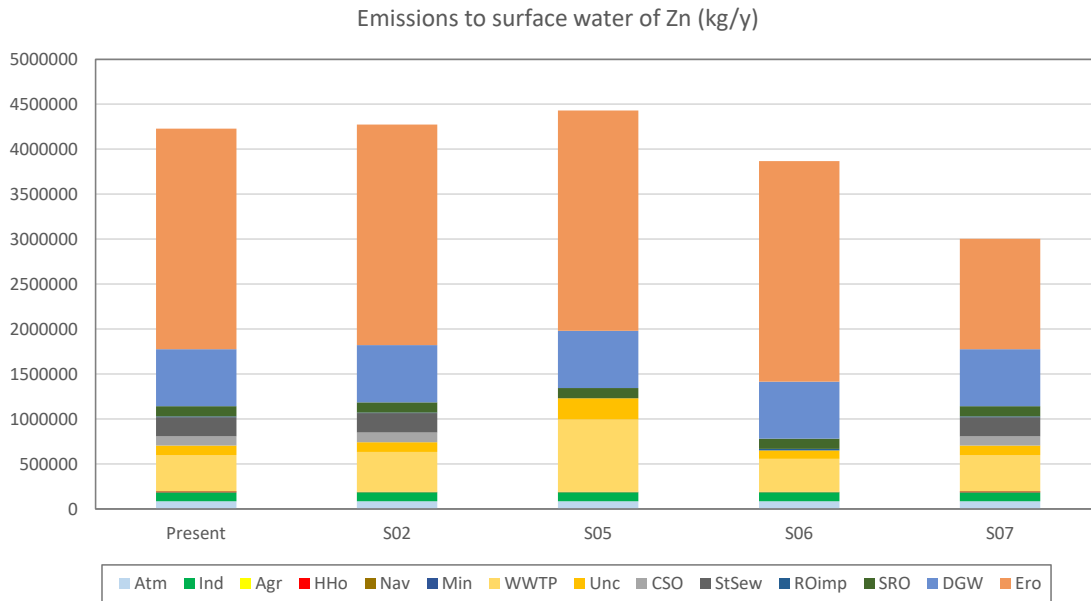


Figure 59: Changes in basin-wide emissions of zinc in four exploratory scenarios.

Figure 60 shows the simulated effect in in-stream concentrations along the Danube River of these four scenarios. River concentrations in S02 are the same as for present emissions. S05 and S06 show a small increase and a small decrease respectively, both well within the range of natural variability. The reduction by erosion control measures under the current assumptions is substantial, and exceeds the range of natural variability.

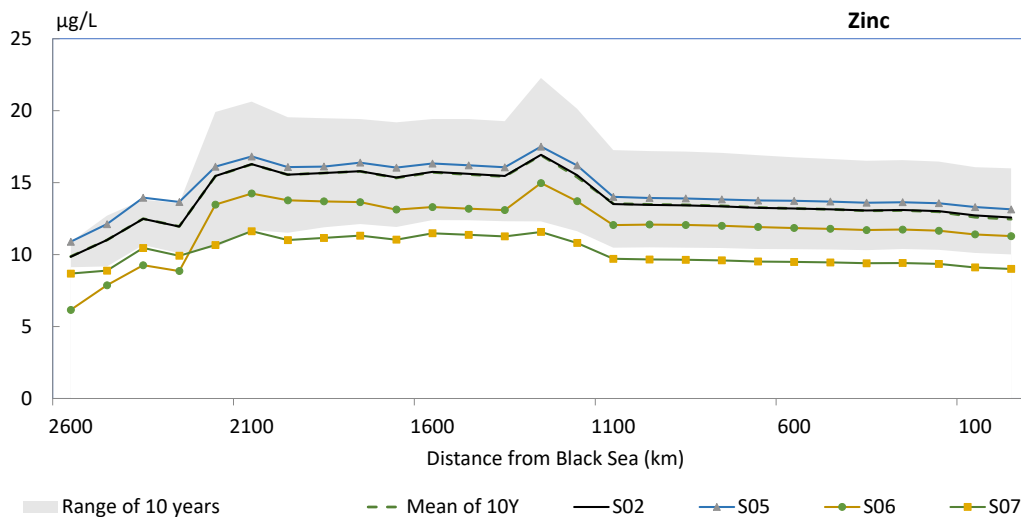


Figure 60: Concentrations along the Danube River of zinc in the various scenarios.

4.4.2 Other scenarios

MoRE

The implementation of climate change scenarios in MoRE was not within the scope of Danube Hazard m³c. However, such scenarios can be easily implemented. Two main areas in modelling must be adapted:

- Water balance (driven by information on hydrology and affecting most other pathway calculations).
- Soil loss from agricultural land (annually varied by the R-factor (rainfall erosivity) of the Revised Universal Soil Loss Equation used in the model).

Consequently, for each sub catchment the run off has to be modified based on climate scenarios. The effluent from WWTPs has to be revised, as input data directly derived from the documentation of wastewater treatment plant effluents. The monthly precipitation needs to be adapted (based on climate scenarios), from which the R-factor is calculated and the long-term annual soil loss is modified based on annual time steps.

The direct effects of climate change can finally be shown by comparing the results of the model in the actual state and the results adapted via the climate change scenarios.

DHSM

The implementation of climate change scenarios in DHSM was not within the scope of Danube Hazard m³c. The DHSM has two inputs that are expected to respond to climate change: (i) the hydrology input and (ii) the sediment delivery to streams. Both inputs are available for the period 2003-2013 with a daily interval. This implies that the climate-induced interannual variability of emissions and in-stream concentrations could be quantified. For in-stream concentrations, this was already illustrated by the range of 10 years indicated in Figure 58 and Figure 60. For the emissions, this is illustrated in Figure 61.

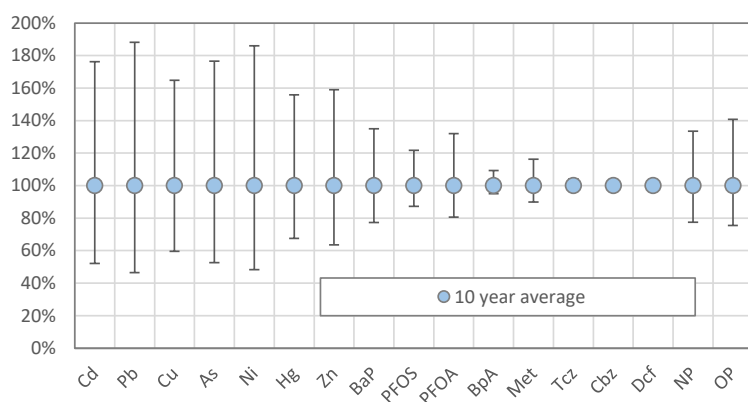


Figure 61: Range of 10 annual emission estimates for different hydrological years for 17 HS.

A tentative assessment of possible climate change impacts on emissions and in-stream concentrations is discussed in the Danube River Basin Scale Assessment Report. The elaboration of true global (including climate) change scenarios for the Danube River Basin is anticipated in the EU R&D project DOORS (<https://www.doorsblacksea.eu/>).

4.4.3 Stakeholder involvement

Danube countries should establish a dialogue and partnership among the relevant key sectors (e.g. water management, industry, agriculture) to seek mutually agreed actions to be jointly implemented. Efforts have to be made to strengthen coordination, consistency and complementarity between all relevant sectorial policies and funding schemes to ensure that they work in an integrated way and in good synergy (Output T3.3).

A first step in this direction was taken in this project. For each pilot region, a questionnaire was developed, translated into national language and submitted to regional and national stakeholders from different sectors. The questionnaire includes specific questions on potential mitigation measures, which might lead to a significant reduction of investigated substances that exceed the EQS (PFOS in five out of seven pilot regions, s-Metolachlor in an intensively used agricultural catchment area and dissolved Cadmium, dissolved Copper and dissolved Zinc in a catchment influenced by abandoned mining).

The proposed measures were described in detail, introduced, and characterized by the gained knowledge from monitoring and modelling to increase the specific insights of all Stakeholders. Detailed information addressed the conditions in the pilot region itself, the origin and effect of the relevant pollutants, considered the factor of exceedance and characterised the relevant emission pathways.

The proposed measures, which were put up for discussion include:

PFOS

- Advanced wastewater treatment at medium (10.000-1000.000 PE) and large (>100.000 PE) treatment plants by implementing quaternary treatment (activated carbon).
- Increased retention capacity of combined sewer overflow and increased treatment on WWTP due to extended storage.

Pesticides (Exceedance of the EQS from the National Substance List)

- Source control - reduction of s-Metolachlor application by 50% on all relevant crops.
- Erosion control: reduction of erosion from arable land.

Heavy Metals

- Short – term perspective: collecting and treatment of abandoned mining effluent from well-known and prospected locations (five effluents direct discharging into tributaries and main river).
- Mid-term/Long-term perspective: Groundwater remediation by restoration of most relevant diffuse sources from abandoned mining sites.

The Stakeholder were given the opportunity, to assess the proposed measure, make a statement on the technical feasibility and the extent of implementation. They were encouraged to document additional ideas or feasible measures by means of text fields that could be filled in freely.

For the seven pilot regions a total of 59 questionnaires were received. Most answers were submitted in the Viseu catchment, were abandoned mining leads to significant pollution, which might raise the awareness and the will or desire to share his opinion.

The technical implementation of the survey was realized by two main documents. One document including information on the pilot region, the relevant substances, the dominant pathways and characterization of possible measures were prepared in a pdf format attached to the mails. The questionnaire itself was prepared in Windows Forms and was kept very short.

The intended employment time was estimated at about 15 minutes, with the documentation estimated at about 10 minutes and the questionnaire itself at about 5 minutes.

The following figure gives an impression of the appearance of the questionnaire (Figure 62).

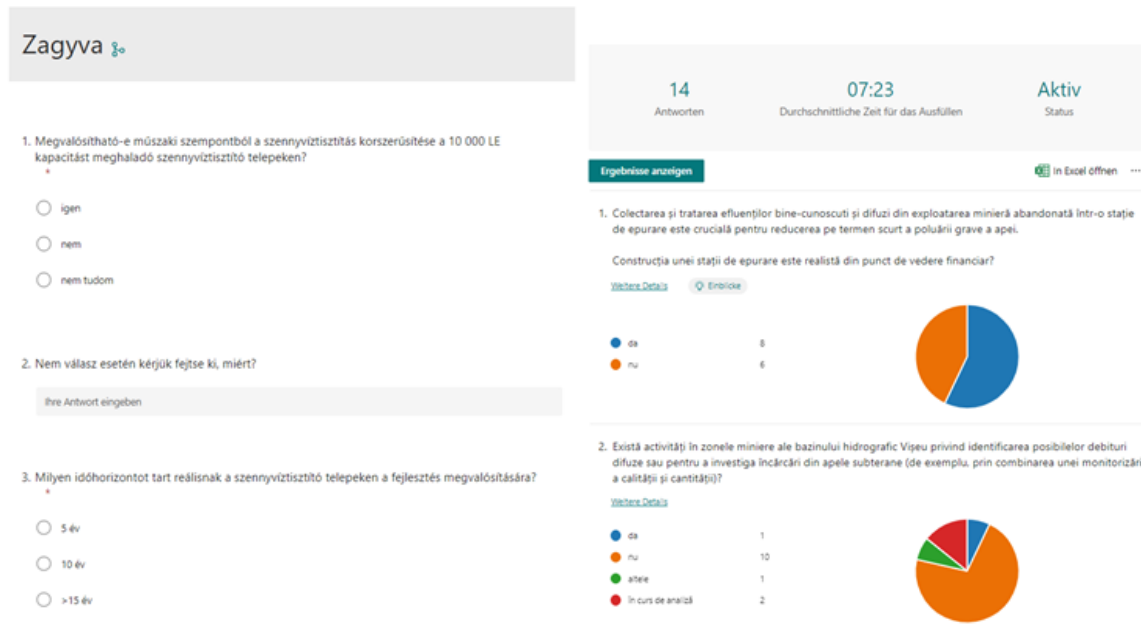


Figure 62: Screenshots of one Hungarian questionnaire (left side) and on the automated analyses prepared by Windows Forms for the Romanian Viseu pilot region.

5 References

- 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.
- Amodio, M., Catino, S., Dambruoso, P. R., de Gennaro, G., Di Gilio, A., Giungato, P., . . . Tutino, M. (2014). Atmospheric Deposition: Sampling Procedures, Analytical Methods, and Main Recent Findings from the Scientific Literature. *Advances in Meteorology*. doi:10.1155/2014/161730
- Behrendt H., Huber P., Kornmilch M., Opitz D., Schmoll L O., Scholz G., Uebe R., 2002. Estimation of the nutrient inputs into river basins – experiences from German rivers. *Regional Environmental Changes* 3: 107–117
- Budai, P., M.K. Kardos, M. Knolmár, G. Szemán, J. Turczel, A. Clement (2020). Development of an autonomous flow-proportional water sampler for the estimation of pollutant loads in urban runoff. *Environ. Monit. Assess.* 192: 572 <https://doi.org/10.1007/s10661-020-08536-3>
- European Commission. (2022). *Statistical regions in the European Union and partner countries*. Retrieved from NUTS and statistical regions 2021: <https://ec.europa.eu/eurostat/documents/3859598/15193590/KS-GQ-22-010-EN-N.pdf/82e738dc-fe63-6594-8b2c-1b131ab3f877?t=1666687530717>
- Food and Agriculture Organization of the United Nations (FAO). (2006). *Guidelines for soil description*. Rome.
- Fuchs, S.; Kaiser, M.; Kiemle, L.; Kittlaus, S.; Rothvoß, S.; Toshovski, S. et al. (2017a): Modeling of Regionalized Emissions (MoRE) into Water Bodies. An Open-Source River Basin Management System. In: *Water* 9 (4), S. 239. DOI: 10.3390/w9040239
- Stephan Fuchs, Tatyana Weber, Ramona Wander, Snezhina Toshovski, Steffen Kittlaus, Lucas Reid, Martin Bach, Laura Klement, Thomas Hillenbrand, Felix Tettenborn (2017b): Effizienz von Maßnahmen zur Reduktion von Stoffeinträgen. UBA-Texte 05/2017. Dessau-Rosslau. <https://www.umweltbundesamt.de/publikationen/effizienz-von-massnahmen-zur-reduktion-von>; page 82
- Fuchs, S.; Kaiser, M.; Reid, L.; Toshovski, S.; Nickel, J. P.; Gabriel, O.; Clara, M.; Hochedlinger, G., Trautvetter, H.; Hepp, G.; Kittlaus, S.; Zessner, M.; Kaul, U.; Forstner, S. (2019): Grenzüberschreitende Betrachtung des Inn-Salzach-Einzugsgebietes als Grundlage für ein transnationales Gewässermanagement. Final report
- Gasperi, J., Sebastian, C., Ruban, V., Delamain, M., Percot, S., Wiest, L., Mirande, C., Caupos, E., Demare, D., Kessoo, M.D. K., Saad, M., Schwartz, J. J., Dubois, P., Fratta, C., Wolff, H., Moilleron, R., Chebbo, G., Cren, C., Millet, M., Barraud, S., & Gromaire, M. C. (2014). Micropollutants in urban stormwater: occurrence, concentrations, and atmospheric contributions for a wide range of contaminants in three French catchments. *Environmental Science and Pollution Research*, 21, 5267–5281. <https://doi.org/10.1007/s11356-013-2396-0>

- Göbel, P., Dierkes, C., & Coldewey, W. G. (2007). Storm water runoff concentration matrix for urban areas. *Journal of Contaminant Hydrology*, 91, 26–42.
<https://doi.org/10.1016/j.jconhyd.2006.08.008>
- Helsel, D. R. (2006). Fabricating data: how substituting values for nondetects can ruin results, and what can be done about it. *Chemosphere*.
- Helsel, D. R. (2012). *Statistics for censored environmental data using Minitab and R* (2 Ausg.). Hoboken, N.J.: Wiley. doi:10.1002/9781118162729
- Hundecha, Y., Arheimer, B., Donnelly, C., Pechlivanidis, I., 2016. A regional parameter estimation scheme for a pan-European multi-basin model, *Journal of Hydrology: Regional Studies*, 6, 90-111, DOI: 10.1016/j.ejrh.2016.04.002
- Julian, P., & Helsel, D. R. (2023). R package: NADA2: Data Analysis for Censored Environmental Data. Von <https://cran.r-project.org/web/packages/NADA2/index.html> abgerufen
- Kittlaus, S., Fuchs, S. (2015). Using large volume samplers for the monitoring of particle bound micro pollutants in rivers. Poster at EGU General Assembly.
- Kittlaus, S., Reid, L., Fuchs, S. (2016). Grenzüberschreitende Modellierung von Stoffeinträgen im Einzugsgebiet des Inn. Vortrag zum Tag der Hydrologie 2016. http://nordsee.lfi.rwth-aachen.de/fghw-uploads/fghw-hefte/Tagungsband_TdH_2016.pdf
- Kosztra, B., Büttner, G., Hazeu, G., & Arnold, S. (10. 05 2019). *Updated CLC illustrated nomenclature guidelines*. Von <https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines/html/index.html> abgerufen
- Lee, L. (2020). R package: NADA: Nondetects and Data Analysis for Environmental Data. Von <https://CRAN.R-project.org/package=NADA> abgerufen
- Lewis, J. (2006). Turbidity-controlled suspended sediment sampling for runoff-event load estimation. *Water Resources Research*, 32(7), 2299-2310, 2006.
- Lewis, J., Eads, R. (2009). *Implementation Guide for turbidity threshold sampling - Principles, Procedures and Analysis*. USDA Forest Service, 2009.
- Moser, A., Bader, H.-P., Fenicia, F., Scheidegger, R., & Stamm, C. (2015). CrossWater - Modelling micropollutant loads from different sources in. *Geophysical Research Abstracts Vol. 17*. EGU, Vienna
- Pekey, B., Karakas, D., Ayberk, S. (2007). Atmospheric deposition of polycyclic aromatic hydrocarbons to Izmit Bay, Turkey. *Chemosphere*, 67(3), 537-547.
- Phillips, J., Russell, M., Walling, D. (2000). Time-integrated sampling of fluvial suspended sediment, a simple methodology for small catchments. *Hydrological Processes*.
- R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. Wien. Von <http://www.r-project.org/>
- Rode, M. and Suhr, U. (2007) Uncertainties in Selected River Water Quality Data. *Hydrology and Earth System Sciences*, 11, 863-874.
<https://doi.org/10.5194/hess-11-863-2007>

- Rocco, C., Duro, I., Di Rosa, S., Fagnano, M., Fiorentino, N., Vetromile, A., Adamo, P. (2016). Composite vs. discrete soil sampling in assessing soil pollution of agricultural sites affected by solid waste disposal. *Journal of Geochemical Exploration*, 30-38.
- Sarkadi, J., Németh, T., Kádár, I. (1986). Heterogeneity of the easily soluble phosphorus content of soils (in Hungarian: A talaj könnyen oldható tápanyagtartalmának heterogenitása). *Agrokémia és talajtan*, 295-306.
- Simpson, B. (2017). *Wastewater Sampling*. Athens, Georgia: US-EPA Science and Ecosystem Support Division.
- Tóth, G., Jones, A., Montanarella, L. (2013). *Lucas Topsoil Survey*. Ispra: JRC TECHNICAL REPORTS, JRC.
- Van den Roovaart, J., Ullrich, A., Duijnhoven, N. van, Prchalová, H., Bleser, J., Whalley, C., 2022, Calculating emissions to water – a simplified method. ETC/ICM Technical Report 03/2022: European Topic Centre on Inland, Coastal and Marine Waters, 205 pp
- Whalley, C., Mohaupt, V., Busch, W., van den Roovaart, J., van Duijnhoven, N., Kirst, I., Schmedtje, U., Altenburger, R., Sommer, L., 2018. *Chemicals in European waters, Knowledge developments*. EEA Report No 18/2018, 76 pp., doi: 10.2800/265080