www.napsea.eu

DELIVERABLE D5.4

FINAL DATA MANAGEMENT PLAN

Work Package 5
Overall communication and dissemination

30-09-2025

Grant Agreement number	**101060418**
Project title	**NAPSEA: the effectiveness of Nitrogen And Phosphorus load reduction measures from Source to sEA, considering the effects of climate change**
Project DOI	10.3030/101060418
Deliverable title	Final Data Management Plan
Deliverable number	D5.4
Deliverable version	Version 1.0
Contractual date of delivery	30-09-2025
Actual date of delivery	30-09-2025
Document status	Concept (of initial version)
Document version	Concept
Online access	Yes
Diffusion	Public
Nature of deliverable	Report
Work Package	WP5: Overall communication and dissemination
Partner responsible	Deltares
Contributing Partners	Deltares
Author(s)	Van Der Heijden, L.H.
Editor	Van Der Heijden, L.H.
Approved by	Blauw, A.
Project Officer	Blanca Saez-Lacava
Abstract	The DMP is a document that outlines how data is handled after the project, including what data is collected, processed, and shared, as well as how it is curated and preserved. The NAPSEA project will work with existing and new datasets, generate various types of data, and make the data findable, accessible, interoperable, and re-usable (FAIR data principle) through published deliverables. The deliverable includes lessons learned and barriers overcome.
Keywords	Data management; FAIR; Model output;

Contents

1. Acronymes

C^nANDY	Coupled Complex Algal-Nutrient Dynamics
DCSM	Dutch Continental Shelf Model
FAIR	Findability, Accessibility, Interoperability, and Reusability
mQM	multiscale water Quality Model
mHM	mesoscale Hydrological Model
WP	Work package
HRU	Hydrologic Response Unit
TN	Total Nitrogen
TP	Total phosphorus
NO_3	Nitrate
PO_4	Ortho-phosphate
WWTP	Wastewater Treatment Plant
SRTM	Shuttle Radar Topography Mission
UWWTD	Urban Waste Water Treatment Directive (EU database)
EEA	European Environment Agency
BOKU	University of Natural Resources and Life Sciences, Vienna (source of climate data)
SCHISM-ECOSMO	Coupled model: SCHISM (Semi-implicit Cross-scale Hydroscience Integrated System Model) + ECOSMO (Ecological System Model)

2. Executive summary

This deliverable outlines the data management strategy for the NAPSEA project, which investigates nutrient load reduction from source to sea under climate change. It describes how data is collected, processed, stored, and shared in accordance with the FAIR principles (Findable, Accessible, Interoperable, Reusable).

The project primarily uses existing datasets for hydrological and water quality modeling, including models such as mHM, mQM, CnANDY, SWAT, and DCSM. No new measurements were taken, and model outputs are managed and disseminated through project deliverables.

Data security is ensured through encrypted storage, secure transfer protocols, and regular backups. Ethical standards are upheld in all data-related activities, including interviews and workshops. The plan adheres to Horizon Europe guidelines and does not adopt additional national or sector-specific data policies.

3. Introduction

As mentioned in Wilkinson et al. (2016): “Good data management is not a goal in itself, but rather is the key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse by the community after the data publication process”. The data management plan (DMP) is an important component in successfully managing data. They describe the data management life cycle, including discovery, collection, processing, generation and organization. Managing data and setting standards from the start will increase the value of the data that will be collected, processed and generated. The European Commission (EC) states that, under guidelines for the Horizon Europe (HE) projects, a good DMP should include information on:

- The handling of research data during and after the end of the project;
- What data will be collected, processed and/or generated;
- Which methodology and standards will be applied;
- Whether data will be shared/made open access and;
- How data will be curated and preserved (including after the end of the project).

This deliverable serves as a DMP that will help and guide the handling of research data after the end of the project. Our DMP follows the general template recommended by the EC HE guidelines that includes:

- a data summary, specifying data collection objectives, data types, size, storage and utility;
- a description of FAIR data use, with a description of how we keep data findable, accessible, interoperable and re-usable according to the FAIR data policy (Wilkinson et al., 2016).

Furthermore, chapters about allocation of resources, data security and ethical aspects are presented.

4. Data summary

4.1. Input data

The main data that was used in the NAPSEA project exists of model input data. This can be data that was converted to suit the modelling tool or directly applicable data. Merely existing data is used on which the FAIR data principle is already applied. In Table 1 an overview of used model applications and the reference/deliverable/paragraph in which data is described further. These refer to the reference scenario data, so existing data of past periods in time.

Table 1. Overview of used models and when feasible the reference in which further details of input data are described.

Model	Deliverable/reference/paragraph
mHM	Paragraph *4.1.1. Hydrology and (inland) waterquality model* describing the data used for the mHM.
mQM	Paragraph *4.1.1. Hydrology and (inland) waterquality model* describing the data used for the mHM. Data used for model parametrization is described in more detail in D3.2 (Musolff & Ledesma, 2024).
C^{n}ANDY	Paragraph *4.1.1. Hydrology and (inland) waterquality model* describing the data used for the mHM. Data used for model parametrization is described in more detail in D3.2 (Musolff & Ledesma, 2024).
SCHISM-ECOSMO model	As hydrodynamical core the SCHISM-ECOSMO model uses the Semi-implicit Cross-scale Hydroscience Integrated System Model (SCHISM, Zhang et al., 2016), which is coupled with an ecological model, the ECOlogical System Model (ECOSMO, Yumruktepe, 2022).
SWAT-model	Paragraph *4.1.2. SWAT model* describing the data used.
DCSM waterquality model	As hydrodynamical core the DCSM model uses the D-Flow Flexible Mesh module[1] from the Delft3D FM Suite[2] online coupled with a waterquality model (D-Water

[1] https://www.deltares.nl/en/software-and-data/products/delft3d-fm-suite/modules/d-flow-flexible-mesh
[2] https://www.deltares.nl/en/software-and-data/products/delft3d-fm-suite/modules

	Quality[3]). The model parametrization and input data are described in Zijl et al. (2023) for hydrodynamics and in Vilmin et al. (in progress) for waterquality. More info in paragraph *4.1.3. DSCM model*.

4.1.1. Hydrology and (inland) waterquality model

In this paragraph the datasets used for the reference model runs for mHM, mQM and C^{n}ANDY, both input data and validation data, are defined (Table 2).

Table 2. Description of model input data and model validation data for the mHM, mQM and C^{n}ANDY simulations.

Data type	Variable	Resolution	Period	Source
		Model input data		
Meteorological data	Precipitation Temperature (mean, min, max)	Daily (1km)	1950-2020	BOKU: High-resolution (1 km) daily climate data (precipitation, minimum and maximum temperatures) (Cruz-Alonso et al., (2023)). https://doi.org/10.1016/j.envsoft.2023.105627
Hydrological data	Predicted discharge	Daily at 5x5 km	1950-2020, Scenarios until 2100	Predicted discharge using the mHM hydrological model results https://doi:10.1029/2008WR007327 https://doi:10.1029/2012WR012195
	Measured discharge	Daily at gauging stations	1950-2020	https://portal.grdc.bafg.de/
Soil data	Soil type			European Soil Database https://esdac.jrc.ec.europa.eu/content/european-soil-database-v20-vector-and-attribute-data
Agricultural data, Nitrogen inputs	- Nitrogen surplus including diffuse sources from non-agricultural land, - Application of fertilizer/manure, - Dates of farming practices.	Annual at 10 x 10 km	1850-2019	Batool et al. (2022) https://doi.org/10.1038/s41597-022-01693-
Land use depended diffuse phosphorous inputs	Dissolved phosphorus inputs	Average values	-	Yang et al. (2021) https://doi.org/10.1016/j.watres.2021.116887
Nutrient point sources	Total N and total P discharge from point sources	Only WWTP plants with population equivalents (PE) > 2000 are considered	average	In Germany Büttner et al. (2022) https://doi.org/10.1016/j.watres.2022.118382 EU scale: EEA (2020) UWWTD database
Morphological data	Digital Elevation Model	Resampled to (100 m)	-	SRTM (https://earthexplorer.usgs.gov/) A terrain elevation model was obtained from the Shuttle Radar Topography Mission (SRTM) sensor.

[3] https://www.deltares.nl/en/software-and-data/products/delft3d-fm-suite/modules/d-water-quality

	Land use/ land cover		-	CORINE https://land.copernicus.eu/pan-european/corine-land-cover Land-cover data were derived from CORINE Land Cover 10 ha (https://gdz.bkg.bund.de/index.php/default/open-data.html, last accessed 20 Dec. 2022). These datasets were resampled to a spatial resolution of 100 m × 100 m for model simulations.
		Model validation data		
Measured Nitrogen and phosphorus concentra-tions, surface water	**Elbe and Rhine** Nitrogen and Phosphorous species concentration		1968-2020	**Elbe and Rhine** Ebeling et al. (2022) https://doi.org/10.5194/essd-14-3715-2022 Virro et al. (2021) https://doi.org/10.5281/zenodo.5097436
	Hunze TN, TP and other water quality parameters (such as Chloride, dissolved oxygen, pH, NO_3, PO_4 etc)	Weekly/ monthly	2000-2023	**Hunze** Biweekly meaured concentrations data at nine gauging stations from upstream to downstream area of the Hunze catchment was collected from the responsible authorities.
Measured nitrogen concentration, groundwater	Nitrogen species concentration	bi-annually to annualy	1990-2017	EEA database https://www.eea.europa.eu/data-and-maps/data/waterbase-water-quality-icm-1

4.1.2. SWAT model

In this paragraph the datasets used for the reference model runs for existing model of the Hunze catchment are defined (Table 3).

Table 3. Description of model input data for the SWAT simulations.

Data type	Variable	Resolution	Period	Source
		Model input data		
Meteorological data	Precipitation, Evaporation	Monthly averages (Netherlands)	2050–2055	Derived from IPCC RCP 4.5 and RCP 8.5 scenarios (see Musolff et al., 2025 for more details)
Meteorological data	Precipitation, Temperature	Daily (original model setup)	2007–2016	Inherited from Gevaert & Waterloo (2019)
Hydrological data	Discharge (simulated)	Subbasin-level (33 subbasins)	2007–2016	SWAT model output
Nutrient data	NO_3^-, TP concentrations	HRU-level (1962 HRUs)	2012–2016	Calibrated against measured concentrations
Land use data	Land use types (agriculture, nature)	HRU-level	2007–2016	Gevaert & Waterloo (2019)
Soil data	Soil types	HRU-level	2007–2016	Gevaert & Waterloo (2019)
Drainage system	Tile drainage depth	Set to 0.6 m	Updated model	Model refinement to simulate shallow lateral discharge
Scenario data	Land use change, wastewater treatment	Various spatial resolutions	Future scenarios	Scenario definitions in Hunze case study (Musolff et al., 2025)
Scenario data	Nature-based solutions	Stream and riparian zone level	Future scenarios	Scenario definitions in Hunze case study (Musolff et al., 2025)

4.1.3. DCSM model

The model input for DCSM are described in Zijl et al. (2023) for hydrodynamics and in Vilmin et al. (in progress) for waterquality. The only difference between the validated model (Vilmin et al., in progress) and the model used for NAPSEA is the addition of nutrient reduction scenarios. These reduction scenarios are based on the outcomes of the mHM, mQM and C^nANDY models. A relative reduction, obtained from those model, was applied on the validated reference for the DCSM model. This meant a reduction (most of the time) of nutrient concentrations as well as discharge into the greater North Sea, applied on the river in the Netherlands and Germany (Rhine, Elbe, Ems, Weser, Meuse, and more).

5. Research outputs

Most of the research outputs that are generated (and potentially re-used) throughout the NAPSEA project is model output. Several model simulations were conducted throughout the course of the NAPSEA project. The data management of these newly generated model consists of a similar principle as the FAIR data principle.

6. Allocation of resources

Since no new data has come in via measurements, no cost were made for producing, storing and managing data. Existing data is already stored and managed. Data coming from model output will be stored and made FAIR via the delivered products (e.g., reports and models).

7. Data security

The security of research data is of utmost importance to ensure that the data is not lost, stolen, or accessed by unauthorized parties. Therefore, several measures will be taken to secure the data throughout the project.

7.1. Privacy

Any sensitive data, such as personally identifiable information (PII), will be handled with extra care. Access to this data will be restricted to authorized personnel only, and the data will be encrypted at rest and in transit. Data that is no longer needed will be securely deleted using appropriate methods.

7.2. Data Security

Secure Storage

All research data will be stored in a secure and backed-up location. The data will be stored on a password-protected network drive with restricted access, and the server will be backed up regularly to ensure that no data is lost in the event of a system failure. data security and storage, at this stage a mix of storage technologies will be explored (cloud, no cloud, SQL, NoSQL, etc.) in order to achieve the most effective solution. The most obvious solution in the future is to move the storage to the cloud but at this stage this solution may not be the most cost effective and other alternatives may need to be considered. In any case the data storage solution will ensure that the new paradigm of cloud computing is compatible with a future deployment via Data and Information Access Service (DIAS). Apart from this, several existing marine data infrastructure platforms will be considered that have a solid data security. For example, SeaDataNet is including the important issues of trust which are addressed in data-based research: security, confidentiality, ownership, assured provenance, authenticity, as well as the quality of the data and the metadata" (SeaDataNet, 2010).

Data Transfer

To ensure secure transfer of data, all files will be encrypted using the Advanced Encryption Standard (AES) algorithm. File transfer will be performed via secure protocols such as SSH or SFTP to ensure that no data is intercepted by unauthorized parties.

Data Recovery

In the event of data loss or corruption, provisions will be in place for data recovery. Backups of the server will be created regularly, and the data will be stored off-site in a secure location. In addition, the data recovery process will be tested periodically to ensure its effectiveness.

8. Ethics

Under the NAPSEA project several interviews and workshops will be conducted. Several ethical key aspects were identified in this respect. In the process of conducting interviews and workshops, collecting data, NAPSEA will ad-here to the ethical principles: the principle of proportionality, the right to privacy, the right to protection of personal data, the right to physical and mental integrity of a person, the right to non-discrimination and the need to ensure high levels of human health protection.

The NAPSEA team is also committed to upholding high standards of academic integrity and will ensure that all sources referenced in the project are properly cited. NAPSEA will adhere to all relevant referencing guidelines and best practices. Accurate referencing is essential to the credibility and legitimacy of the project's findings, and all necessary steps will be taken to avoid any issues of plagiarism or academic misconduct.

9. Other issues

NAPSEA won't make use of other national/funder/sectorial/departmental procedures for data management.

10. References

Büttner O, Jawitz JW, Birk S, Borchardt D. Why wastewater treatment fails to protect stream ecosystems in Europe. Water Research 2022; 217: 118382.

Gericke, A. & Leujak, W. (2024). Model input of selected scenarios. EC report of grant 101060418 Deliverable 3.4. https://napsea.eu/wp-content/uploads/2024/10/D3.4_Model-input-of-selected-scenarios.p

Gericke, A., Leujak, W., Musolff, A. & Geidel, T. (2024). Set of Scenarios. EC report of grant 101060418 Deliverable 3.3. https://napsea.eu/wp-content/uploads/2024/04/D3.3-Set-of-Scenarios_NAPSEA.pdf

Gevaert, A., M.J. Waterloo, (2019). Topsoil duurzame waterkwaliteit Drenthe. Effecten van klimaat- en beheerscenario's op de export van nutriënten en gewasbeschermingsmiddelen in de stroomgebieden van de Drentsche Aa en de Hunze. RPS-report 1603390A00-R19-710.

Jomaa, S. & Musolff, A. (2023). Data collection. EC report of grant 101060418 Deliverable 3.1. https://napsea.eu/wp-content/uploads/2024/02/D3.1_Data_overview_NAPSEA_final.pdf

Musolff, A. & Ledesma, J. (2024). Calibrated models. EC report of grant 101060418 Deliverable 3.2. https://napsea.eu/wp-content/uploads/2024/04/D3.2.-DEM_Calibrated_models_NAPSEA.pdf

Musolff, A., Ledesma, J., Gericke, A., Rozemeijer, J., Troost, T., and Liu, X. (2025). Effectiveness of measures. EC report of grant 101060418 Deliverable 3.5. https://napsea.eu/wp-content/uploads/2025/09/D3.5_Effectiveness-of-scenarios_update_incl_appendix.pdf

SeaDataNet (2010). Quality control procedures.

Vilmin, L., di Paola, F., Heye, S., and Blauw, A. (in progress). D-Flow FM and D-Water Quality 3D North Sea model. 3D DCSM-FM with water quality: Model set-up and validation.

Wilkinson, M. D. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. doi:10.1038/sdata.2016.18

Yang, S., Bertuzzo, E., Büttner, O., Borchardt, D., & Rao, P. S. C. (2021). Emergent spatial patterns of competing benthic and pelagic algae in a river network: A parsimonious basin-scale modeling analysis. Water Research, 193. https://doi.org/10.1016/j.watres.2021.116887

Yumruktepe, V. Ç., A. Samuelsen, and U. Daewel. 2022. ECOSMO II (CHL): a marine biogeochemical model for the North Atlantic and the Arctic. Geoscientific Model Development 15:3901-3921.

Zhang, Y. J., F. Ye, E. V. Stanev, and S. Grashorn. 2016. Seamless cross-scale modeling with SCHISM. Ocean Modelling **102**:64-81.

Zijl, F., Zijlker, T., Laan, S., & Groeneboom, J. (2023): 3D DCSM-FM: a sixth-generation model for the NW European Shelf : 2022 release, Deltares, 11208054-004-ZKS-0003.