SOS-Water deliverable report

D4.2 Introducing innovative water indicators towards a system of indicators building frameworks

Lead beneficiary	4-POLIMI	Due Date	30 March 2025
WP no	4	New due date (if delay)	
Task no	4.2 and 4.3	Actual Delivery Date	28 March 2025
Dissemination level	Public - PU	Status	FINAL - submitted

Authors

Authors	Partner no	Partner organisation	Name of author
Main author	4	POLIMI	Wei Xia, Sandra Ricart, Matteo Giuliani, Andrea Castelletti
Contributing author(s)	3	UPV	Najib Boubakri

Review

Authors	Partner no	Partner organisation	Name of author
Technical review	3	UPV	Miguel Angel Jiménez Bello
Language review – *if applicable*			

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Document history

Date	Version	Chapters affected	Description of change	Author	Document status
4/12/2024	0.1	Outline	Table of Content	Wei Xia, Sandra Ricart, Matteo Giuliani, Andrea Castelletti	DRAFT
29/01/2025	0.2	Section 3	Indicator aggregation	Wei Xia, Sandra Ricart, Matteo Giuliani, Andrea Castelletti	DRAFT
20/02/2025	0.3	Section 2	Indicator assessment and innovation	Najib Boubakri, Hector Macian-Sorribes	DRAFT
28/02/2025	0.4	Section 1 and 4	Introduction and discussion next steps	Wei Xia, Sandra Ricart, Matteo Giuliani, Andrea Castelletti	DRAFT
24/03/2025	0.5	Section 2	indicator assessment and innovative indicators	Hector Macian-Sorribes	DRAFT
28/03/2025	1	All	Final formatting & PDF creation	Alberto Fresolone	FINAL
20/01/2026	2.0	All	Update based on reviewers feedback	Hector Macian-Sorribes, Bruno Invernizzi	

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Project funded by

Publishable Executive Summary

This deliverable, D4.2: Introducing Innovative Water Indicators Towards a System of Indicators Building Frameworks, presents the methodological framework for developing a comprehensive and integrated system of water indicators within the SOS-WATER project. The deliverable builds upon Task 4.2 and Task 4.3 by critically assessing existing water indicators, identifying key gaps, proposing guidelines for the development of innovative indicators, and developing systematic aggregation mechanisms to improve decision-making in water resource management.

The first part of the deliverable focuses on indicator assessment and innovation (Section 2). Our approach builds upon the initial selection and evaluation of key water indicators documented in D4.1. This process was further refined through interactive workshops with project partners, which provided valuable insights into which indicators are both practical and preferred across different contexts. These workshops were complemented by a detailed survey conducted among project partners to rigorously assess the suitability of the indicators for water resource management. Through these multi-step assessments, this section identifies critical limitations in current approaches, particularly their ability to capture the complex spatial and temporal variability inherent in water systems. In response to these gaps, we outline strategic directions for the development of innovative indicators that prioritize multi-dimensional integration, enhanced data resolution, and climate-sensitive frameworks. Finally, preliminary thresholds for the proposed indicators, derived from an extensive review of the literature, are provided. These thresholds will need further refinement and adjustment to align with the specific characteristics and contextual requirements of each case study.

The second part of the deliverable explores indicator aggregation mechanisms (Section 3), recognizing the complexity and multi-dimensional nature of water systems. Two complementary approaches are introduced. First, data-driven methods for indicator reduction (Section 3.1) utilize techniques such as Non-negative Principal Component Analysis (NPCA) and evolutionary algorithms (PCSEA, ORA-RPSS, and δ-MOSS Greedy) to reduce dimensionality and redundancy while preserving key information. Second, Multi-Attribute Value Theory (MAVT) (Section 3.2) is employed to integrate diverse indicators into a decision-support framework, allowing the assignment of weights based on stakeholder preferences and facilitating the development of meta-indicators. Together, these approaches ensure that the indicator system remains both analytically rigorous and operationally useful.

The discussion and next steps (Section 4) highlights remaining challenges and areas for further development. While the methodological framework is well-established, additional calibration and validation are required across multiple case studies to refine the innovative indicators and aggregation techniques. Future work will focus on integrating these indicators into the broader SOS-WATER framework to ensure compatibility with modelling and water system performance assessment efforts.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Table of Content

1. INTRODUCTION 6

2. INDICATOR ASSESSMENT AND INNOVATIVE INDICATORS 7

2.1 INDICATOR GAPS IDENTIFICATION 7
2.1.1 Critical review of indicators 7
2.1.2. Development and distribution of an indicator evaluation survey 9
2.1.3. Analysis of survey results and indicator assessment 11
2.1.4. Identified gaps 13
2.2 INNOVATIVE INDICATORS 17
2.3 PRELIMINARY THRESHOLDS 18
2.3.1. Identification of indicators and information sources on thresholds 19
2.3.2. Comparison of thresholds provided by different sources and final choice 21
2.3.3. Spatio-temporal aggregation and uncertainty on thresholds 22

3. INDICATOR AGGREGATION MECHANISMS 24

3.1 DATA-DRIVEN METHODS 24
3.1.1. Non-negative Principal Component Analysis 24
3.1.2. PCSEA 25
3.1.3. ORA-RPSS 25
3.1.4. δ-MOSS Greedy 26
3.1.5. Numerical experiments 27
3.1.5.1. Benchmarking using DTLZ5(I,M) Problem 27
3.1.5.2. Experimental setup 28
3.1.5.3. Numerical results 29
3.2. MULTI-ATTRIBUTE VALUE THEORY 31
3.2.1. Evaluating the alternatives 31
3.2.2. Giving weights to alternatives 34
3.2.3. Global and partial value functions 35
3.2.4. Evaluation 37
3.2.5. Numerical experiments 38

4. DISCUSSION AND NEXT STEPS 41

REFERENCES 43

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

List of Figures & Tables

Figure 1. Interactive workshop discussion on indicator assessment using Miro Board. 8
Figure 2. Interactive workshop at Utrecht for indicators discussion..... 8
Figure 3. Objectives hierarchy and indicator framework from WP1..... 9
Figure 4. Indicator evaluation in the Jucar River basin by category and sector 11
Figure 5. Indicator evaluation in the Mekong basin by category and sector 12
Figure 6. Performance on DTLZ5(I,M) problems. Number of runs with the essential set correctly identified. ORA-RPSS consistently returns the correct essential indicators in a high number of runs... 30
Figure 7. Objectives hierarchy for the main objective of co-developing a water management plan for the Vietnamese Mekong Delta. Details on main and secondary objectives and indicators. Further details in Deliverable 1.1. Case study-specific stakeholder engagement roadmaps..... 39
Figure 8. Objectives Hierarchy for the Mekong Delta case study 40
Figure 9. Objectives Hierarchy for the Mekong Delta case study – Focus on the Healthy Delta dimension 40
Figure 10. Objectives Hierarchy for the Mekong Delta case study – Focus on the Liveable Delta dimension 41

Table 1. Summary of the highest and lowest rated indicators for the Jucar River basin 12
Table 2. Summary of the highest and lowest rated indicators for the Mekong basin 13
Table 3. Summary of main indicators' strengths and identified gaps 14
Table 4. Pre-selected indicators for the Jucar River Basin 19
Table 5. Comparison of thresholds per selected indicator and final choice 22
Table 6. Complexity of the algorithms as stated in the sources. Algorithm parameters (P) chosen for benchmarking tests. 29
Table 7. Setup of the four DTLZ5(I, M) problems..... 29
Table 8. Principles to be ensured when applying an ordering scale. 32
Table 9. Methods for eliciting weights to objectives 35

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by

Description of deliverable

1. Introduction

Water system management today faces unprecedented challenges due to climate variability, growing demands, and increasingly complex socio-economic interactions. In this context, developing a robust system of water indicators that not only captures the multi-dimensional nature of water systems but also supports effective decision-making is critical. The SOS-WATER project aims to address these challenges by providing a comprehensive and integrated indicator framework. Deliverable D4.2 aims to provide the theoretical groundwork for a joint system of water indicators by linking an in-depth assessment of existing indicators with innovative approaches to indicator aggregation.

The motivation for this work arises from the need to overcome limitations in traditional indicator systems. Conventional approaches have often been constrained by a narrow focus on individual aspects of water management, whether hydrological, meteorological, or socio-economic, resulting in fragmented assessments that fail to capture the full complexity of water systems. Recognizing these shortcomings, our work began with a multi-stage process for indicator assessment (described in detail in Section 2). Building on the initial selection and evaluation of indicators documented in D4.1, we conducted a comprehensive assessment of the existing water indicators. Interactive workshops with project partners were organized to identify which indicators are both practical and preferred across various contexts, while a detailed survey further assessed their suitability for water resource management. These multiple assessment steps revealed critical limitations in the current approaches, especially regarding their ability to capture spatial and temporal variability. In response, we have outlined key directions for developing innovative indicators that prioritize multi-dimensional integration, enhanced data resolution, and climate-sensitive frameworks. The feedback from these sessions was instrumental in restructuring the indicator list, ensuring that the emerging framework is not only comprehensive but also adaptable to the specific needs of different case studies. Additionally, preliminary thresholds, derived from an extensive literature review, have been proposed to support these new indicators, although they will require further refinement to align with the specific characteristics and contextual requirements of each case study.

While Section 2 establishes the foundation by addressing what the indicators should measure and where improvements are needed, Section 3 of the report takes a further step by exploring how these individual indicators can be aggregated into a coherent system. Given that water management issues are inherently multi-faceted, the ability to synthesize numerous indicators into a unified meta-indicator is essential for providing a clear, decision-relevant picture of system performance. However, the aggregation of indicators presents its own set of challenges, including the need to reduce data redundancy and balance conflicting information.

To tackle these issues, Section 3 is structured into two complementary parts. Section 3.1 focuses on data-driven methods for indicator reduction and aggregation. Techniques such as Non-negative Principal Component Analysis (NPCA) are employed to reduce a high-dimensional dataset into a more manageable set of essential indicators while preserving critical inf critical information. Other methods like PCSEA, ORA-RPSS, and δ-MOSS Greedy, are also discussed as means to identify and remove

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by Swiss Confederation

redundant indicators. These methods facilitate the transformation of a complex array of individual measurements into a compact set of essential indicators that can reliably reflect the overall state of the water system.

Section 3.2 introduces an alternative, yet complementary, perspective based on Multi-Attribute Value Theory (MAVT). While the data-driven methods provide a statistically robust mechanism for reducing dimensionality, MAVT addresses the challenge from a decision-making standpoint. By integrating both quantitative and qualitative assessments, MAVT enables the assignment of weights to individual indicators based on stakeholder preferences and the relative importance of various water management objectives. This approach leads to the formulation of a value function that aggregates individual indicator performances into a single composite index. Such a value function is particularly valuable in contexts where trade-offs coexist between different water system indicators.

In summary, this report presents a comprehensive approach to developing a system of water indicators that is both theoretically sound and practically applicable. By first diagnosing the shortcomings of current indicator systems and then introducing innovative measures to fill these gaps, we establish a solid foundation for improved water management. Building on this, the report details advanced aggregation mechanisms that integrate diverse data sources and stakeholder preferences, ultimately yielding a unified meta-indicator framework.

2. Indicator assessment and innovative indicators

2.1 Indicator gaps identification

2.1.1 Critical review of indicators

The assessment of identified indicators and the development of innovative indicators have been carried out through several stages. First of all, a comprehensive assessment of existing water indicators and the identification of critical gaps were conducted, as documented in deliverable D4.1. This task focused on an extensive literature review of the most commonly used and well-known indicators in water resource management. This review has been conducted using diverse information sources, including policy reports and guidelines, scientific articles, project outputs such as deliverables, legislative documents (e.g., Water Management Plans and water legislation), and others. The evaluation of these indicators focuses on their applicability, advantages, and limitations, considering several key factors, such as the variables involved in calculating the indicator, information sources, temporal and spatial scales, and their relevance to water management. The identified indicators are categorized based on their scope, including water resources indicators, water demand indicators, and indicators for integrated water resources management at basin scale (See deliverable D4.1). Following the initial selection and evaluation of key indicators, the selected indicators were further refined through a series of practical activities, interactive workshops, and collaborative engagements with project partners. A key activity was the interactive workshop conducted during the general assembly in Utrecht (M12), which facilitated in-depth discussions on the applicability and robustness of the previously identified indicators using tools such as the Miro board and paper diagrams (Figure 1 and Figure 2). The enriched dialogue allowed a nuanced understanding of which indicators are practical and preferred across different contexts. The

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Project funded by Swiss Confederation

feedback collected during the session led to restructuring the indicator list, ensuring its alignment with the specific needs of case studies and enhancing its contextual relevance.

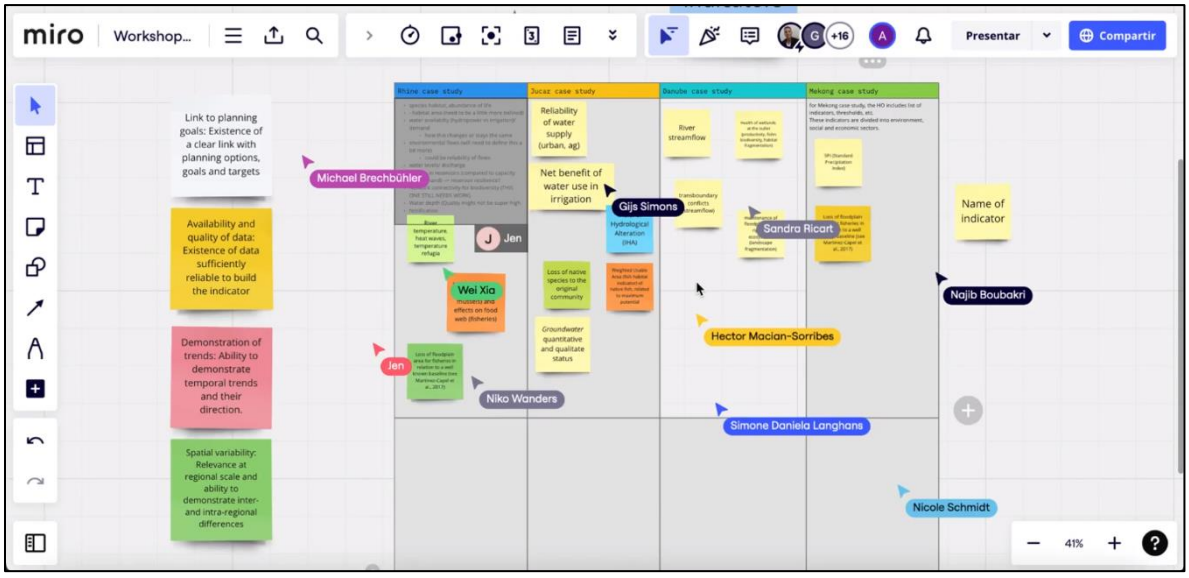

Figure 1. Interactive workshop discussion on indicator assessment using Miro Board.

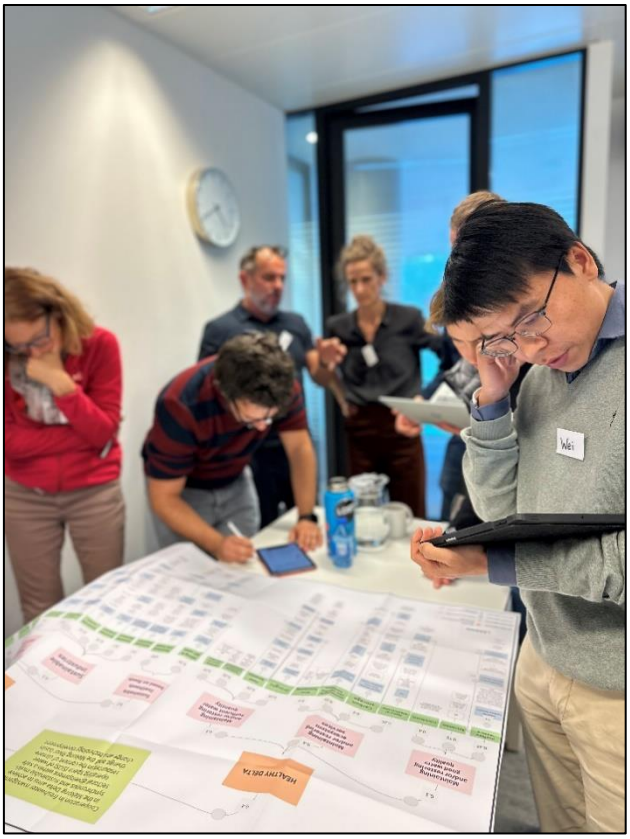

Figure 2. Interactive workshop at Utrecht for indicators discussion

Following identifying and evaluating of the initial set of indicators, integration meetings were conducted with WP1 to align and harmonize the indicators developed in WP4 with those from WP1. The WP1

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

indicators are organized based on an objective hierarchy comprising three primary levels: Objectives, sub-objectives, and indicators (Figure 3). At the objectives level, several goals are defined, such as ensuring good ecological status and enhancing the efficiency of water supply, aligned with SOS-WATER's overarching aims. The sub-objectives level decomposes these goals into specific outcomes, grouped by environmental, social, and economic dimensions. The indicators level assigns quantitative and qualitative metrics to each sub-objective. This collaboration with WP1 aimed to ensure computational feasibility and alignment with stakeholder requirements and overall project objectives, focusing on selecting indicators that provide unique and relevant insights while eliminating redundancy. These discussions facilitated refining the indicator list, ensuring consistency and coherence across work packages. As a result, a harmonized and well-defined set of indicators was established, enhancing the SOS-WATER project's analytical coherence and improving its analytical tools' applicability and utility.

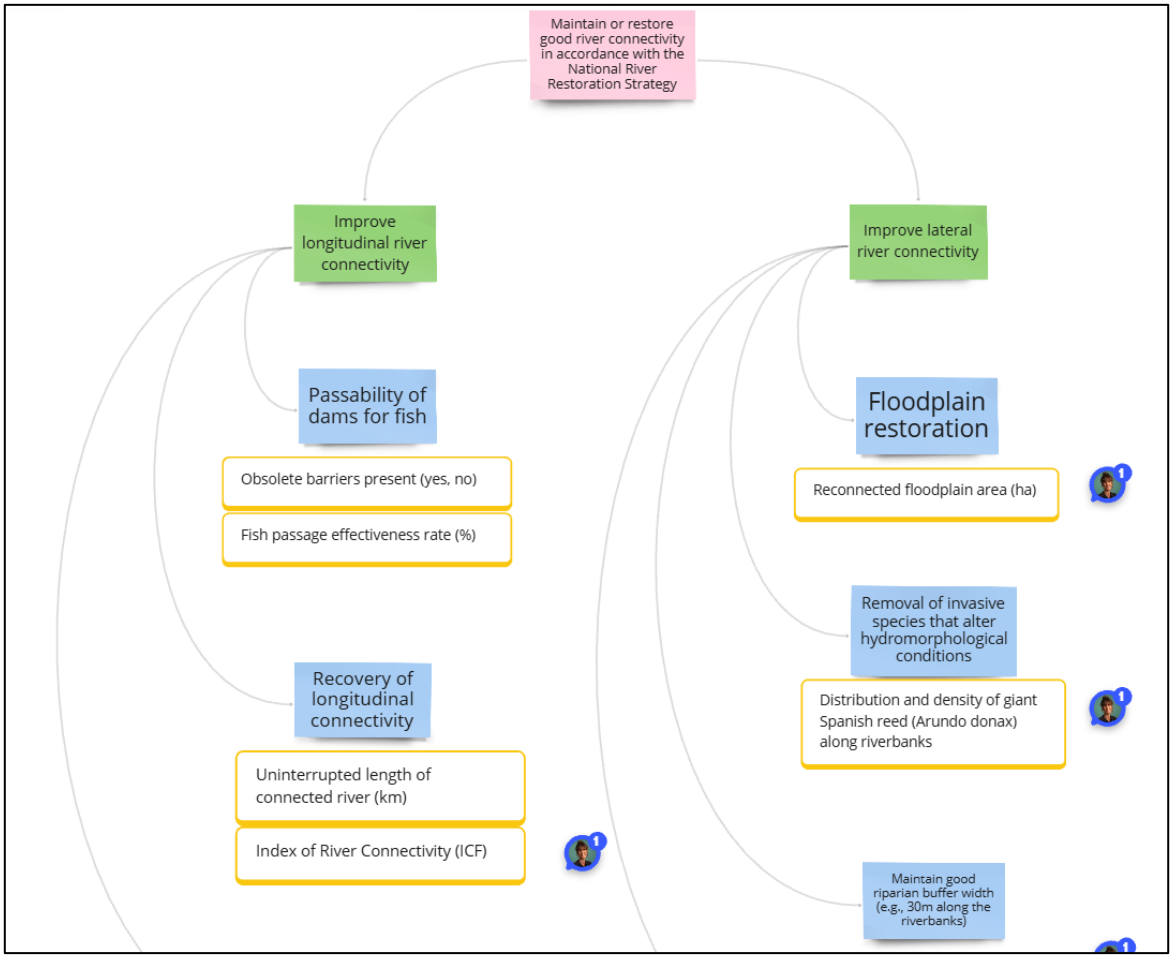

Figure 3. Objectives hierarchy and indicator framework from WP1

2.1.2. Development and distribution of an indicator evaluation survey

After identifying, cataloguing, and evaluating multiple water indicators used to assess the different dimensions and stress zones within the water system, a detailed survey was conducted to evaluate and classify potential indicators. The survey to assess identified indicators was designed to conduct a multicriteria evaluation of selected indicators, assessing their suitability for water resources

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by

management. The main objective of this survey was to identify gaps in the existing framework by assessing each indicator against key criteria, including:

- Relevance to policy, planning, and decision-making processes
- Data availability and quality to ensure feasibility in implementation
- Interpretability and clarity for effective communication with stakeholders
- Flexibility to adapt across different contexts and environmental conditions
- Long-term viability to remain effective under changing conditions
- Spatial variability to assess differences at regional and local scales
- Cost-effectiveness in terms of data collection, processing, and maintenance
- Policy alignment with existing regulations and management strategies

The survey was distributed among project partners responsible for completing it as a group. Stakeholders not directly involved in the project were not consulted during this process, due to three main reasons:

1. A mixed group gathering together modellers and stakeholders could have masked the results of the survey, which was intended to reflect the points of view of modellers
2. Stakeholders are not usually familiar with the indicators, apart from those that are directly related to their expertise or to their decision-making capacity. They might not be able to provide comprehensive evaluations of many indicators
3. The inclusion of stakeholders, considering their time constraints, would prevented the elaboration of a detailed survey.

In spite of not being directly involved in the filling of the survey, stakeholders from all case studies validated the selected indicators during the follow-up workshops held in all case studies.

Each partner group compiled responses based on internal discussions among team members, consolidating their collective expertise into a single response per group. The assessment was shared in an excel-based format on June 26, 2024. Along with the survey file, partners received a detailed presentation outlining the objectives, methodology, and instructions for completion. The excel sheets contained interactive and dynamic questions, allowing respondents to assess indicators according to pre-defined criteria systematically. The deadline for submission was July 19, 2024, providing sufficient time for group discussions and responses.

Each indicator was assessed using a two-step rating process: first, each criterion was rated on a scale from 1 to 5, where 1 represents low effectiveness or relevance, and 5 represents high effectiveness or relevance. Then, each indicator was evaluated across all criteria using the same 1 to 5 scale. The final indicator rating was calculated as a weighted average of the criteria scores. The threshold for classifying indicators based on their evaluation across criteria was 200. Indicators with a rating above 200 were considered relevant, while those scoring below this threshold were considered less relevant. The criteria and survey questions are detailed in the supplementary material.

By systematically reviewing the final list of indicators, the survey aimed to pinpoint areas where current assessments fall short and highlight opportunities for integrating more effective and innovative indicators. This evaluation provides a critical foundation for addressing identified gaps and aligning the

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by

indicators with the diverse needs of water management systems across varying contexts and stress zones. The survey was instrumental in classifying and selecting the most effective indicators for assessing various dimensions across case studies.

2.1.3. Analysis of survey results and indicator assessment

Feedback from the survey responses guided the selection of the most pertinent indicators for each case study and provided critical insights into its evaluation. Indicators were assessed across three main categories: water resources indicators, water demand indicators, and integrated water resource management (IWRM) indicators at the basin scale.

For the Jucar River basin (Figure 4), the water resources indicators category received particularly high ratings for meteorological (Meteo), hydrological (Hydro), groundwater (Gw), surface water (Sw), and water stress and use intensity indicators, reflecting their importance in assessing resource availability and stress levels. In the water demand indicators category, agricultural and socio-economic demand indicators were rated the highest, highlighting the significant role of these sectors in water allocation and planning for the basin and within the IWRM indicators category, reliability, resiliency, exploitation, and sustainability indicators received the top ratings, demonstrating their critical relevance to integrated water management approaches.

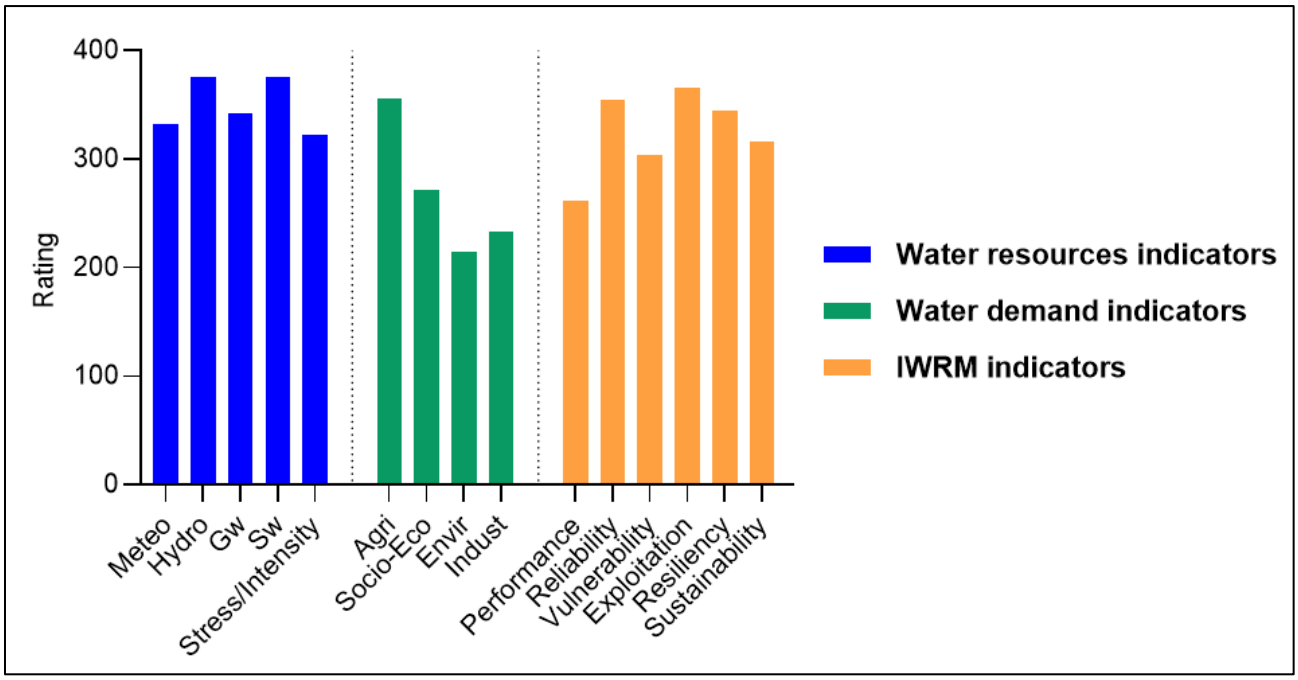

Figure 4. Indicator evaluation in the Jucar River basin by category and sector

Through a detailed indicator-level analysis of the survey results, 64 indicators were evaluated, with ratings ranging from 89 to 386. For the Júcar river case study (Table 1), the survey results highlight the importance of hydrological and meteorological indicators. The highest-rated indicators include Low Flow Index (LFI) and Drought Duration, both scoring 386, reflecting their critical role in assessing water availability and drought impacts in the basin. Additionally, the Standardized Runoff Index (SRI) and Consecutive Dry Days (CDD) received high ratings, underlining their relevance for detecting hydrological

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by

extremes. Conversely, indicators such as the River Sinuosity Index and Channel Width-to-Depth Ratio, both scoring 89, were rated low, indicating limited applicability for the Júcar river's management needs. The relatively low score of the Industrial Water Footprint (rating 179) suggests that industrial water use is of secondary concern compared to agricultural and ecological indicators in this basin. This ranking underscores the priority of drought and water availability indicators, aligning with the region's water scarcity challenges.

Table 1. Summary of the highest and lowest rated indicators for the Jucar River basin

Indicators	Rating
Standardized Precipitation Index (SPI)	363
Standardized Runoff Index (SRI)	372
Low Flow Index (LFI)	386
Drought Duration	386
Consecutive Dry Days	379
Crop Water Stress Index (CWSI)	361
River Sinuosity Index	89
Channel Width-to-Depth ratio	89
Average Fish Index of Biotic Integrity (F-IBI)	89
Industrial Water Footprint	179

Similarly, in the Mekong basin (Figure 5), water resources indicators such as hydrological, groundwater, and surface water indicators were rated highly, indicating their importance in addressing the water challenges of this region. In the water demand category, agricultural, socio-economic and industrial demand indicators received notable ratings, underscoring the balance needed between these three sectors. For the IWRM indicators category, all indicators were consistently highly rated, indicating their collective importance in achieving integrated and adaptive water management strategies in the basin.

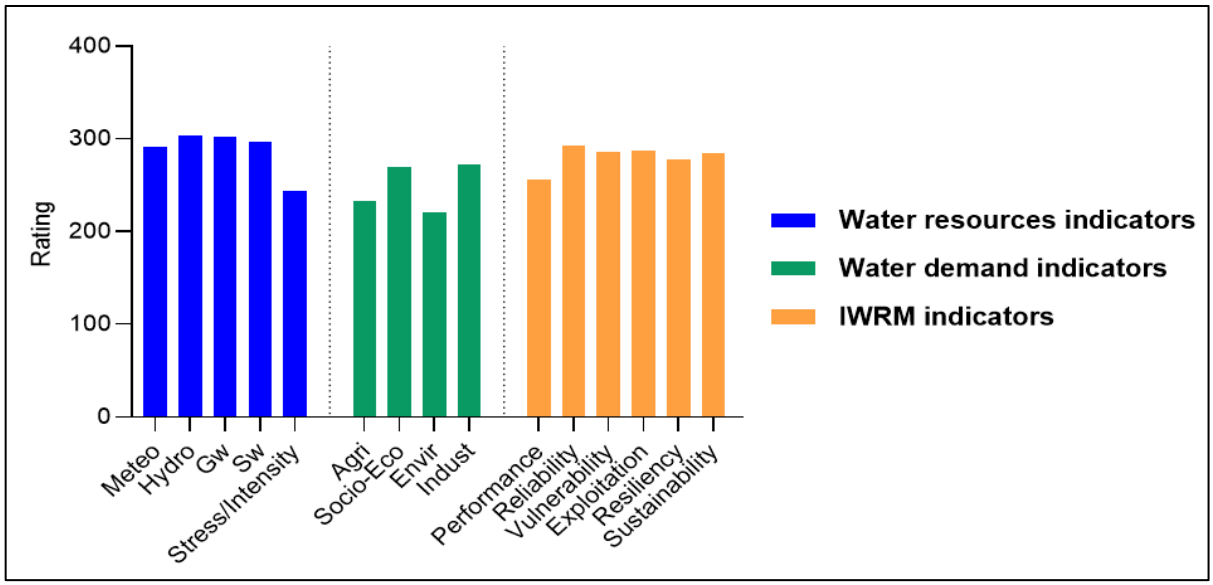

Figure 5. Indicator evaluation in the Mekong basin by category and sector

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

The indicators rating for the Mekong case study range from 74 to 321 (Table 2). The highest-rated indicators emphasize hydrological and water quality metrics, reflecting their importance for water resource management in the basin. The Water Stress Index (WSI), rating 321, and Rainfall Anomaly Index (RAI), rating 307, scored highly, highlighting water availability and variability concerns. In addition, the Aquifer Recharge Rate (ARR), rating 306, and Water Quality Index (WQI), rating 302, were recognized as crucial for groundwater and water quality assessments. However, the lowest-rated indicators include the Average Fish Index of Biotic Integrity (F-IBI) and Warm Spell Duration Index (WSDI), rating 74 and 97, respectively, suggesting a lower perceived relevance of ecosystem health and temperature anomalies in comparison to hydrological factors. These indicators ranking underscore the Mekong river's primary concerns related to hydrological variability, water stress, and groundwater sustainability.

Table 2. Summary of the highest and lowest rated indicators for the Mekong basin

Indicators	Rating
Standardized Precipitation Index (SPI)	293
Heavy Precipitation Days (HPD)	300
Rainfall Anomaly Index (RAI)	307
Water Quality Index (WQI)	302
Water Stress Index (WSI)	321
Aquifere Recharge Rate (ARR)	306
Average Fish Index of Biotic Integrity (F-IBI)	74
Beneficial Water Consumption	168
Non-beneficial Water Consumption	168
Warm Spell Duration Index (WSDI)	97

Based on this analysis of indicators, the results provide valuable insights into the interconnections, synergies, and complementarities among different indicators and sectors. These insights guide the development of innovative indicators, incorporating complementary indicators, and refining existing ones to ensure they comprehensively address linked aspects of water resource management. Additionally, this analysis identified further gaps and indicators that needed to be included to meet specific requirements and contextual challenges. By identifying these relationships and addressing the gaps, the analysis enhances the overall coherence and effectiveness of the indicator system, creating the foundation for a more integrated and robust assessment framework.

2.1.4. Identified gaps

From the survey results, a clear gap was not identified in the indicators collected in D4.1. Demand indicators performed slightly below the rest, in particular for ecosystems, but their marks were not bad enough to consider that a gap has been identified. Consequently, we identified gaps at the individual indicator level, reviewing their main features to locate drawbacks. Evaluating the current set of water indicators reveals several critical gaps that limit their effectiveness in providing a comprehensive understanding of water resources management. These gaps are particularly evident when considering the complex interplay between meteorological, hydrological, socioeconomic, and environmental factors. The following table summarizes the strengths and identified gaps of the indicators.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by Swiss Confederation

Table 3. Summary of main indicators' strengths and identified gaps

Indicators	Strengths	Identified gaps
1- Meteorological Indicators		
Standardized Precipitation Index (SPI)	Provides early warning signals for droughts by measuring precipitation variability	• **Limited Scope**: These indicators primarily focus on precipitation and evapotranspiration but do not account for other critical hydrological variables such as soil moisture, groundwater recharge, or surface water flow. This limits their ability to provide a holistic view of water availability and drought impacts. • **Spatial and Temporal Resolution:** The applicability of SPI and SPEI is often constrained by the spatial and temporal resolution of available data. For instance, SPI may not capture local variations in precipitation, and SPEI requires accurate evapotranspiration data, which is often estimated rather than measured. • **Climate Change Adaptation:** Both SPI and SPEI assume stationarity in climate conditions, which may not hold true under climate change. This limits their effectiveness in predicting future water resource availability and drought risks.
Standardized Precipitation-Evapotranspiration Index (SPEI)	Incorporates both precipitation and evapotranspiration, making it more suitable for arid regions and climate change contexts	
Heavy Precipitation Days (HPD)	Tracks the frequency and intensity of heavy rainfall events, useful for flood risk assessment	
2- Hydrological Indicators		
Standardized Runoff Index (SRI)	Measures hydrological drought conditions by assessing runoff variability	• **Regional Variability**: The applicability of SRI and LFI is limited by regional differences in hydrological behaviour and data availability. For example, SRI may not accurately reflect the hydrological conditions in regions with complex river systems or significant human interventions. • **Human Interventions**: These indicators do not account for human activities such as water abstraction, reservoir management, or land-use changes, which can significantly alter hydrological regimes. • **Data Availability**: Accurate runoff data, required for SRI and LFI, is often lacking in data-scarce regions, limiting their applicability.
Low Flow Index (LFI)	Tracks water deficits in rivers during low flow periods, providing insights into hydrological droughts	

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

3- Reservoirs and Groundwater Indicators

Groundwater level Index (GLI)	Tracks changes in groundwater levels, essential for assessing aquifer status	• **Data Availability and Monitoring**: The effectiveness of GLI and ARR depends on continuous monitoring, which is often lacking in many regions. This limits their ability to provide real-time assessments of groundwater sustainability. • **Integration with Surface Water**: These indicators do not explicitly account for interactions between groundwater and surface water systems, which are critical for integrated water resource management. • **Water Quality:** GLI and ARR focus on water quantity but do not address water quality issues, such as contamination or salinization, which can significantly impact groundwater sustainability.
Aquifer Recharge Rate (ARR)	Measures the rate at which aquifers are replenished, crucial for sustainable groundwater management	

4- Surface Water Indicators

Surface Water Supply Index (SWSI)	Assesses surface water availability, particularly in snowpack runoff areas	• **Water Quality and Ecological Health**: SWSI focuses on water quantity but does not account for water quality or the ecological health of surface water bodies. This limits its ability to assess the overall sustainability of surface water resources. • **Reservoir Management:** While SWSI considers reservoir storage, it does not account for the complex management practices that influence reservoir operations, such as flood control, hydropower generation, and environmental flow releases.

5- Water Stress and Use Intensity Indicators

Multivariate Standardized Drought Index (MSDI)	Combines precipitation and soil moisture data to assess drought conditions	• **Sectoral Integration:** These indicators focus on specific aspects of water stress (e.g., agricultural drought) but do not integrate across multiple sectors (e.g., urban, industrial, environmental). This limits their ability to provide a holistic view of water stress across different water users. • **Water Quality and Ecosystem Health**: MSDI, PDSI, and SMAI focus on water quantity but do not address water quality or ecosystem health, which are critical for sustainable water resource management.
Palmer Drought Severity Index (PDSI)	A long-standing drought index that considers temperature and precipitation	
Soil Moisture Anomaly Index (SMAI)	Measures deviations in soil moisture from normal conditions, useful for agricultural drought assessment	

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

		• **Complexity and Data Requirements**: These indicators can be complex to calculate and require extensive data, which may not be available in all regions.

6- Socioeconomic Indicators

Economic losses due to water-related disasters (ELS)	Measures the economic impact of water-related disasters, such as floods and droughts	• **Environmental and Social Dimensions**: These indicators focus on economic impacts but do not account for the environmental and social dimensions of water demand. For example, they do not consider the ecological impacts of water use or the social equity of water allocation. • **Sectoral Coverage**: While these indicators provide insights into specific economic sectors, they may not cover all categories of water users.
Water use intensity by economic activities	Assesses the efficiency of water use across different economic sectors	

7- Agricultural Water Demand Indicators

Crop Water Stress Index (CWSI)	Measures water stress in crops, aiding in irrigation scheduling	• **Complex Soil - Plant - Atmosphere Interactions**: These indicators focus on crop water stress but do not account for the complex interactions between soil, plants, and the atmosphere, which can significantly influence water demand. • **Climate Change Adaptation:** While WSDI provides insights into the impact of warm spells on crop growth, it does not account for other climate change impacts, such as changes in precipitation patterns or increased frequency of extreme weather events.
Warm Spell Duration Index (WSDI)	Tracks the duration of warm spells, which can impact crop growth and water demand	

8- Environmental Water Demand Indicators

Mean Species Abundance (MSA)	Measures biodiversity in aquatic ecosystems, providing insights into ecological health	• **Meteorological and Hydrological Factors**: These indicators focus on ecological health but do not explicitly account for the

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Suitable Habitat Area Index (SHAI)	Assesses habitat availability for species, useful for conservation planning	meteorological and hydrological factors that influence habitat availability and environmental flow. • **Integration with Water Resource Management**: MSA and SHAI are often used in isolation from water resource management practices, limiting their ability to inform integrated water resource management decisions.
9- Industrial Water Demand Indicators		
Industrial Water Footprint (IWF)	Assesses water consumption in industrial processes, useful for sustainability reporting	• **Sectoral Coverage**: IWF focuses on industrial processes but may not cover all categories of water users. • **Water Quality and Pollution**: IWF measures water consumption but does not address water quality issues or the potential for pollution associated with industrial activities.
10- Joint Resource-Demand Indicators		
Reliability, Vulnerability, Exploitation, Resilience, Sustainability, and Performance Indicators	Provide a comprehensive assessment of water resource management	• **Complexity and Data Requirements**: These indicators can be complex to calculate and require extensive data, which may not be available in all regions. • **Sector-Specific Issues**: While these indicators provide a holistic view of water resource management, they may not address sector-specific issues, such as the unique challenges faced by agriculture, industry, or urban water users.

The identified gaps highlight the necessity for innovative, adaptive indicators that address the complexities of water management across diverse spatial and temporal scales. Key areas for improvement include the integration of water quality across all sectors, the consideration of cross-sectoral environmental impacts, the inclusion of social equity and economic trade-offs in water demand assessments, and the development of indicators that bridge supply and demand dynamics to evaluate system resilience under both long-term and extreme scenarios. Addressing these gaps will enable a more holistic and adaptive approach to water resource management, ensuring sustainability and resilience in the face of growing water challenges.

2.2 Innovative indicators

To address the gaps identified in current water resource indicators, the development of innovative, adaptive indicators must prioritize multi-dimensional integration, enhanced data resolution, and

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

climate-sensitive frameworks. For meteorological indicators, integrating soil moisture, groundwater levels, and surface water flow with precipitation and evapotranspiration data will enable a holistic assessment of water availability, while leveraging high-resolution remote sensing and climate projections will improve their sensitivity to non-stationary climate conditions (Vicente-Serrano et al., 2022b, 2022a). Hydrological indicators require regional calibration to account for spatial variability and human interventions (e.g., water abstraction and reservoir operations), supported by advanced monitoring networks to enhance runoff data quality in underrepresented regions (Konar et al., 2016). Groundwater sustainability can be better captured by combining groundwater recharge rates and depletion trends with water quality parameters (e.g., salinity, contaminants), while integrating surface-groundwater interactions into indices will reflect systemic resource availability (Famiglietti, 2014). For surface water systems, indicators must incorporate reservoir management practices (e.g., flood control, environmental flow releases) and water quality metrics (e.g., nutrient loads, pollutants) to balance operational and ecological needs (Tramblay et al., 2020).

Water stress and demand indicators should adopt cross-sectoral approaches, linking agricultural, industrial, and urban water use with soil-plant-atmosphere dynamics and climate-driven projections to predict future demand under changing conditions (Kim et al., 2020). Socioeconomic indicators must expand coverage to include marginalized users (e.g., smallholder farmers) and integrate social equity metrics (e.g., access disparities, economic trade-offs) (Sullivan, 2002). Environmental flow indicators need to embed meteorological and hydrological drivers (e.g., precipitation extremes, flow regimes) into habitat suitability models, ensuring ecological thresholds inform water allocation decisions (Richter et al., 2012). Industrial water use assessments must extend beyond consumption to include pollution footprints (Mekonnen and Hoekstra, 2015), while sector-specific indices should address unique challenges (e.g., circular water use in manufacturing). Finally, integrated resource-demand indicators must adopt simplified, scalable frameworks that bridge supply-demand gaps through probabilistic models and multi-scale spatial-temporal aggregations, enabling resilience evaluations under long-term and extreme scenarios.

2.3 Preliminary thresholds

A threshold is a limiting value associated with an indicator that establishes the frontier between safety and unsafety in terms of system performance. The value of a threshold depends primarily on the features of each water system. Those can be mainly divided into:

- Resource characteristics: meteorological and hydrological patterns, regulation facilities such as reservoirs and aquifers, existence of water transfers, non-conventional water source availability, system connectivity, etc.
- Demand characteristics: living population and population pattern, agricultural surface, crops planted, industrial demand, hydropower facilities, etc.
- Ecosystem features: number of protected ecosystems related with water, degree on which ecosystems depend on water, current status, ecosystem relevance, etc.
- Other features: contribution of each system to the regional and/or national gross domestic product, legislative frameworks, social perception on water uses, power dynamics between users and between users and administration, stakeholder profiles, etc.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Project funded by Swiss Confederation

The combination of all these features drives, for each system, a unique vision on what is safe or unsafe. Water systems with distinct surface storage capacities or with large portfolios of available water sources (surface, ground, transfers, reuse, desalination) might be more tolerant to extreme droughts than those with limited water supply choices. Similarly, systems with an intensive water use would be more vulnerable to decreases in hydrological discharges than others whose water consumption remains at low levels compared to water resources. Setting appropriate thresholds require an in-depth analysis at the case study level. Consequently, the rest of this sub-section will demonstrate the setting of thresholds for a case study, the Jucar River Basin.

2.3.1. Identification of indicators and information sources on thresholds

Based on the indicator evaluation survey and the available mathematical models, a pre-selection of indicators to be used to compute the Safe Operating Space in the Jucar River Basin is provided in Table 2.4. For each indicator, an analysis was carried out to determine if a threshold is provided by law, by bibliography or by the expert knowledge of stakeholders or modellers. In the Jucar River Basin, both the Jucar River Basin Management Plan (JRBMP) and the Jucar Drought Management Plan (JDMP) set thresholds to some of them, attached to either the activation of drought management measures (JDMP) or to the planning and construction of measures to enhance water supply or reduce water demand (JRBMP). However, these thresholds set by law might not imply that safety is achieved if not or only occasionally violated.

Table 4. Pre-selected indicators for the Jucar River Basin

Indicator	Threshold information source				Spatial scale	Observations
	Law	Stks	Mod.	Biblio		
Standardized Precipitation Index (SPI)	X			X	Sub-basin	Threshold established by the Jucar Drought Management Plan JDMP (10^{th} percentile of SPI12 time series)
Standardized Precipitation-Evapotranspiration Index (SPEI)				X	Sub-basin	Stakeholders are not very familiar with this indicator
Standardized Runoff Index (SRI)	X			X	Particular locations (JDMP considers 5 gauges)	Threshold defined to normalized 3-month aggregated streamflows, part of the official Scarcity State Index of the JDMP
Low Flow Index (LFI)				X	Sub-basin	Shows overlaps with the SRI

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Indicator	Threshold information source				Spatial scale	Observations
	Law	Stks	Mod.	Biblio		
Standardized Water Level Index (SWI)	X			X	Main reservoirs (Alarcon, Contreras, Tous) and Forata	Threshold defined to normalized storages, part of the official Scarcity State Index of the JDMP
Crop Water Stress Index (CWSI)			X	X	Agricultural demand	
Water use productivity		X	X	X	Demand	Defined based on benefits, benefits per unit of water applied, or crop production
Groundwater Level Index (GLI)	X			X	Particular locations (2 piezometers in Mancha Oriental aquifer, 1 in Requena-Utiel aquifer)	Threshold defined to normalized piezometric levels, part of the official Scarcity State Index of the JDMP
Surface Water Supply Index (SWSI)			X	X	Demand	Shows overlaps with reliability
Water Use Intensity	X			X	Demand	The Jucar River Basin Management Plan (JRBMP) provides estimations only for hydropower plants
Magnitude of Monthly Flow Alteration (MMFA)	X			X	Reach with minimum streamflow	Criteria and thresholds established by the JRBMP, used together with the suitable habitat area to calculate minimum streamflows to be imposed
Suitable Habitat Area	X	X	X	X	Reach with minimum streamflow	Criteria and thresholds established by the JRBMP, used together with the MMFA to calculate minimum streamflows to be imposed
Water Supply Reliability	X	X	X	X	Demand	Criteria and thresholds established by the JRBMP

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

2.3.2. Comparison of thresholds provided by different sources and final choice

From the pre-selected indicators, a grouping of them was made according to the information sources and overlaps:

1. Indicators for which stakeholders could provide thresholds (regardless of thresholds provided by law): water use productivity, suitable habitat area, water supply reliability
2. Indicators for which law could provide thresholds, but not stakeholders: SPI, SRI, SWI, GLI, water use intensity, MMFA
3. Indicators for which thresholds could be provided by modellers' knowledge of the system or bibliography: CWSI
4. Indicators for which only bibliography could provide thresholds: SPEI
5. Indicators that were considered to overlap other indicators: LFI, SWSI

From these groups, indicators belonging to the first group were considered to be the preferred ones to infer the safe operating state in the Jucar River Basin. The rest of indicators could be adequate to evaluate the system performance compared to law requirements (second group and first group if only law thresholds were considered); for research purposes (third and fourth group); or not required due to overlapping other ones (fifth group). Consequently, the final choice of indicators was:

- Water use productivity: this indicator was primarily defined based on benefits, although some stakeholders also expressed the convenience of evaluating them based on crop productions.
- Suitable habitat area: defined as a percentage of the maximum possible suitable habitat area per reach with minimum streamflow downstream of a dam and with fish habitat models.
- Water supply reliability: defined as volumetric reliability.

Concerning water supply reliability, stakeholders from the farming sector considered that its response would be similar to water use productivity, which they would prefer. Table 2.5 compares the thresholds from law and stakeholders and makes a final choice. It should be noted that thresholds proposed by stakeholders are still preliminary thresholds, to be confirmed in further stakeholder workshops. It should also be noted that the water supply reliability indicator, as reported by the Jucar River Basin Agency, is currently achieved at the expenses of aquifer overexploitation (that the model would not acknowledge). This might imply that the current situation is unsafe and the application of this threshold might result into a fully unsafe space. Moreover, stakeholders from the agricultural sector would prefer the water use productivity over this indicator to define their safe or unsafe status. Consequently, the representativity of water supply reliability as an indicator of safety or unsafety will be assessed in future workshops with stakeholders and based on modelling results.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by

Table 5. Comparison of thresholds per selected indicator and final choice

Indicator	Law threshold	Threshold proposed by stakeholders	Threshold chosen
Water use productivity	None	Average annual benefits should be at least 80% of the current ones.	Stakeholders
Suitable habitat area	30% for a chosen fish species. It could be lowered down to 25% in case of non-protected areas under drought conditions. No violations allowed	50% for the most restrictive fish species at the monthly scale should be fulfilled at least 80% of time	Stakeholders
Water supply reliability	For urban demands: • Monthly deficits should not be higher than 10% of monthly demand • Deficits in 10 years should not be higher than 8% of annual demand For agricultural demands: • Deficit in 1 year should not be higher than 50% of annual demand • Deficit in 2 years should not be higher than 75% of annual demand • Deficit in 10 years should not be higher than 100% of annual demand For the nuclear power plant of Cofrentes: • No deficits allowed No violations are allowed.	None	Law. See the body text for considerations on this indicator

2.3.3. Spatio-temporal aggregation and uncertainty on thresholds

Each threshold, as well as the computation of the indicator to which it is associated, should refer to a spatial domain and a temporal scale. In the case of the Jucar River Basin, the temporal scales of the indicators and thresholds selected is provided in Table 2.5, consisting in temporal aggregation and violation ratio. For example, water use productivity is aggregated at the annual scale with no violation, while the suitable habitat area is defined at the monthly scale with an allowed violation of 20%. Water

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

supply reliability, on the other hand, is defined for three temporal aggregations, being each one of them associated with its own violation ratio.

In terms of spatial aggregation, water use productivity is primarily defined per agricultural demand and hydropower plant, water supply reliability is defined per consumptive demand, while habitat suitable area refers to each particular modelled river reach with habitat models and to each fish species. However, for agricultural demands, an aggregation was performed to simplify the evaluation of the safe operating space, grouping agricultural demands in three areas: Mancha Oriental, Jucar-Turia Canal and Traditional demands. This grouping was possible given the similarities shown by these demands in terms of crops planted, water sources and constraints, as well as water rights. Traditional demands, in the downstream part of the basin, rely entirely on surface water and citrus crops, while they possess senior rights. Jucar-Turia Canal combines surface and groundwater, is placed in the downstream basin, relies on citrus crops and its supply is subject to water availability. Mancha Oriental, on the other hand, uses primarily groundwater, although being entitled to some amount of surface water, it is placed in the middle basin, relies mostly on annual crops and is currently being subject to a reduction in its pumping allowance to guarantee the sustainability of the Mancha Oriental aquifer. In case of hydropower plants, a straightforward aggregation was performed by adding the benefits provided by all plants, since they belong to the same operator and energy market, being operated in full coordination.

It should be noted that all thresholds adopted are subject to uncertainty in three main ways. First of all, spatio-temporal aggregation obviously induces uncertainties associated with differences among the aggregated demands. Secondly, the adoption of a unique threshold for each indicator neglects the diversity noticed among demands, which might react differently to changes in benefits and supply levels. In particular, the benefit levels that each agricultural demand can bear depend on its farming structure and the capacity of farmers to resist temporary shocks on economic returns. Moreover, each crop responds in a different way to water deficits. In the Jucar River Basin, the adoption of a unique threshold for all demands is a temporary oversimplification that will be refined in further workshops with stakeholders in order to set a threshold value that is representative for each demand group. Finally, each threshold sets a frontier between safe and unsafe that does not acknowledge any difference within each zone, while this “all or nothing” approach could be too strong in some situations. This issue could be alleviated, e.g., by adopting a multi-threshold approach, in which further distinctions within the safe or the unsafe areas could be established.

While the previous discussion focused on the causes of uncertainty in threshold definition, it is important to clarify to what extent such uncertainty can be estimated. Uncertainty in indicators can be more readily quantified through model-based approaches—for instance, by running simulations under alternative model parameterizations or different climate projections (GCMs), resulting in a distribution of indicator values.

In contrast, uncertainty in thresholds is inherently more difficult to estimate. Thresholds are often derived from expert judgment or stakeholder consultation rather than direct empirical evidence, and actual threshold-crossing events (e.g. system failures) are rarely observed in water systems. Moreover, the complex and often nonlinear behavior of socio-hydrological systems can blur the distinction between gradual change and abrupt tipping points. As a result, the precise location of a threshold is uncertain and may vary depending on the conceptual framework, stakeholder perceptions, or natural variability.

Funded by the European Union

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Project funded by Swiss Confederation

Nevertheless, this uncertainty can still be characterized—for example, by using multiple threshold definitions (multi-threshold analysis), by eliciting uncertainty ranges from experts, or by performing sensitivity analyses on the criteria used to define safe and unsafe conditions. These approaches will be explored in the next project phases to better represent the uncertainty surrounding thresholds, and how this uncertainty interacts with uncertainty in indicator values.

3. Indicator aggregation mechanisms

3.1 Data-driven methods

This section introduces several data-driven methods that can be employed for indicator aggregation within the framework of a system of indicators. These methods are particularly valuable in addressing challenges such as high dimensionality, redundancy, and interdependence among indicators, which are common in complex datasets. By leveraging these techniques, analysts can streamline indicator systems, reduce redundancy, and enhance interpretability without losing critical information.

To illustrate the utility of these methods, we draw parallels to many-objective optimization (MaOO) problems, where multiple objective functions must be simultaneously optimized. In these problems, the objective functions often display redundancy, correlation, and even conflict, much like the relationships observed among indicators in complex systems. By viewing the objectives in MaOO as analogous to indicators, it becomes easier to conceptualize the role of these data-driven methods in reducing the dimensionality of indicator datasets. For the remainder of this text, the term “indicator” will be used as a synonym for “objective” in the optimization context.

This analogy underscores the relevance of these methods in tackling real-world challenges. Just as dimension reduction techniques in MaOO help identify a more compact yet informative set of indicators to guide decision-making, similar approaches can aid in creating simplified and meaningful representations of indicator systems. This not only improves computational efficiency but also provides clearer insights for decision-makers, enabling more effective evaluations and strategies in complex systems.

3.1.1. Non-negative Principal Component Analysis

Non-negative Principal Component Analysis (NPCA) is a variant of the widely used Principal Component Analysis (PCA) (Zass and Shashua., 2006), designed to address specific constraints and challenges in data analysis. While regular PCA identifies principal components as linear combinations of the original variables that maximize variance, NPCA adds a critical condition: the components must be non-negative. This ensures that the aggregated indicators derived from the method maintain interpretability, as negative values might not make sense in many real-world contexts, such as physical, economic, or environmental indicators.

Originally introduced by Zass and Shashua (2006), NPCA was later applied to a water resource management (WRM) problem within the many-objective optimization (MaOO) domain by Giuliani et al. (2014). The method has proven effective in situations where the variables or indicators under consideration must be combined in a manner that preserves their non-negativity while capturing the essential patterns and relationships in the data.

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

In the context of indicator aggregation, NPCA works by identifying new variables—referred to as aggregated indicators—that are linear combinations of the original indicators. These aggregated indicators are constructed such that they capture a specified fraction of the total variance (explained variance, EV) in the dataset. This process reduces the dimensionality of the indicator system, simplifying its structure while retaining as much critical information as possible.

The fraction of EV to retain is determined by selecting the number of principal components (PCs) . For example, setting a high EV threshold ensures that most of the variability in the data is captured, preserving more detailed information. However, this often results in a larger number of aggregated indicators, meaning that dimensionality reduction is less pronounced. Conversely, setting a lower EV threshold results in fewer aggregated indicators, achieving greater dimensionality reduction but at the potential cost of omitting some variability in the data.

The trade-off lies in balancing the need for dimensionality reduction with the requirement to maintain the integrity and usefulness of the data. A well-chosen EV threshold minimizes the risk of significant information loss while ensuring that the number of aggregated indicators is manageable for subsequent analysis and decision-making.

3.1.2. PCSEA

The Pareto Corner Search evolutionary algorithm has been introduced by Singh et al. (2011). It uses an evolutionary search to find the corners of the Pareto Front (PF). Then it estimates the redundancy of the indictors by introducing the metric $R \in [0, 1]$, which measures the change in the number of non-dominated solutions from the reference set F (computed using the corners of the PF from the step before) and the set of relevant indictors F_R, when one of the indicators in F_R is omitted. A value of R close to 1 indicates a relatively small change in the number of solutions after omitting the indicator, therefor the indicator is considered redundant and permanently removed from the F_R. Initially F_R = F, then sequentially indicators are removed until $R < C$ for all indicators in F_R. The threshold C is defined by the user. The sequence of omitting the indicators can have an impact on the final result, which is one of the drawbacks of the method.

In the context of indicator aggregation, this method can identify and remove redundant indicators by evaluating their impact on the diversity of solutions. The metric R, quantifies redundancy by measuring the change in non-dominated solutions when an indicator is omitted. A high value indicates redundancy, prompting removal.

3.1.3. ORA-RPSS

The Objective Reduction Algorithm using Representative Pareto Solution Search (ORA-RPSS), introduced by Guo et al. (2016), is a sophisticated method designed for reducing the dimensionality of multi-objective datasets while retaining the most relevant and non-redundant indicators. It is particularly effective in scenarios where datasets exhibit significant redundancy and conflict among indicators, which can complicate analysis and decision-making processes.

The method begins by selecting a representative set of non-dominated solutions from the Pareto Front (PF). This is accomplished using the Multi-Objective Evolutionary Algorithm based on Decomposition (MOEA/D), a widely used optimization technique. Once the representative solutions are identified, they

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by Swiss Confederation Federal Department of Economic Affairs, Education and Research EAER State Secretariat for Education, Research and Innovation SERI

are analysed to detect redundancy and conflict among indicators. Redundancy refers to instances where multiple indicators convey overlapping or similar information, whereas conflict arises when indicators are negatively correlated or inherently compete against each other. Identifying these relationships is critical to reducing dimensionality without compromising the dataset's integrity.

ORA-RPSS employs a redundancy metric R, similar in concept to the one used in PCSEA, to quantify the overlap or similarity between pairs of indicators. Pairwise comparisons are conducted to evaluate these relationships, enabling the method to systematically isolate and eliminate redundant indicators. This process ensures that only the most distinct and informative indicators are retained.

Based on the redundancy analysis, ORA-RPSS constructs a subset of non-redundant indicators. This construction is iterative, meaning the method refines the subset across multiple iterations to achieve an optimal balance between indicator diversity and relevance. The refinement process is guided by a user-defined threshold parameter C, which controls the level of redundancy tolerated in the final subset. A lower C value enforces stricter reduction criteria, resulting in fewer, more diverse indicators, while a higher C value allows for greater redundancy, retaining a larger set of indicators. The iterative refinement process ensures that the final set of aggregated indicators balances diversity and relevance. This is particularly important in tasks involving high-dimensional datasets, where simplifying the dataset must not come at the expense of losing critical information or introducing bias. By preserving the essential characteristics of the dataset while eliminating unnecessary complexity, ORA-RPSS facilitates more effective analysis and decision-making.

ORA-RPSS offers several advantages, including effective dimensionality reduction, enhanced interpretability of the remaining indicators. Its ability to simplify complex datasets without sacrificing critical trade-offs makes it an invaluable tool in many-objective optimization and indicator aggregation tasks. This method is especially useful in applications where high-dimensional datasets are common, such as environmental management, engineering design, and economic modelling. By focusing on the most critical aspects of the data, ORA-RPSS empowers decision-makers to make more informed and efficient analyses.

3.1.4. δ-MOSS Greedy

This algorithm was introduced by Brockhoff and Zitzler (2009) and is a method to solve the basic dimensionality reduction (DR) problem δ-MOSS described in the same paper. The δ-MOSS problem focuses on identifying a subset of indicators from a larger set within a user-defined tolerance. This algorithm provides a greedy alternative to exact solution methods, which tend to scale poorly as the number of indicators increases, making them impractical for high-dimensional datasets.

It is one of the first methods of DR introduced and commonly used as a reference for other algorithms. The basic concept of the algorithm is to pairwise compare solutions from a representative solution set obtained from a PF approximated by a Multi-Objective Evolutionary Algorithm (MOEA) and thereby create a subset $F' \subset F$, which is δ-nonconflicting with F. The user has to specify the parameter δ, which measures the change in the dominance structure based on ϵ-dominance, and represents the accepted error in the algorithm.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Its scalability and practical implementation have made it a commonly referenced method in both academic research and applied settings. By offering an effective balance between computational efficiency and the preservation of essential data characteristics, the algorithm is well-suited for tasks involving high-dimensional datasets, such as indicator aggregation in complex systems.

3.1.5. Numerical experiments

To illustrate the practical application of these dimensionality reduction (DR) methods, we employ the DTLZ5 problem from the well-known DTLZ test suite (Deb et al., 2005). The DTLZ test suite is a standard benchmarking framework extensively used in the evaluation of many-objective optimization (MaOO) algorithms due to its versatility and ability to represent various characteristics and challenges encountered in real-world optimization problems. The DTLZ5 problem, in particular, is well-suited for demonstrating DR methods, as it is designed to highlight scenarios involving both redundancy and conflict among objective functions. These characteristics make it an ideal surrogate for systems of indicators, where the challenge often lies in managing similar redundancies and trade-offs.

In the context of the DTLZ5 problem, the objective functions can be viewed as analogous to indicators in a system of indicators. This analogy stems from the fact that both objectives in MaOO problems and indicators in complex systems are often correlated, redundant, or conflicting. These parallels provide a meaningful framework for testing and analysing the performance of DR methods when applied to high-dimensional datasets.

By using the DTLZ5 problem as the basis for numerical experiments, this section aims to demonstrate how DR methods perform under controlled yet challenging conditions. The ability to manipulate the problem setup ensures a thorough and systematic evaluation of the methods, highlighting their strengths and limitations in various scenarios. This approach not only provides insights into the theoretical capabilities of the DR methods but also sheds light on their potential applicability to real-world systems of indicators.

3.1.5.1. Benchmarking using DTLZ5(I,M) Problem

The DTLZ problem suite from Deb, Thiele, Laumanns and Zitzler (Deb et al., 2005) is one of the most commonly used test suites for MOEAs and contains 8 different problem sets. In the DTLZ test suite, a set of problems is provided which allows the user to define the number of indicators and decision variables. For these problems, the POF is known. The test suite contains problems with redundant and non-redundant features. For the experiment, the redundant DTLZ5 problem is chosen.

The DTLZ5 problem is a redundant DTLZ problem set with a dimensionality of the POF smaller than the number of original indicators. The size of the set F_T (I) and the size of F_O (M) can be specified by the user, where F_T (I) denotes the set of I essential indicators and F_O (M) denotes the set of M original indicators. In the following the according problem set will be referred to as DTLZ5(I, M) or D(I, M) (e.g. D(3, 5) refers to a problem with five original indicators, with two redundant and three essential indicators). The objectives of the problem are built as follows:

$$Min.f_1(x) = (1 + g(x_M))cos(q_1\pi/2)...cos(q_{M-2}\pi/2)cos(q_{M-1}\pi/2),$$

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

$Min.f_2(x) = (1 + g(x_M))cos(q_1\pi/2)...cos(q_{M-2}\pi/2)sin(q_{M-1}\pi/2)$, $Min.f_3(x) = (1 + g(x_M))cos(q_1\pi/2)...sin(q_{M-2}\pi/2)$,

...

$Min.f_M(x) = (1 + g(x_M))sin(q_1\pi/2)$,

with $q_I = \pi/4/(1+g(x_M))1+2g(x_M)x_i)$, for i =2, 3, …, (M-1),

$g(x_M) = \sum_{x_i \in x_M}(x_i - 0.5)^2$, 0 £ x_i £ 1, for i = 1, 2, …, n.

All of the indicators have to be minimized. The POF is obtained with g = 0, which corresponds to x_i = 0.5 for all $x_i \epsilon x_M$ with $x_i \epsilon [0, 1]$. The number of decision variables n is set to $n = M + k - 1$ with k=10 according to the suggestion for optimal performance from (Deb et al., 2005). For every combination of I and M the indicators $f = \{1, ..., h\}$ with h = M-I+1 are perfectly correlated, while the rest of the indicators is conflicting with every other indicator (Saxena et al., 2013). Therefor the set of essential indicators consists of $F_T = \{f_h, f_{M-I+2}, ..., f_M\}$.

3.1.5.2. Experimental setup

In this experiment, four different DR algorithms are tested on four combinations of I and M for the DTLZ5 problems shown in Table 3.2 Two different setups are tested for either feature selection or extraction. All methods tested are feature selection algorithms or can be used as such, but the NPCA-based method can also be used as feature extraction algorithm. An initial set of solutions is required as input for the algorithm, therefore a set of 200 random solutions is generated by applying an NSGA-II algorithm for only one generation and using the solutions returned as the set of initial solutions. For each combination of I and M 30 independent sets of solutions are created and then all methods are applied to every set independently.

The DR methods δ-MOSS, ORA-RPSS, and PCSEA were run on 30 independent problem sets, evaluating the frequency of success in identifying the essential set F_T. The specific parameters chosen for this study for each technique are shown in Table 3.1. Two steps were used in the evaluation of the performance of NPCA. In the first step, the number of components to be extracted as principal components is chosen based on the explained variance using a threshold of EV > 0.8. The threshold was selected as the highest number of runs finding the correct number of essential indicators was obtained while testing the thresholds 0.7, 0.75, and 0.8 as suggested in Giuliani et al. (2014). In the second step indicators are selected by looking at the coefficients of the principal components and selecting the indicator with the highest contribution for each PC.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Project funded by Swiss Confederation

Table 6. Complexity of the algorithms as stated in the sources. Algorithm parameters (P) chosen for benchmarking tests.

Algorithm	Complexity	P	Setting
NPCA		EV	0.8
δ-MOSS Greedy	$O(\min n^2m^3, n^4m^2)$	δ	0
PCSEA	$O(Nm^2 \log N)$	C	0.8
ORA-RPSS (3, 5 | 3,10)	$O(N^2n^2)$	C	0.7
ORA-RPSS (5,10 | 5,20)		C	0.8

N = size of non-dominated set, m = number of original indicators, n = number of solutions

Table 7. Setup of the four DTLZ5(I, M) problems

I	M	Solutions	k
3	5	200	14
3	10	200	19
5	10	200	19
5	20	200	29

I = size of F_T , M = size of F_O, k = number of decision variables

3.1.5.3. Numerical results

The ORA-RPSS method shows the overall best performance as for every problem F_T is correctly identified in a minimum of 83 % of the runs (Figure 3.1). The NPCA method, shows a very good performance for D(3,5) and a medium performance for D(5,20). This indicates that this method performs better if the size of F_T is close to F_O or $F_O >> F_T$. The PCSEA method shows a low performance for the two sets with I = 3. This could be related to the choice of C as ORA-RPSS performs well with an adjusted value for C. In the analysis, it was also visible that the size of F returned by PCSEA is on average very close to the correct size of F_T . The reference values for δ-MOSS Greedy are taken from Yuan et al. (2018) and have been reproduced similarly in other papers. The algorithm identifies the correct set in only 17 % on the best-performing problem and in 0 % of the runs for the worst case.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Project funded by Swiss Confederation

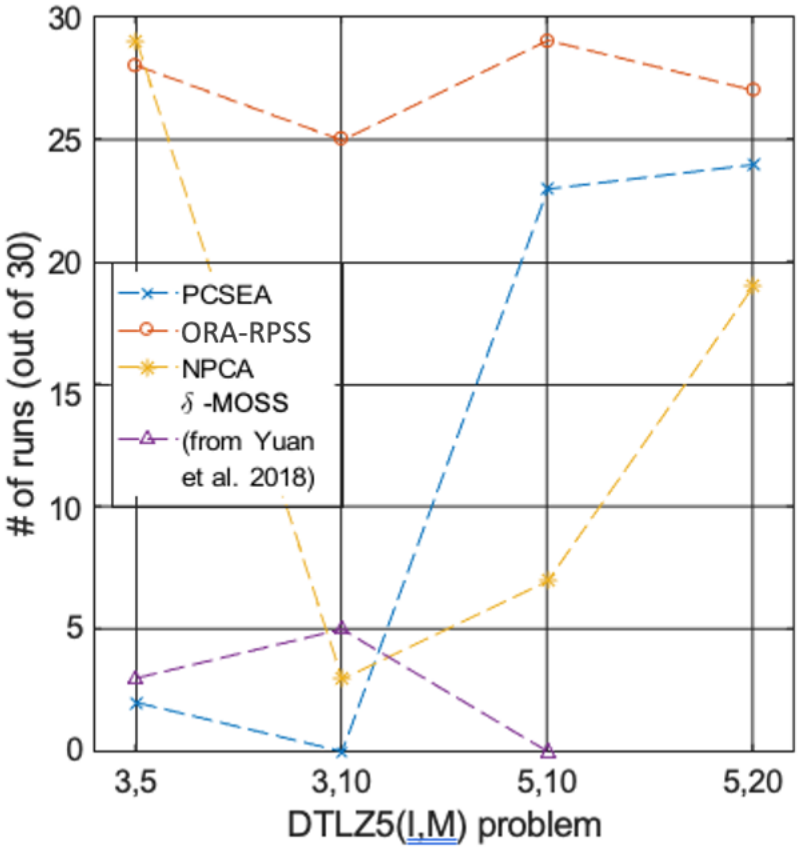

Figure 6. Performance on DTLZ5(I,M) problems. Number of runs with the essential set correctly identified. ORA-RPSS consistently returns the correct essential indicators in a high number of runs.

The results of the benchmarking tests showed strong deviations between the performance of the methods. ORA-RPSS showed a stable high performance over all problems, while the other algorithms showed low performance for at least one problem set. The approximations of the PF obtained from the NPCA method showed good solution diversity and convergence. A short runtime for the NPCA itself was observed, but the full runtime significantly increased adding to the runtime of the optimization with Borg MOEA. This increase could be handled with an improved coupling of the problem with the algorithm, parameter tuning, parallelization of the optimization or using a different MOEA which performs well on the type of problem. Two problem features were observed to have an effect on changes in performance: the increasing number of indicators and the fraction of redundant indicators in a set in relation to the number of original indicators. This aspect could be relevant for future analysis of real-world problems when searching for a fitting DR method.

The findings from the experiments conducted are subject to several important limitations that should be acknowledged. First, the performance of DR was tested within the context of a many-objective optimization (MaOO) problem, where the primary goal is to identify the optimal solutions—specifically, those on the Pareto front. This differs from the context of SOS-Water evaluation in the project, where the goal is not solely to pinpoint the best solutions based on all indicators but to analyse and assess the entire solution set comprehensively. In SOS evaluation, every solution (e.g., the SOS evaluation with a specific local management pathway under a specific scenario) holds value, not just those deemed

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by

optimal in terms of Pareto efficiency. Despite this contextual difference, the insights derived from MaOO problems offer meaningful guidance for real-world indicator systems, where redundancy and conflicts among indicators are common challenges.

Second, the analysis was limited to a single redundant benchmarking problem with a specific problem structure. While this allowed for demonstrating the method's potential, the results may be influenced by the unique characteristics of the selected problem. To fully understand the performance of DR in practical applications, further evaluation of real-world problems is necessary. This would provide a more comprehensive assessment of its effectiveness and applicability in diverse, real-world scenarios.

Finally, additional factors that could impact performance were not extensively explored. For instance, the tuning of parameters was limited, and the influence of different solution set sizes was not fully investigated. Similarly, the potential effects of varying levels of noise in the initial solution set were not accounted for, despite prior research (e.g., Saxena et al., 2013; Yuan et al., 2018) highlighting the significance of these factors. Future studies should address these aspects to provide a more comprehensive understanding of the method's performance across diverse scenarios and conditions.

3.2. Multi-Attribute Value Theory

Multi-Attribute Value Theory (MAVT), introduced by Keeney and Raiffa (1976), guides decision-makers in finding the ideal alternative when there are several options to evaluate. It serves as a type of multi-criteria decision analysis (MCDA) with the ability to aggregate and evaluate both quantitative and qualitative information or attributes, which is essential in the realm of indicators determining the multi-objective nature of the SOS water framework. MAVT can be used to handle challenges that involve a finite and discrete set of choices or alternatives that must be judged and prioritized (e.g. by assigning weights according to their performance and relevance). The method revolves around defining a value function V(·), which takes a vector of indicators that measure the impacts and benefits of a given alternative and produces a single value that reflects the satisfaction level of the decision-maker regarding that alternative. When the value function is defined, it allows for the systematic prioritization of various alternatives, considering that an increase in V leads to enhanced satisfaction levels.

3.2.1. Evaluating the alternatives

MAVT relies on a limited number of robust principles of rationality when addressing alternatives, primarily focusing on completeness and transitivity (Eisenführ et al. 2010). Completeness means that a decision-maker is able to express her likes or neutrality regarding any two possible alternatives, while transitivity suggests that when a decision-maker chooses alternative A over alternative B and prefers B over C, she ought to prefer A over C as well.

Alternatives are feasible solutions to address the decision-related challenge, indicating the multiple choices or options that are reviewed considering various standards or attributes (e.g., indicators) (Aubert et al. 2018). The assessment procedure entails analysing different options by evaluating their effectiveness in relation to multiple characteristics, listing them according to their diminishing satisfaction, ultimately opting for the one that ranks highest. Therefore, selecting involves establishing a hierarchy among priorities. For this purpose, it is necessary to establish scales, which are important for gauging the effectiveness of alternative solutions based on diverse criteria. Scales can fall into two

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by Schweizerische Eidgenossenschaft Confédération suisse Confederazione Svizzera Confederaziun svizra Swiss Confederation Federal Department of Economic Affairs, Education and Research EAER State Secretariat for Education, Research and Innovation SERI

categories—quantitative or qualitative—based on the specific criteria under review. Ensuring uniform scaling for every criterion is essential to uphold the validity of the decision-making process. This process ensures that the total scores faithfully capture the decision-maker's priorities, avoiding the risk of any single factor skewing the overall judgment due to differences in how measurements are expressed or the ranges used.

Ordinal scales serve to arrange options or levels of attributes according to a hierarchy of preferences, without measuring the extent of differences among them (Soncini-Sessa et al. 2007). For example, given four alternatives (A1, A2, A3, and A4), an ordinal ranking can be expressed in the following form:

$$\{A2 > A1 \sim A3 > A4\}$$

In which A2 would be preferred over A1 (that is, A2 > A1), A1 is equivalent (∼) to A3, which is preferred over A4. Consequently, we can conclude that the chosen alternative will be A2.

This implies that although you can identify which choices are preferred compared to others, the measurement does not reveal the extent to which one option is favoured over another. In other words, ordinal scales establish an order (such as first, second, third) but fail to indicate the degree of preference variations between the ranks, which restricts the types of mathematical analyses that can be performed (only non-parametric techniques are appropriate for examining ordinal data). This means that when it comes to selecting just the best option, using ordinal scales for overall preference values is enough (Choo et al. 1999). However, it is important that an ordering principle is not just explicitly stated; it should also reflect democratic values (specifically when there are three or more options to arrange), as this will ensure that stakeholders and decision-makers embrace it. According to Arrow's (1951) mathematical findings, democracy will be ensured if five reasonable fairness principles (Table 3.3) are accomplished when merging personal preferences into a group ranking.

Table 8. Principles to be ensured when applying an ordering scale.

Principle	Description
Universality	Adjusts to the personal preference hierarchies of all users.
Pareto efficiency (unanimity)	Values the consensus of preferences among everyone involved.
Independence of irrelevant alternatives	The hierarchy of two alternatives is influenced exclusively by the preferences held for each of those alternatives.
Non-dictatorship	No individual voter holds the authority to influence the group's hierarchy.
Transitivity (implied consistency)	The hierarchy within the group must be consistent in its logic.

However, the Arrow's paradox highlights that no method of aggregating individual preferences into a collective decision can satisfy all reasonable fairness principles simultaneously when there are at least three alternatives and two or more decision-makers. For example, consider three decision-makers and three options A, B, and C:

Decision-maker 1: A > B > C

Decision-maker 2: B > C > A

Decision-maker 3: C > A > B

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

If we try to aggregate these preferences into a single group ranking, at least one of Arrow's criteria will be violated:

A > B, B > C, and C > A = leads to a cyclical ranking (violating the 'transitivity' principle).

Another option is to apply absolute scales, in which the values given to characteristics can be directly understood and interpreted without requiring any modifications or adjustments. On an absolute scale, cardinal rankings are used to express the numerical relevance of each alternative (with a concrete zero point, making interpretation straightforward). Following the previous example for ordinal scales, we can express a cardinal ranking in this way:

{16, 24, 16, 13}

where the first number is the value attributed to alternative A1, the second to A2, and so on. A key point when applying absolute scales is the need to determining both the measurement unit for the values but also the zero point, which is interpreted as the absence of value, and this is a complex endeavour due to the inherently subjective nature of the attributes. Therefore, it could be a better option to use interval scales: in this case, preferences are rated to incorporate all possible value interpretations, but a true zero point does not exist. An interval scale allows for an endless variety of equivalent representations of a value, all interconnected through a linear transformation that utilizes positive coefficients. For example, in an interval scale, the ranking could be expressed in this way:

{1650, 2450, 1650, 1350}

which is obtained by applying the transformation to the previous formula:

$$i' = 100i + 50$$

but it is also equivalent to the following ranking:

{10, 14, 10, 8.5}

which is obtained by applying another transformation pattern:

$$i' = 0.5i + 2$$

Therefore, the cumulative values of the choices will take the form of intervals to analyse the trade-offs among various attributes (Mustajoki et al. 2006). Take, for example, a situation where boosting one variable by a certain extent compensates for a diminishment in another; the interval scale will confirm that such trade-offs are proportional and stable. The process to establish an interval scale is simple:

- Establish the spectrum: Determine the optimal and least favourable points for the alternative (for instance, the lowest and highest values).
- Set the scale limits: Assign numerical values to the scale's endpoints (generally using 0 to represent the least favourable outcome and 1 or 100 for the most favourable).
- Gather preferences: Employ techniques such as direct scoring, comparative pairs, or swing weighting to identify intermediate values.
- Maintain uniformity: Be sure that the intervals indicate consistent differences in preference.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Project funded by

3.2.2. Giving weights to alternatives

The weights indicate how stakeholders prioritize trade-offs among different goals, representing their assessment of the significance of each objective (Morton and Fasolo, 2009). They quantify what matters in the decision to prioritize among different alternatives. Requesting the weights to decision-makers through the direct-ratio method is the easiest way to go about it (Lin, 2013); however, main disadvantages of this method include subjectivity (relies on the ability of the decision-maker to provide accurate and consistent ratios), cognitive load (it becomes challenging to apply with many attributes), and range is not considered (it can distort the relative importance), making other options more advisable, such as the *swing* and *trade-off* methods (Riabacke et al. 2012).

- The Swing weighting method is a participatory strategy that emphasizes the effect of adjusting an attribute from its lowest point to its highest point, in comparison to other attributes that are defining a specific alternative. Once the worst and best levels for each attribute are determined (e.g., worst = 0, best = 1), it can be fixed a hypothetical baseline scenario by assuming all attributes are at their worst levels, to then evaluate the impact of improving each attribute (e.g. ask the decision-maker to imagine swinging one attribute from its worst level to its best level while keeping the rest at their worst levels). After replicating the process for all attributes, a rank can be established in terms of how much the swing from worst to best improves the overall outcome by assigning weights to the attributes proportionally (the most important swing is expected to be given to the highest weight). The sum of the weights should be normalized to 1 for consistency. By using this method, it is important to consider that the process can be cognitively demanding for decision-makers, particularly if many attributes are considered. Likewise, it is essential to provide clear definitions of the worst and best levels meaning for each attribute to avoid misunderstandings of the attributes' impacts, and consequently, of the alternatives.
- The trade-off weighting method requires that the decision-maker evaluates different attributes by imagining situations in which compromises need to be fixed between them. The process starts by defining two hypothetical scenarios where an improvement in one attribute is balanced by a decrease in another attribute. At this point, we should ask the decision-maker which scenario is the preferred one (e.g. *Would you rather see a major enhancement in attribute A paired with a minor reduction in attribute B, or vice versa?*). Trade-offs can be adjusted until the decision-maker is indifferent between both scenarios, and relative weights can be calculated according to the relative importance of the attributes. As in the previous method, weights must be normalized to ensure they sum 1. By applying this method, it is important to carefully frame scenarios to avoid bias in their interpretation, while being aware of its time-consuming nature when considering multiple and complex trade-offs.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Project funded by Swiss Confederation

Table 9. Methods for eliciting weights to objectives

Aspect	Swing weighting	Trade-off weighting
Focus	Impact of improving attributes from worst to best.	Direct comparisons between attributes.
Ease of use	Intuitive but can be demanding for many attributes.	More challenging, especially with many trade-offs.
Attribute – Ranges	Explicitly considers attribute ranges.	Implicitly considers ranges through trade-offs.
Cognitive load	Moderate.	High, especially with complex scenarios.
Preferred context	When clear ranges and priorities are known.	When direct comparisons between attributes are important.

3.2.3. Global and partial value functions

The goal of the MAVT is to systematize the decision-maker's preferences, meaning it aims to collect the relevant information required to rank the various alternatives based on their level of satisfaction. The top option in the list represents the most suitable compromise, considering the various assessment criteria. As anticipated, the decision-maker's satisfaction is based on a preference structure that can be characterized by a function V (·), known as the *global value function*, which assigns a value to each option A, reflecting the decision-maker's level of satisfaction with that choice. The level of satisfaction is influenced by the outcomes generated by the alternative, meaning that this function essentially reflects the value represented by the vector *i*(A) of indicators associated with alternative A:

$$V\,(i(A)) = V\,(i_1(A), \ldots, i_m(A))$$

The value *V (i(A))* is also called the *Project Index* and enjoys the property that

$$V\,(i(A1)) > V\,(i(A2)) \text{ à } i(A1) > i(A2)$$

$$V\,(i(A1)) \geq V\,(i(A2)) \text{ à } i(A1) \geq i(A2)$$

Where *i*(A1) > *i*(A2) signifies that the decision-maker favours the outcomes of option A1 over those of option A2, indicating a preference for the first alternative; while *i*(A1) ≥ *i*(A2) indicates that the decision-maker has a mild inclination towards the first over the second or perceives them as being on par. Therefore, the most effective alternative A_c is the one that increases the value of V (*i*(A)) as much as possible.

However, the overall objective's value function is usually established by deriving separate (or partial) value functions for each sub-objective at the lowest level, which depend on a limited number of attributes, and then integrating these values at more advanced levels. This necessitates outlining value functions for all fundamental goals and creating aggregation guidelines for higher-level objectives. Different procedures exist, two of the most illustrative procedures being the *midvalue splitting* and the *significant points* methods.

The *midvalue splitting* method is a systematic approach that uses the Matrix of the Effects. Its configuration is based on the estimation of the values i_j *(A)* of the *m* evaluation indicators (*j* = 1, … , *m*) for a given alternative (*A*), duly organized in a vector *i*(A). The set of the vectors *i*(A), which were

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

calculated for each of the alternatives, forms a matrix in which the rows correspond to the indicators and the columns to the alternatives. For a given indicator, we need to define the domain [a, b] of the function v_k (·) based on the values in the matrix and the decision maker's (or the expert's) experience. It is essential for the domain to be extensive enough to include every possible situation that could arise, not limited to what is shown in the matrix, since there may be a need to create and assess additional alternatives later; concurrently, it should not be so expansive that it leads to confusion for the decision-maker. Then, we should ask the decision-maker if the indicator needs to be maximized, obtaining three possible responses:

(a) the decision-maker wants to maximize the indicator: the partial value function is monotonically increasing and the values of v_k (·) at the two extremes of the domain [a, b] are set as follows:

$$v_k(a) = 0 \quad \& \quad v_k(b) = 1$$

(b) the decision-maker wants to minimize the indicator: the partial value function is monotonically decreasing and the values of v_k (·) at the two extremes of the domain [a, b] are set as follows:

$$v_k(a) = 1 \quad \& \quad v_k(b) = 0$$

(c) the decision-maker neither wants to maximize nor to minimize the indicator: in this case, the function is not monotonic, and it is not possible to associate a value to the points a and b as in the previous cases.

At this point, we could ascertain that the function is increasing (a) or nonmonotonic (c). In both cases, we should ask the decision-maker to indicate the greatest value ∈ [a, b] for which the value of the function is zero and the smallest value ∈ [a, b] for which its value is equal to one. Then consider the interval [,] and ask the decision-maker to specify the value such a way that:

$$V_{k\,i}\,\frac{2}{k} = 0.50$$

and repeat the process asking for i_k^3 and i_k^4 such that:

$$v_k\, i_k^3 = 0.25 \quad \& \quad v_k\, i_k^4 = 0.75$$

and so forth, consistently splitting the intervals of the value function into smaller halves.

It would be a matter of asking "What fraction of the total value should be allocated to attribute 'A' versus all other attributes?"? After identifying a sufficient set of points, the v_k (·) function can be formulated using interpolation techniques. To accomplish this task, the decision-maker is required to connect the indicator's values with a range of function values that were not utilized in the identification process; her responses are then assessed in comparison to the values produced by the estimated function. The function ought to be improved as needed, until the decision-maker feels confident that it reflects her choices accurately.

The advantages of this method are diverse: it is a structured process, which means it provides a step-by-step framework that simplifies weight elicitation; it is consistent, that is, it ensures logical and proportional weight distribution; and encourages active participation and reflection on preferences by the decision-maker. However, it faces cognitive demands (it requires the decision-maker to make multiple proportional judgments, which can be challenging for complex problems), is subjective when

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by Swiss Confederation

processing with splitting, which may introduce bias; and is time-consuming, that is, it can be lengthy if there are many attributes to evaluate.

The *significant points* method is a structured approach mostly appropriate when collaborating with an expert from whom to inquire about the key values of the indicator along with the corresponding function values. This means that rather than beginning with the $v_k(\cdot)$ values and progressing to the i_k values, one should initiate the process from i_k and link it to $v_k(i_k)$. This method focuses on assigning relevant points to the attributes, reflecting their contribution to the overall decision outcome. We should ask the decision-maker or expert to identify a subset of minimum and maximum values of i_k for each attribute, which provides a range of possible values for each attribute (e.g., worst to best levels). These values ('significant points') can range between 0 and 1 or among 0 and 100 to facilitate the assignment of points among attributes (at the end we can convert weights to sum 1). The decision-maker will allocate points to each attribute based on its perceived importance. Attributes that are more critical to the decision receive more points, reflecting their relative contribution to achieving the desired outcome.

The main advantages of this method are its simplicity (easy to understand and apply, especially for decision-makers unfamiliar with complex elicitation techniques), flexibility (it works well for a wide range of decision problems) and direct elicitation (allows decision-makers to intuitively express their preferences and reshape them after being tested). However, some limitations are noted: subjectivity (it relies on the decision-maker's subjective judgment, which may introduce bias or inconsistency), no explicit range consideration (it does not inherently account for the magnitude of change within each attribute's range, which is arbitrary), and limited precision (it may oversimplify complex decision-making problems with interdependent attributes).

Once all the partial value functions $v_k(\cdot)$ have been identified, the global value function $V(i_1, \ldots, i_m)$ is defined by their sum. The global value function can be two-dimensional when it is the composition of only *two* partial value functions $v_1(\cdot)$ and $v_2(\cdot)$, whose arguments are the indicators i_1 and i_2 respectively. In this case, for estimating these values, it would be sufficient to identify two pairs (i_1', i_2') and (i_1'', i_2'') of indicator values that the decision-maker considers to be equivalent. In the case of multi-dimensional value functions, we can repeat the reasoning used in the two–dimensional case, but in this case the decision-maker should identify $m - 1$ pairs of equivalent points in the space of the values. As can be expected, the identification of these pairs becomes more difficult as the number m of indicators increases. To simplify this process, we can ask the decision-maker to specify the relative importance between pairs, that is, carry out a pairwise comparison between the criteria associated to each pair of indicator values. To do that, we can establish a hierarchy by applying the Analytic Hierarchy Process through which the decision-maker can compare attributes or alternatives two at a time to determine their relative importance or preference.

3.2.4. Evaluation

Once the global value function is determined and the hierarchy on attributes or alternatives is fixed, we can simplify the decision-making process and transform the multi-objective challenge into an evaluation of the indicator values based on the decision-maker's preference and satisfaction. Therefore, the alternatives are prioritized by sorting them according to the lower values of the function. However, the method used to derive the rankings is inherently marked by a degree of randomness and ambiguity,

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by

evident in the assessment of impacts, the determination of individual value functions, and the assignment of weights. Intrinsic uncertainty exists since the decision-making estimates the effects of something that does not yet exist, indicators used to characterize the partial value functions can be more or less representative and accurate, and the weights attribution are challenging when translating qualitative comparisons into quantitative values. Consequently, the decision-maker can perceive that the ranking is too fragile to be robust, as changing their answers a little bit could provide another type of ranking.

A sensitivity analysis can help to ascertain whether the decision-maker is confident with the results by evaluating how changes in the inputs (attribute weights, value functions, ranges) affect the overall decision outcome (Wieckowski et al. 2024). For example, the decision-maker can gradually increase or decrease the weight of a specific attribute and observe its impact on the overall scores; modify the risk aversion level or linearity of a partial value function to reflect different preferences; expand or contract the range of an attribute to assess how it affects the calculated weights and rankings; identify the weight at which one alternative overtakes another in ranking; or even create a best-case, worst-case, and most-likely scenarios to test the robustness of the decision.

3.2.5. Numerical experiments

To illustrate the practical application of some of the methods considered in the MAVT, we conducted a participatory process focused on establishing a co-developed and comprehensive objectives hierarchy (i.e. *Project Hierarchy*) to organize the multi-attribute value functions associated to the Mekong Delta case study. An objectives hierarchy provide a structured framework that organizes objectives into a logical sequence, starting from broad, overarching goals, and refining them into specific actionable targets and attributes whose degree of success is assessed using measurable indicators (Langhans and Lienert 2016; Zheng and Lienert 2018; Marttunen et al. 2019). To this regard, an in-person workshop was conducted in October 2023 to enhancing stakeholder's engagement and capturing decision-maker's viewpoints and preferences, while consolidating shared understanding of the main objective of co-developing a water management plan for the whole basin context, which ensures to stay within the Safe Operating Space (SOS) now and in the future. The goal was threefold: 1) Identify main values and translate them into objectives, 2) Define an objectives hierarchy, and 3) Agree on indicators to monitor each objective.

Following an iterative approach, decision-makers (i.e. stakeholders representing public administration, industry, academia, and NGOs) were able to identify main water management challenges and share their own visions for the future of the Delta and related water values. Consequently, the main objective aligned with the SOS water framework was divided in two sub-objectives or dimensions: 1) Ensuring a healthy delta (e.g. maintaining and/or restoring good water quality, sufficient water quantity, and ecosystem services) and 2) promoting a liveable delta (e.g. promoting sustainable industries and livelihoods, enhance resilience to extreme events and maintaining local culture). For each dimension, water values were identified, and decision-makers collaboratively draft an objectives hierarchy, arranging the objectives at different levels, from general up to four levels, becoming more specific at each level. Tentative measurable attributes (i.e. indicators) are assigned to each of the lowest-level sub-objectives to monitoring the objective's achievement (see Figure 3.2).

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Project funded by

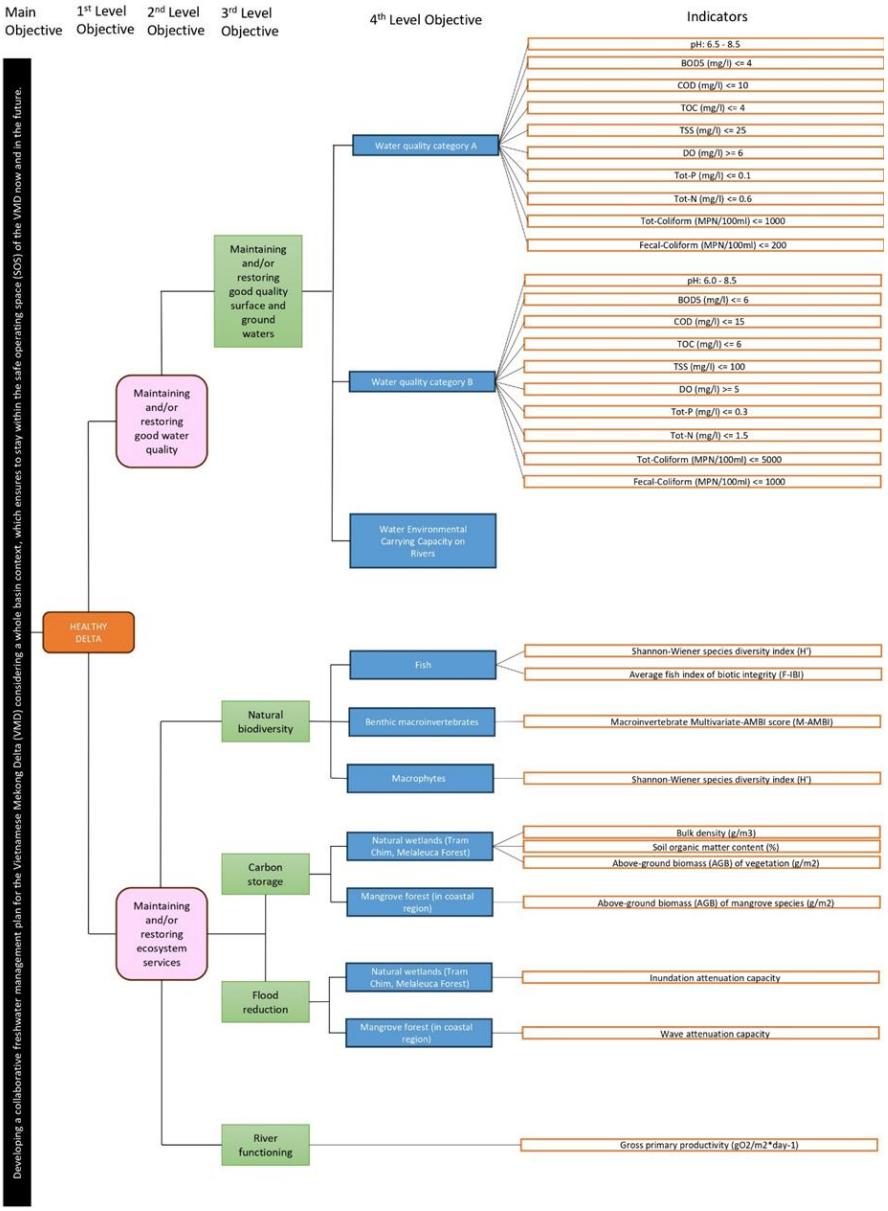

Figure 7. Objectives hierarchy for the main objective of co-developing a water management plan for the Vietnamese Mekong Delta. Details on main and secondary objectives and indicators. Further details in Deliverable 1.1. Case study-specific stakeholder engagement roadmaps.

These indicators were used to quantify the relationship between objectives and attributes through the partial value functions. By assuming an equal relevance for both dimensions (i.e. healthy and liveable delta), stakeholders firstly assigned weights to the objectives at the higher levels (e.g., water quality, ecosystem services, extreme events, and cultural values) through complementary one-on-one interviews (Figure 3.3). The aim was to discuss and validate the initial importance (e.g. values) assigned to each objective to then proceed with the internal distribution of this relevance moving down with the lower-level objectives.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by

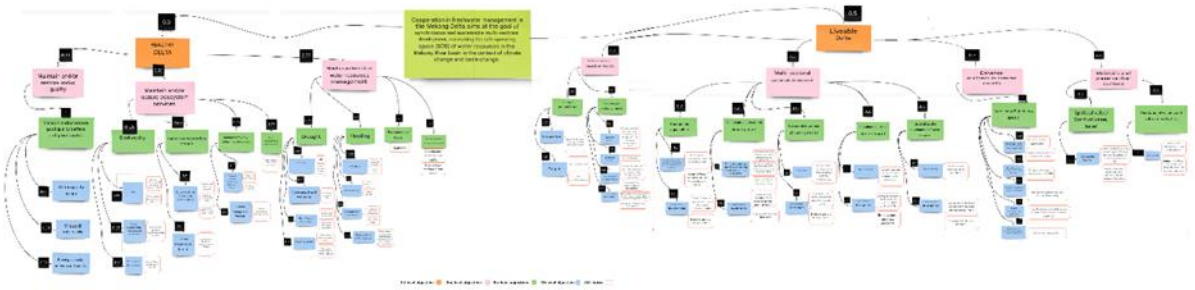

Figure 8. Objectives Hierarchy for the Mekong Delta case study

It is interesting to note that weights are equally distributed across the attributes/indicators representing the Healthy Delta dimension, that is, each of the three sub-objectives (water quality, ecosystem services, and water resources management) received the same relevance, as well as the indicators used for their accomplishment (Figure 3.4).

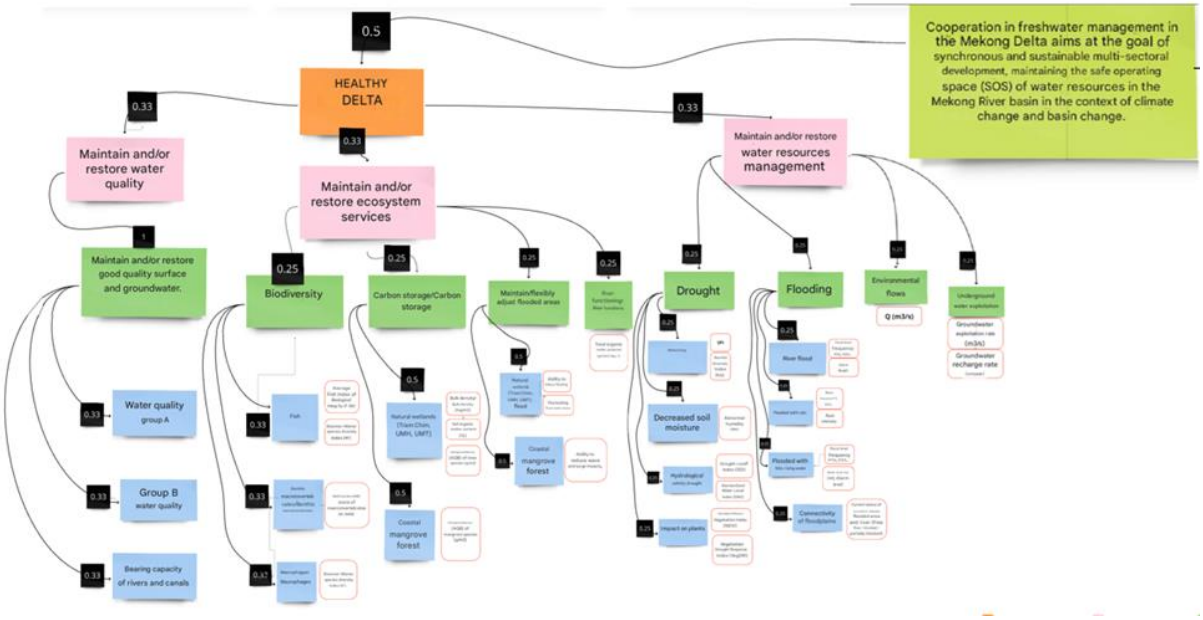

Figure 9. Objectives Hierarchy for the Mekong Delta case study – Focus on the Healthy Delta dimension

Interestingly, the weights assigned to the attributes defining the Liveable Delta dimension are not balanced: two sub-objectives –promoting multi-sectoral sustainable development and enhancing resilience to extreme events– are considered more relevant than safeguarding sustainable livelihoods and preserving river culture (Figure 3.5). This imbalance also occurs in the assessment of some lower-level objectives: to achieve multi-sectoral sustainable development decision-makers prioritize flood-based economic production and the agricultural and urban sustainable development. These differences in attribute relevance will determine the meaning of the value functions.

It is noteworthy that the importance given to the factors characterizing the Liveable Delta dimension is uneven: two specific goals—fostering sustainable development across various sectors and improving resilience against extreme events—are deemed more significant than protecting sustainable livelihoods and maintaining river culture (Figure 3.5). This imbalance also occurs in the assessment of some lower-

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

level objectives: decision-makers prioritize flood-based economic production and agricultural and urban sustainable development to achieve multi-sectoral sustainable development instead of focusing on tourism development.

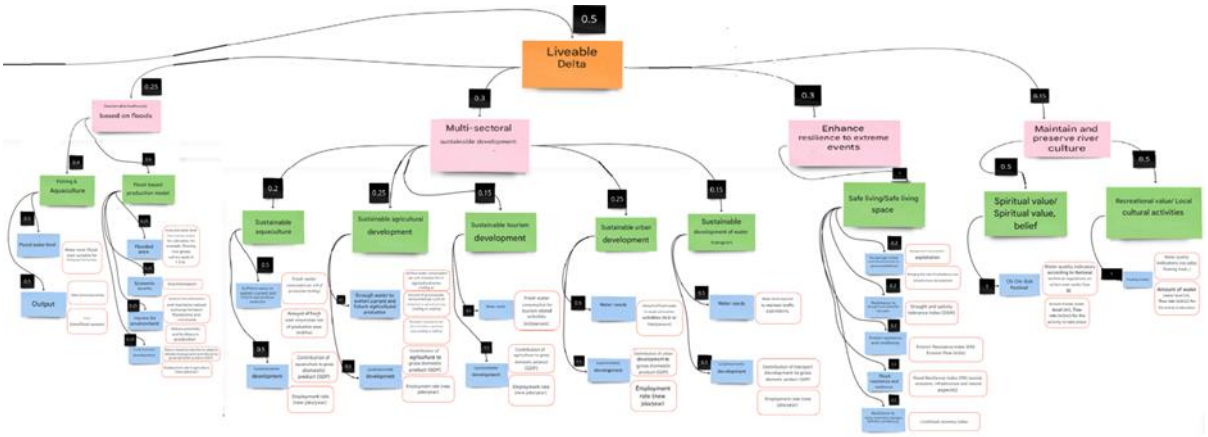

Figure 10. Objectives Hierarchy for the Mekong Delta case study – Focus on the Liveable Delta dimension

After establishing the objectives' hierarchy, the next phase will focus on utilizing a swing weighting approach to adjust each attribute in relation to a hypothetical baseline scenario. This provides the decision-maker with the opportunity to manipulate one feature from its least to its most extreme value and to organize the features based on this analysis.

4. Discussion and Next Steps

The work presented in this deliverable represents a significant advancement toward developing a comprehensive system of water indicators for the SOS-WATER project. In our approach, we have first established a thorough assessment of the current indicator landscape, identifying key gaps in traditional approaches, and then point out directions to design innovative indicators that can capture the spatial, temporal, and probabilistic nuances inherent in water resource management. Complementing this, we have explored advanced aggregation mechanisms, ranging from data-driven reduction techniques to frameworks based on Multi-Attribute Value Theory (MAVT), to synthesize the diverse set of indicators into a coherent meta-indicator system. While these efforts mark important progress, the discussion of our findings and the identification of future work reveal both strengths and areas requiring further development.

Our analysis in Section 2 revealed that existing water indicators, although well-established in many contexts, often fall short in addressing the multi-dimensional and dynamic nature of water systems. Traditional indicators tend to focus on individual aspects such as hydrological variability or meteorological extremes, without adequately integrating spatial patterns and temporal fluctuations of water variables. By conducting an extensive literature review and engaging project partners through interactive workshops, we have been able to pinpoint these critical gaps. As a result, we have outlined strategic directions for the development of innovative indicators that aim to overcome these limitations. These directions underscore the necessity to develop innovative, adaptive indicators that prioritize multi-dimensional integration and enhanced data resolution. For example, new meteorological indicators should merge measurements of soil moisture, groundwater levels, and surface water flows

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).
Project funded by

with traditional metrics like precipitation and evapotranspiration, thereby enabling a holistic assessment of water availability. Similarly, hydrological, groundwater, and surface water indicators must incorporate factors such as regional calibration, human interventions, and climate projections to capture spatial variability and dynamic system changes effectively.

Section 3 builds on these insights by addressing the equally challenging task of indicator aggregation. Given the high dimensionality and potential redundancy among individual indicators, our data-driven methods (e.g., NPCA, PCSEA, ORA-RPSS, and δ-MOSS Greedy) have proven effective in reducing complexity while preserving critical information based on a benchmarking numerical example. These methods show potential to facilitate the transformation of a large set of indicators into a more manageable collection of essential indicators that still encapsulate the essential characteristics of the water system. Meanwhile, the application of MAVT in Section 3.2 introduces a decision-making framework that integrates both quantitative and qualitative aspects by assigning weights to individual indicators based on stakeholder preferences. This dual approach not only provides flexibility in addressing the inherent trade-offs among different water management objectives but also enhances the transparency of the overall evaluation process.

Despite these promising developments, several challenges remain. First, the data-driven aggregation methods are sensitive to parameter choices and require careful calibration. The performance of techniques such as NPCA and ORA-RPSS can vary depending on the underlying data structure and the selected thresholds for explained variance or redundancy metrics. Similarly, while the MAVT approach is conceptually robust, it relies heavily on subjective weighting and the accurate elicitation of stakeholder preferences (factors that may introduce variability in the final aggregated outcomes). These challenges underscore the importance of further validation and iterative refinement, particularly as we move from controlled numerical experiments to real-world case studies.

In terms of the next steps, our immediate focus will be on the implementation of these methodologies across all the case studies in SOS-WATER project. So far, preliminary thresholds have been established for certain indicators based on extensive literature review, setting precise threshold values remains challenging with the current level of information. These thresholds are inherently dependent on specific system characteristics and performance metrics, a topic that will be further explored in Task 4.4. For the indicator aggregation methods presented in Section 3, the next phase will focus on applying these approaches using real-world data from each case study. This process may reveal that different case studies benefit from different aggregation strategies. Some may adopt data-driven techniques, while others might prefer the Multi-Attribute Value Theory (MAVT) approach, particularly if there is strong stakeholder engagement to guide the weighting process. Overall, these steps are crucial for refining the indicator framework and ensuring that both the innovative indicators and their aggregation methods are robust, context-sensitive, and fully integrated into the SOS-WATER project's broader evaluation framework.

Another critical next step is the integration of our work with the broader SOS-WATER framework. The indicator system developed in this deliverable is intended to feed into subsequent stages of the project, such as the comprehensive SOS evaluation in WP5. Ensuring seamless compatibility with other work packages will be essential for establishing a unified operational framework.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Project funded by Swiss Confederation

In summary, while the deliverable has successfully laid the theoretical and methodological groundwork for a new system of water indicators, significant work remains to fully operationalize this framework. Continued validation, methodological refinement, enhanced stakeholder engagement, and integration with the overall project architecture are essential to ensure that the final indicator system is both comprehensive and practical. These next steps will drive the iterative improvement of the framework, ultimately supporting more resilient and sustainable water resource management in the face of growing environmental challenges.

References

Alexander, L. V, Arblaster, J.M., 2009. Assessing trends in observed and modelled climate extremes over Australia in relation to future projections. Int J Climatol 29, 417–435.

Arrow, K. (1951). *Social Choice and Individual Values*. John Wiley & Sons: New York.

Aubert, A.H., Bauer, R., Lienert, J. (2018). A review of water-related serious games to specify use in environmental Multi-Criteria Decision Analysis. *Environmental Modelling & Software* 105: 64-78. https://doi.org/10.1016/j.envsoft.2018.03.023

Brockhoff, D., & Zitzler, E. (2009). Objective reduction in evolutionary multiobjective optimization: Theory and applications. *Evolutionary computation*, *17*(2), 135-166.

Choo, E.U., Schoner, B., Wedley, W.C. (1999). Interpretation of criteria weights in multicriteria decision making. *Computers & Industrial Engineering* 37(3): 527-541. https://doi.org/10.1016/S0360-8352(00)00019-X

Deb, K., Thiele, L., Laumanns, M., & Zitzler, E. (2005). Scalable test problems for evolutionary multiobjective optimization. In *Evolutionary multiobjective optimization: theoretical advances and applications* (pp. 105-145). London: Springer London.

Eisenfuhr, F., Weber, M., Langer, T. (2010). *Rational Decision Making*. 1st edition. Springer-Verlag, Berlin.

Famiglietti, J.S., 2014. The global groundwater crisis. Nat Clim Chang 4, 945–948.

Giuliani, M., Galelli, S., & Soncini-Sessa, R. (2014). A dimensionality reduction approach for many-objective Markov Decision Processes: Application to a water reservoir operation problem. *Environmental Modelling & Software*, *57*, 101-114.

Guo, X., Wang, Y., & Wang, X. (2016). An objective reduction algorithm using representative Pareto solution search for many-objective optimization problems. *Soft Computing*, *20*, 4881-4895.

Guttman, N.B., 1998. Comparing the palmer drought index and the standardized precipitation index 1. JAWRA Journal of the American Water Resources Association 34, 113–121.

Jackson, R.D., Idso, S.B., Reginato, R.J., Pinter Jr, P.J., 1981. Canopy temperature as a crop water stress indicator. Water Resour Res 17, 1133–1138.

Keeney, R.L., Raiffa, H. (1976). *Decision with Multiple Objectives: Preferences and Value Trade-offs*. John Wiley & Sons: New York.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Kim, Y.-H., Min, S.-K., Zhang, X., Sillmann, J., Sandstad, M., 2020. Evaluation of the CMIP6 multi-model ensemble for climate extreme indices. Weather Clim Extrem 29, 100269.

Konar, M., Evans, T.P., Levy, M., Scott, C.A., Troy, T.J., Vörösmarty, C.J., Sivapalan, M., 2016. Water resources sustainability in a globalizing world: who uses the water? Hydrol Process 30.

Langhans, S.D., Lienert, J. (2016). Four common simplifications of Multi-Criteria Decision Analysis do not hold for river rehabilitation. *PLoS ONE* 11(3): e0150695. https://doi.org/10.1371/journal.pone.0150695

Lin, S-W. (2013). An investigation of the range sensitivity of attribute weight in the analytic hierarchy process. *Journal of Modelling in Management* 8(1): 65-80. https://doi.org/10.1108/17465661311311987

Marttunen, M., Haag, F., Belton, V., Mustajoki, J., Lienert, J. (2019). Methods to inform the development of concise objectives hierarchies in multi-criteria decision analysis. *European Journal Operational Research* 277: 604-620. https://doi.org/10.1016/j.ejor.2019.02.039.

McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales, in: Proceedings of the 8th Conference on Applied Climatology. Boston, pp. 179–183.

Mekonnen, M.M., Hoekstra, A.Y., 2015. Global gray water footprint and water pollution levels related to anthropogenic nitrogen loads to fresh water. Environ Sci Technol 49, 12860–12868.

Morton, A., Fasolo, B. (2009). Behavioural decision theory for multi-criteria decision analysis: a guided tour. *Journal of the Operational Research Society* 60(2): 268-275. https://doi.org/10.1057/palgrave.jors.2602550

Mustajoki, J., Hamalainen, R.P., Lindstedt, M.R.K. (2006). Using intervals for global sensitivity and worst-case analyses in multiattribute value trees. *European Journal of Operational Research* 174: 278-292. https://doi.org/10.1016/j.ejor.2005.02.070

Riabacke, M., Danielson, M., Ekenberg, L. (2012). State-of-the-art prescriptive criteria weight elicitation. Advances in Decision Sciences, 2012: 276584. https://doi.org/10.1155/2012/276584

Richter, B.D., Davis, M.M., Apse, C., Konrad, C., 2012. A presumptive standard for environmental flow protection. River Res Appl 28, 1312–1321.

Saxena, D. K., Duro, J. A., Tiwari, A., Deb, K., & Zhang, Q. (2012). Objective reduction in many-objective optimization: Linear and nonlinear algorithms. *IEEE Transactions on Evolutionary Computation*, *17*(1), 77-99.

Shafer, B.A., Dezman, L.E., 1982. Development of a surface water supply index (SWSI) to assess the severity of drought conditions in snowpack runoff areas.

Shukla, S., Wood, A.W., 2008. Use of a standardized runoff index for characterizing hydrologic drought. Geophys Res Lett 35.

Singh, H. K., Isaacs, A., & Ray, T. (2011). A Pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems. *IEEE Transactions on Evolutionary Computation*, *15*(4), 539-556.

Soncini-Sessa, R., Weber, E., Castelletti, A. (2007). *Integrated and Participatory Water Resources Management – Theory*. Volume 1A. 1st edition. Elsevier Science: Amsterdam.

Sullivan, C., 2002. Calculating a water poverty index. World Dev 30, 1195–1210.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).

Project funded by Schweizerische Eidgenossenschaft Confédération suisse Confederazione Svizzera Confederaziun svizra Swiss Confederation Federal Department of Economic Affairs, Education and Research EAER State Secretariat for Education, Research and Innovation SERI

Tramblay, Y., Koutroulis, A., Samaniego, L., Vicente-Serrano, S.M., Volaire, F., Boone, A., Le Page, M., Llasat, M.C., Albergel, C., Burak, S., 2020. Challenges for drought assessment in the Mediterranean region under future climate scenarios. Earth Sci Rev 210, 103348.

Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., 2010. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. J Clim 23, 1696–1718.

Vicente-Serrano, S.M., Domínguez-Castro, F., Reig, F., Beguería, S., Tomas-Burguera, M., Latorre, B., Peña-Angulo, D., Noguera, I., Rabanaque, I., Luna, Y., 2022a. A near real-time drought monitoring system for Spain using automatic weather station network. Atmos Res 271, 106095.

Vicente-Serrano, S.M., García-Herrera, R., Peña-Angulo, D., Tomás-Burguera, M., Domínguez-Castro, F., Noguera, I., Calvo, N., Murphy, C., Nieto, R., Gimeno, L., 2022b. Do CMIP models capture long-term observed annual precipitation trends? Clim Dyn 58, 2825–2842.

Wieckowski, J., Kizielewicz, B., Salabun, W. (2024). A multi-dimensional sensitivity analysis approach for evaluating the robustness of renewable energy sources in European countries. *Journal of Cleaner Production* 469: 143225. https://doi.org/10.1016/j.jclepro.2024.143225

Yuan, Y., Ong, Y. S., Gupta, A., & Xu, H. (2017). Objective reduction in many-objective optimization: evolutionary multiobjective approaches and comprehensive analysis. *IEEE Transactions on Evolutionary Computation*, *22*(2), 189-210.

Zass, R., & Shashua, A. (2006). Nonnegative sparse PCA. *Advances in neural information processing systems*, *19*.

Zheng, J., Lienert, J. (2018). Stakeholder interviews with two MAVT preference elicitation philosophies in a Swiss water infrastructure decision: Aggregation using SWING-weighting and disaggregation using UTA. *European Journal of Operational Research* 267: 273-287. https://doi.org/10.1016/j.ejor.2017.11.018

Disclaimer

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union, REA or the Swiss State Secretariat for Education, Research and Innovation (SERI). Neither the European Union, SERI nor the granting authority can be held responsible for them.

Acknowledgement of funding

Funded by the European Union

Project funded by
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Confederation
Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101059264 as well as the Swiss State Secretariat for Education, Research and Innovation (SERI).