



STARS 4 Water

EU-Workshop report on the use of STARS4Water data-driven models

**Deliverable 6.8
September 2026**

EU-workshop report on the use of STARS4Water data-driven models

Lead beneficiary	Jülich Research Centre – FZJ
Lead author(s)	Leandro Avila and Stefan Kollet
Work Package	6: Dissemination and communication
Due date	30 September 2026
Submission date	18 September 2026
Contributors	Héctor Aguilera (IGME), Victor Gomez (UCM), Pedro Martinez (UCM), Helen Baron (UKCEH), Kolbjørn Engeland (NVE), Hege Hisdal (NVE), Trine Hegdahl (NVE), Sami Miaari (FZJ), Stefan Kollet (FZJ), Leandro Avila (FZJ)

Dissemination Level		
PU	Public	X
SEN	Confidential, only for members of the consortium and the granting authority (including other EU institutions and bodies)	
CI	Classified, as referred to EU Decision 2015/444 and its implementing rules	

Version log			
Version	Date	Released by	Nature of Change
0.7	13.09.2026	Leandro Avila	Draft
0.8	21.09.2026	Marek Giełczewski	Review
	21.09.2026	Pedro Martinez-Santos	Review
0.9	23.09.2026	Leandro Avila	Final draft
1.0	23.09.2026	Harm Duel	Accepted

Citation

Avila, L. and S. Kollet, 2026: EU-workshop report on the use of STARS4Water data-driven models. Deliverable D6.8.



The STARS4Water project has received funding from the European Union's Horizon Europe research and innovation program under the Grant Agreement No 101059372

Disclaimer

The content of this deliverable does not reflect the official opinion of the European Union. Responsibility for the information and views expressed herein lies entirely with the author(s).

Summary

This report documents the results of a STARS4Water workshop on the potential for upscaling data-driven models to other river basins and the applicability of these models to ‘what-if’ scenarios for EU-wide assessment of climate change impacts. The workshop was held on 5 February 2026 and brought together researchers and water-sector stakeholders to share recent developments in data-driven modelling and to discuss their potential for transfer and upscaling beyond the river basin scale. The event provided an opportunity to present ongoing work under WP4 and WP5 and to exchange views on how machine-learning approaches can support large-scale water resources assessment and *what-if* scenario analysis.

Four technical presentations covered complementary data-driven approaches for key water-cycle variables. These included an ensemble LSTM approach for large-scale Water Table Depth (WTD) applications, the MLMapper 2.0 tool for spatial prediction of hydrogeological variables, deep-learning frameworks based on ConvLSTM for Snow Water Equivalent (SWE) and WTD, and a transferable deep-learning model for groundwater storage prediction. The examples presented included potential transfers between multiple regions inside and outside Europe. The approaches demonstrated different pathways for extending modelling capabilities to new regions, including transfer learning, fine-tuning, the use of globally available forcing and satellite datasets, and multi-site training.

The open discussion also focused on the use of data-driven models for *what-if* scenario analysis in the context of climate change. Participants considered how surrogate models could support the analysis of climate-change impacts and alternative water-management strategies and discussed the data and adaptation requirements associated with applying such approaches across different regions. Attention was given to groundwater and snow cover in large-scale European modelling frameworks due to their importance for water availability, drought and flood dynamics, and climate adaptation.

Overall, the workshop demonstrated the potential of data-driven approaches to bridge basin-scale modelling and broader-scale water resources assessment. It also highlighted that successful upscaling requires consideration of hydroclimatic similarity, data availability, and the appropriate degree of local adaptation. The workshop therefore served both to communicate progress within STARS4Water and to identify priorities for the further development, evaluation, and uptake of the presented approaches.

Table of contents

Summary	ii
1. Introduction	4
1.1. STARS4Water project	4
1.2. Content of this report.....	5
2. Workshop on potential for upscaling data-driven models	7
3. Workshop approach.....	9
3.1. Main goals and agenda of the workshop	9
3.2. Workshop participants	10
3.3. Overview of the data driven model presentations	11
4. Discussions and main conclusions.....	14
References.....	15
Annex 1: Presentations	16
A. Global Upscaling of Ensemble LSTM Using True Water Table Depth Observation	16
B. Transferable deep learning model for groundwater storage prediction.....	22
C. Application of deep-learning frameworks to estimate groundwater levels and daily snow water equivalent over Europe	27
D. Spatial Prediction of Hydrogeological Variables using ML: Examples of transfer to different regions	32

1. Introduction

1.1. STARS4Water project

The STARS4Water project (Supporting Stakeholders for Adaptive, Resilient, and Sustainable Water Management) is a 4-year initiative funded by the Horizon Europe Program. It focuses on advancing the understanding, observation, and monitoring of water resource availability, contributing to the objectives of the European Green Deal and EU water policies. The project pursues two main goals:

- Enhancing knowledge of how climate change impacts water resources at the river basin level, while identifying vulnerabilities affecting ecosystems, communities, and economic sectors.
- Creating and delivering innovative data services and models to empower decision-makers with tools for planning and implementing adaptive, resilient, and sustainable freshwater management strategies.

Through these efforts, STARS4Water seeks to bridge knowledge gaps and provide actionable insights for sustainable water resource governance. To achieve this, the STARS4Water project is built around two core components, as shown in Figure 1.1. The first focuses on fostering collaboration with seven distinct river basin hubs (RBHs): the Danube, Drammen, Duero, East Anglia, Messara, Rhine, and Seine. Using a living lab-inspired, co-creation approach, the project engages stakeholders in designing tailored services and models. This participatory process ensures that the tools and solutions developed align closely with stakeholder needs, enhancing their utility and longevity beyond the project's timeframe. The second component emphasizes the integration of models into existing tools used for river basin management and decision-making. By leveraging these innovations, the project enables more accurate forecasting of water resource availability and introduces fresh perspectives on the complex relationships between water, nature, and society. These advancements expand the scope of indicators available, supporting more informed and holistic water management strategies.

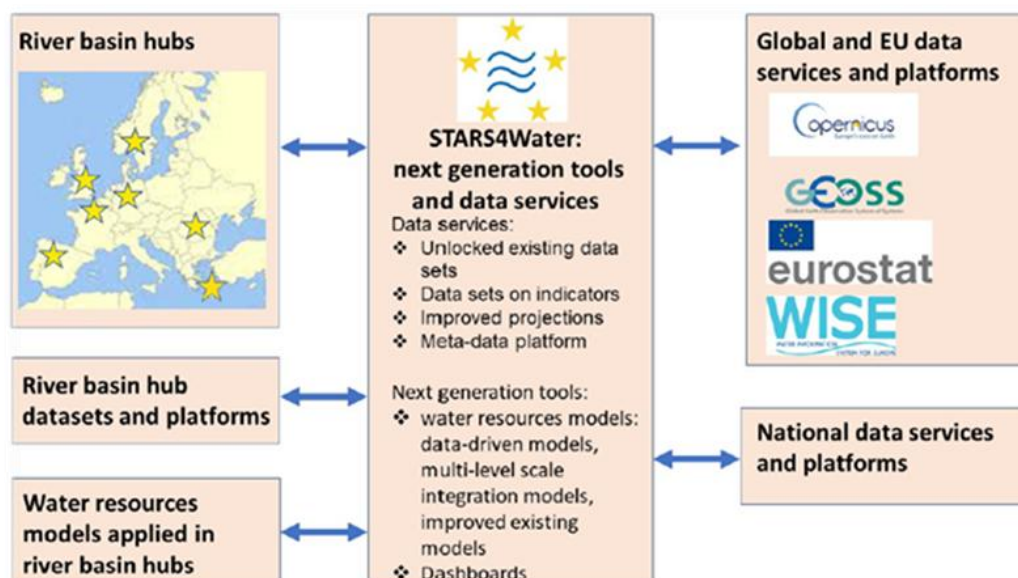


Figure 1.1: Overview of STARS4Water activities within the context of stakeholders and data providers.

Figure 1.2 presents the structure of the activities in the project. The project developed a set of data-driven models (WP3) based on the identified needs and requirements from river basin stakeholders (WP1). This report summarises the outcomes of a workshop organized to explore and discuss the potential for upscaling data-driven models to support integrated assessment of water resources across EU (WP4) and beyond (WP5).

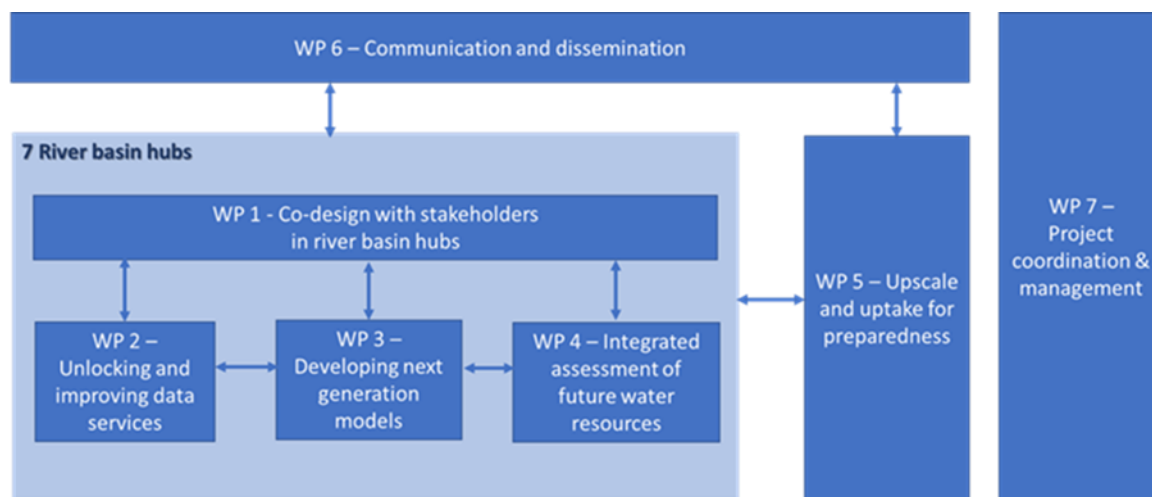


Figure 1.2. Work package structure of the STARS4Water project.

The STARS4Water project is organized in six work packages (WPs) that structure its activities. WP1 gathers the needs and requirements of the seven river basin hubs through a stakeholder co-creation approach, which guides the development of a new generation of data-driven models in WP3, supported by the observation and data infrastructure assembled in WP2. These models feed the integrated assessment of future water resources carried out in WP4, which also evaluates and upscales climate-related water risks and impacts to the European level (Task 4.4), while WP5 synthesizes the results of these assessments and upscales the modelling experiences to other river basins outside Europe (Task 5.2). WP6, to which the present Deliverable 6.8 belongs, is dedicated to the dissemination, communication, and uptake of the project outcomes. Within this structure, this report documents the EU workshop on the use of STARS4Water data-driven models, a communication activity that closes the circle between the technical developments of Task 4.4 and Task 5.2 and the dissemination of the project's results to researchers and water-sector stakeholders.

1.2. Content of this report

This document is organized in the following chapters:

- Chapter 2 is setting the scene of the workshop and highlights the context and ambitions of the workshop.
- Chapter 3 describes the workshop approach, including the date, format, objectives, participant profile, and agenda. It also summarises the four technical presentations and explains how they were followed by an open discussion involving researchers and stakeholders.
- Chapter 4 presents the main conclusions emerging from the workshop discussions, including the value of data-driven models as complementary tools to physically based modelling,

particularly in terms of computational efficiency, flexibility, and transferability. It then discusses their potential for upscaling to the European level and the importance of globally available datasets for enabling applications across large spatial domains. The chapter also discusses next steps, and it links the workshop outcomes to the ongoing activities of Tasks 4.4 and 5.2.

- Finally, the annexes contain the slides of the four presentations delivered during the workshop, which complement the discussion and conclusions documented in the main body of the report.

More details about the data-driven model presented in this report can be found in the project deliverables D4.4 (Avila et al., 2026a) and D5.4. (Avila et al., 2026b).

2. Workshop on potential for upscaling data-driven models

Climate change and increasing socio-economic pressures are altering water availability and increasing the need for improved tools to support water resources assessment and management. Changes in precipitation, temperature, snow accumulation and melt, groundwater recharge, and water demand can affect water availability and increase risks to ecosystems, communities, agriculture and energy production. Addressing these challenges requires modelling approaches that can provide reliable information across different spatial and temporal scales and that can support the development of adaptive and resilient water management strategies.

The STARS4Water project addresses these challenges by combining advanced modelling approaches, data services, and stakeholder co-creation. The project focusses on seven European river basins, where modelling activities are developed and assessed in collaboration with stakeholders. The project seeks both to improve understanding of how climate change affects water resources at the river basin level and to develop innovative data services and models that can support decision-making for adaptive and sustainable freshwater management (Figure 2.1).

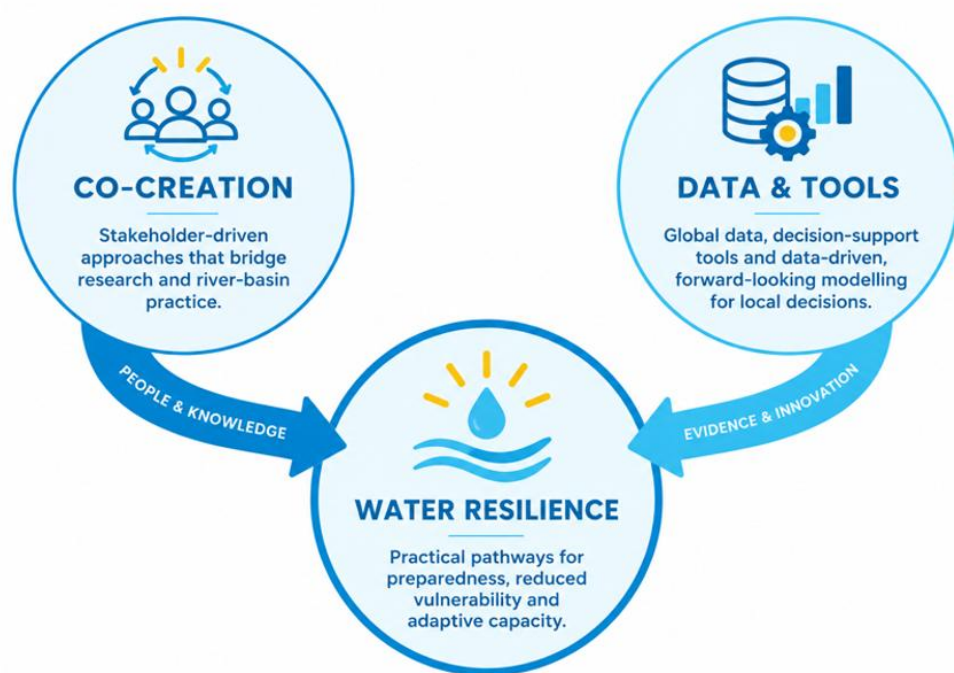


Figure 2.1. Supporting water resilience through data innovation and stakeholder co-creation.

An important challenge within this framework is determining how knowledge and modelling capabilities developed at the river basin scale can be extended to larger spatial domains and to regions where equivalent modelling resources or observational data may not be available. Work Package 4 (WP4) focuses on the integrated assessment of future water resources, including the assessment and upscaling of climate-related water risks (Task 4.4). In parallel, Task 5.3 examines how experiences and methodologies developed within the project can be transferred and upscaled to other river basins beyond Europe.

Data-driven and machine-learning approaches provide one potential pathway towards this objective. Once trained, machine-learning models can provide predictions efficiently and can therefore support applications requiring repeated simulations, rapid assessments, or exploration of alternative scenarios. The increasing availability of spatially extensive datasets, including meteorological reanalysis products and satellite observations, further creates opportunities to apply data-driven approaches beyond the regions in which they were initially developed. At the same time, transferability is not guaranteed: the need for local observations, fine-tuning, or bias correction depends on the characteristics of the target region. Within this context, a workshop was organized to discuss the potential of these approaches for transfer and upscaling. Its objectives, format, and main outcomes are documented in the following chapters.

The main objective of the workshop was to discuss how the presented approaches could be used in broader water resources assessments and decision-support applications, in particular for *what-if* scenario analysis under climate change. Because machine-learning or data-driven models can provide predictions substantially faster than computationally demanding process-based simulations, they offer an opportunity to explore a larger number of alternative climate and water-management scenarios and to assess their potential implications for water availability and related risks.

The workshop placed particular emphasis on groundwater and Snow Water Equivalent (SWE), two components of the water cycle that remain comparatively underrepresented in large-scale European modelling frameworks despite their importance for water resources management and climate-change adaptation. This report documents the workshop, its technical contributions, and the main points emerging from the subsequent discussions. It does not aim to provide a detailed technical description of each individual modelling approach; rather, it records the context in which the approaches were presented, the potential identified for their transfer and upscaling, and the issues raised through interaction with researchers and stakeholders. The report also relates the workshop discussions to the continuing activities of STARS4Water, particularly those concerned with upscaling, *what-if* scenario development, and the further refinement and uptake of data-driven modelling approaches. The workshop provided input for activities in Task 4.4 (Upscaling of basin-wide assessment of climate change risks and impacts to EU-level) and Task 5.2 (Upscaling to other river basins).

3. Workshop approach

3.1. Main goals and agenda of the workshop

The workshop was held online on 5 February 2026 under the title “*What-if Scenarios for EU-wide Assessment of Data-Driven Models*”. It was organized by Forschungszentrum Jülich (FZJ) within the STARS4Water project to share progress in data-driven modelling and to discuss, with researchers and water-sector stakeholders, the potential for transferring and upscaling these approaches to new river basins and regions. The workshop agenda consisted of four technical presentations followed by an open discussion session (Table 3.1). Each presentation was allocated 15 minutes, with the final 55 minutes dedicated to discussion, identification of next steps, and wrap-up.

Table 3.1. Workshop agenda.

Time	Presentation	Presenter
14:00 – 14:05	Welcome & Introduction	Harm Duel (Deltares)
14:05 – 14:20	Global Upscaling of Ensemble LSTM Using True Water	Sami Miaari (FZJ)
14:20 – 14:35	Spatial prediction of hydrogeological variables using ML	Victor Gomez (UCM)
14:35 – 14:50	Application of deep-learning frameworks to estimate SWE and WTD	Leandro Avila (FZJ & NVE)
14:50 – 15:05	Transferable deep learning model for groundwater storage	Héctor Aguilera (IGME)

The presentations covered complementary machine-learning approaches addressing groundwater, snow, hydrogeological variables, and groundwater storage. The first presentation, *Global Upscaling of Ensemble LSTM Using True Water Table Depth Observation*, introduced an approach for extending ensemble LSTM-based groundwater prediction using WTD observations. The second, *Spatial Prediction of Hydrogeological Variables using ML*, presented MLMapper 2.0 demonstrated its application in central Chile (Aliaga et al 2026) and southern Madagascar. The third, *Application of deep-learning frameworks to estimate SWE and WTD*, presented ConvLSTM-based surrogate models for daily SWE and WTD and discussed the transfer of the SWE model from Norway to the western United States (Rocky Mountains). The fourth, *Transferable deep learning model for groundwater storage*, presented a data-driven approach based on GRACE/GRACE-FO and ERA5-Land and discussed its potential for application across regions, together with preliminary results for reservoir storage forecasting. More details related to the outputs and discussion for each modelling approach can be found in D4.4 (Avila et al., 2026a) and D5.3 (Avila et al., 2026b), respectively.

The main purpose of these presentations was not only to communicate modelling results but also to provide a basis for discussing the conditions and challenges associated with model transfer and upscale. The discussion addressed the potential of data-driven approaches to extend basin-scale

modelling to larger spatial domains, the role of globally available datasets, and the extent to which local adaptation may be required in target regions.

A further focus of the workshop was the application of data-driven models to *what-if* scenario analysis in the context of climate change. Participants discussed how these models could support the rapid exploration of climate-change impacts and alternative water-management strategies. The exchange with stakeholders also provided an opportunity to consider practical requirements, data-availability constraints, and priorities for the further development and uptake of the presented approaches within STARS4Water.

3.2. Workshop participants

The workshop brought together 26 participants from 13 institutions and was structured around four technical presentations followed by an open discussion (Table 3.2).

Table 3.2. Participation from institutes and organisations.

Abbreviation	Full name	Country	Participants	Type
FZJ	Forschungszentrum Jülich (Jülich Research Centre)	Germany	8	Research institute
EC / JRC	European Commission – Joint Research Centre	EU	3	Research institute/policy maker
Deltares	Deltares	Netherlands	3	Research institute
BOKU	University of Natural Resources and Life Sciences, Vienna	Austria	2	University
NVE	Norwegian Water Resources and Energy Directorate	Norway	2	Government agency / stakeholder
JKU	Johannes Kepler University Linz	Austria	1	University
IGME	Geological and Mining Institute of Spain	Spain	1	Research institute
UCM	Complutense University of Madrid	Spain	1	University
UKCEH	UK Centre for Ecology & Hydrology	United Kingdom	1	Research institute
INRAE	National Research Institute for Agriculture, Food and the Environment	France	1	Research institute
Landsvirkjun	Landsvirkjun – The National Power Company of Iceland	Iceland	1	Stakeholder
Planet	Planet Labs	USA/international	1	Consultancy
SEVEN	SEVEN – Environmental Consultancy	Greece	1	Consultancy
Total			26	

Participants included representatives from both the research community and water management organisations, ensuring that discussions addressed both the scientific basis and the practical applicability of the presented models. The workshop brought together 26 participants although this number was lower than the total number of invitations issued. While additional researchers and water-sector stakeholders from across Europe were invited, attendance was constrained by scheduling conflicts and availability limitations common to online events of this nature. Consequently, the figures reported in Table 3.2 reflect actual participation on the day of the workshop rather than the broader outreach efforts undertaken during its preparation. Participation was dominated by members of the STARS4Water consortium. Among external stakeholders, only Landsvirkjun, the National Power Company of Iceland, was represented. It should be noted, however, that stakeholder perspectives were also incorporated through several consortium partners that combine research activities with operational water management responsibilities, such as NVE.

While the numerical balance between the research community and the water management stakeholder community was not equal, the level of stakeholder participation was considered sufficient for the purpose and format of the event. The workshop was conceived as a technical and scientific exchange rather than a broad stakeholder consultation, and the limited number of stakeholders did not hinder substantive discussions on the practical applicability, operational relevance, and data requirements of the presented modelling approaches. Nevertheless, this imbalance represents a limitation regarding the representativeness of the conclusions drawn. The practical feedback obtained was largely shaped by the perspectives of the participating institutions, and some of the needs and priorities of the wider European water management community may therefore not be fully reflected in this report. Within this context, the workshop remained focused on identifying technical priorities and opportunities for the transfer, upscaling, and what-if application of data-driven models. These priorities are being further refined and evaluated through Tasks 4.4 and 5.3, as well as through other project dissemination and stakeholder engagement activities.

3.3. Overview of the data driven model presentations

This section provides a concise technical summary of each of the four presentations delivered during the workshop. More details are presented in deliverables D4.4 (Avila et al., 2026a) and D5.3 (Avila et al., 2026b). All slides are presented in Annex 1.

Global Upscaling of Ensemble LSTM Using True Water Table Depth Observations

This presentation introduced an ensemble of Long Short-Term Memory (LSTM) networks designed to predict water table depth (WTD) at global scale, addressing the scarcity and uneven distribution of groundwater observations and the difficulty of assessing prediction quality in unobserved regions (Miaari et al., 2026). The models were trained based on TSMP simulations throughout Europe using dynamic forcing variables and static attributes, with the ensemble spread intended to serve as a quantitative proxy for prediction uncertainty. Transferability was evaluated by applying the pretrained ensemble to globally available ERA5 reanalysis forcing and validating the predictions against independent local WTD observations across Europe.

The results indicated that the model reproduced the hydrological dynamics of water-table fluctuations and that increasing the spatial diversity of the training data improved transfer performance, reducing RMSE by about 12% and increasing the Pearson correlation by about 17%. Remaining biases were attributed mainly to scale mismatches between the gridded model inputs (~12 km resolution) and point-scale observations, as well as to distributional differences between the TSMP training dataset

and the ERA5 evaluation inputs. Ensemble spread (variance and interquartile range) showed a strong correlation with RMSE and mean bias, supporting the use of spread-skill relationships to map prediction confidence in regions lacking observations. Planned next steps include global application of the pretrained ensemble forced with ERA5 and its evaluation against independent observations.

Transferable Deep Learning Model for Groundwater Storage Prediction

This presentation described a transferable deep-learning surrogate model for groundwater storage prediction, developed to meet stakeholder needs for refined groundwater resource estimation under what-if management scenarios, with the Duero aquifer basin as the pilot case study (Aguilera et al., 2026). The modelling workflow combined data correction and a spatio-temporal Transformer (STT): total water storage from terrestrial system simulations was bias-corrected against local observations through spatiotemporal kriging to generate groundwater storage change (GWSC) training targets, and the STT was trained to provide multi-step (12-month) GWSC forecasts from an input window of smoothed historical data, using concatenated spatial embeddings to capture spatiotemporal dependencies and spatial cross-validation to prevent data leakage.

The model performed well during training ($R^2 \sim 0.87$) and showed moderate generalization on validation and test sets ($R^2 \sim 0.53$, MAE ~ 0.03 m). Transferability tests in the Seine and Messara basins were constrained by weak correlations between simulated WTD and total water storage data and local observations, motivating the exploration of alternative global datasets. The presentation outlined applications for GWSC-based what-if analysis, including seasonal forecasts to support groundwater allocation strategies, response to future climate scenarios, and management options dependent on water-use data. Next steps include continued transfer assessments in Seine and Messara, probabilistic uncertainty quantification, and refinement of the modelling targets (e.g., groundwater levels instead of storage).

Application of Deep-Learning Frameworks to Estimate Groundwater Levels and Daily Snow Water Equivalent over Europe

This presentation reported on ConvLSTM surrogate models developed to emulate physically based hydrological models over extended spatial domains, motivated by the limitations of point-focused machine-learning models for large-area forecasting (Avila et al., 2026c). A first ConvLSTM was trained in the Drammen River Basin and later upscaled to Norway based on the SeNorge dataset, and used to forecast daily snow water equivalent (SWE) up to seven days ahead using an autoregressive scheme. Evaluation against in-situ observations showed generally good performance. The framework was also transferred to the Alps by retraining it on LISFLOOD model outputs, where the ConvLSTM outperformed LISFLOOD at several stations.

A second ConvLSTM was developed to estimate water table depth anomalies (WTDa) over Germany from ParflowCLM simulations, using soil moisture and antecedent precipitation as dynamic inputs and topography, slope, and soil textural properties as static inputs. The model reproduced WTDa accurately in shallow groundwater regions, but with lower skill in deep groundwater settings. Planned work includes transferring the ConvLSTM frameworks to regions outside Europe and implementing what-if scenarios involving perturbed precipitation and temperature regimes for snow and simulated pumping for groundwater.

Spatial Prediction of Hydrogeological Variables using ML: Examples of Transfer to Different Regions

This presentation introduced MLMapper 2.0, a Python-based machine-learning tool for the spatial prediction of hydrogeological variables using supervised classification or regression, trained on point datasets and sets of explanatory variables (Gomez et al., 2026; Martinez et al., 2026). The tool applies multicollinearity screening, hyperparameter fitting, variable selection, and scaling and resampling techniques, and it benchmarks 19 supervised classification algorithms to produce spatially distributed predictive maps by identifying areas similar to locations with known outcomes. Within STARS4Water, nitrate contamination classification in East Anglia and the Duero basin achieved accuracy rates of approximately 90% and 85%, respectively, while regression estimates in East Anglia reached a R^2 of about 0.70; key predictors included aquifer geometry and soil properties, distance to agricultural areas, and nitrogen pressures.

Transfer of the approach to southern Madagascar, applied to groundwater potential for solar-powered pumping and to water-quality prediction based on electrical conductivity, achieved success rates of about 80% and 75% using datasets of 775 and 2,314 boreholes, respectively, with tree-based ensemble algorithms consistently showing the best performance. Ongoing work includes the integration of quality and quantity predictions into an integrative groundwater map for Madagascar and knowledge transfer between Spanish basins (Duero to Ebro). The presentation concluded that transferability depends chiefly on the availability and quality of input data in the target region and on the adaptation of target and explanatory variables to the specific hydrogeological context.

4. Discussions and main conclusions

The workshop provided an opportunity to present and discuss recent advances in data-driven modelling within STARS4Water and to examine their potential for application beyond the river basin considered in this project. The presentations and the subsequent discussion with researchers and water-sector stakeholders showed that machine-learning approaches can complement physically based models for large-scale assessment, while also offering opportunities to transfer modelling capabilities between regions with different hydroclimatic conditions.

A central conclusion of the workshop was that data-driven models can provide a valuable complement to physically- based hydrological and hydrogeological models. Physically- based approaches remain important for representing water-cycle processes, but their application over large spatial domains can require substantial computational resources, detailed parameterization, and basin-specific calibration. These requirements can make rapid assessment and repeated scenario exploration challenging, particularly in regions where detailed process information or monitoring data are limited.

The models presented during the workshop demonstrated that, once trained, machine-learning and surrogate models can provide predictions efficiently and can therefore support applications such as rapid assessment, ensemble simulations, and *what-if* scenario analysis. For this last one, the participants concluded to focus on analyzing what-if-scenarios for Snow Water Equivalent (SWE) and Water Table Depth (WTD). Their flexibility of the presented models also allows transferability from one river basin or region to other regions with similar spatial and hydroclimatic contexts. Finally, the workshop discussions made clear that data-driven models should be viewed as complementary tools rather than replacements for physically based modelling.

The examples also illustrated different pathways for achieving upscaling, including transfer between contrasting regions, fine-tuning with local data, and the use of globally available forcing datasets. At the same time, the discussion emphasized that upscaling cannot be understood as simply application the same model everywhere. Model performance depends on hydroclimatic similarity, spatial scale, data availability, and, for some applications, operational characteristics. The appropriate balance between direct transfer and local adaptation is therefore a key consideration for EU-wide application and transferability outside Europe.

Attention during the workshop was given to groundwater and Snow Water Equivalent (SWE), discussing how to improve the representation of those hydrological variables using machine learning models. The presented applications showed how machine-learning models can be combined with high-resolution simulations and reanalysis products to estimate groundwater and SWE over larger spatial domains, developing tools to complement existing modelling frameworks and improve the representation of these hydrological variables in EU-scale assessments as well as in regions outside Europe.

Improving the representation of these hydrological variables is therefore relevant not only for monitoring current conditions but also for assessing potential changes under climate change. In this context, the workshop discussion highlighted the potential of data-driven models to support *what-if* scenario analysis, allowing a larger number of climate and water-management scenarios to be explored and helping to assess their potential consequences for water availability and related risks.

References

- Aguilera, H., Gómez-Escalonilla, V., García Tricás, E., García Menéndez, O., de la Hera-Portillo, Á., Rodríguez del Rosario, M., and Martínez-Santos, P. (2026): A Deep Learning surrogate for groundwater storage change prediction at regional scale (Duero River Basin, Spain), EGU General Assembly 2026, Vienna, Austria, 3–8 May 2026, EGU26-20548, <https://doi.org/10.5194/egusphere-egu26-20548/>
- Aliaga-Alvarado M, Gómez-Escalonilla V, Martínez-Santos P (2023). Identification of non-conventional groundwater resources by means of machine learning in the Aconcagua basin, Chile. *Journal of Hydrology: Regional Studies*, 49, 101502.
- Avila, L., Aguilera, H., Gomez, V., Martinez, P., Baron, H., Engeland, K., Gelati, E., Hisdal, H., Hegdahl, T., Kollet, S., Salamon, P., Grimaldi, S., & Bisselink, B. (2026a). Upscaling of basin-wide assessment of climate change risks and impacts to EU-level. Deliverable 4.4 of the STARS4Water project. Horizon Europe Grant Agreement No 101059372.
- Avila, L. , Aguilera, H., Gomez, V., Martinez, P., Baron, H., Engeland, K., Hisdal, H., Hegdahl, T., Kollet, S., & (2026b). Upscaling to other river basins. Deliverable 5.3 of the STARS4Water project. Horizon Europe Grant Agreement No 101059372.
- Avila, L., Engeland, K., Jahr, T., and Kollet, S. (2026c). A ConvLSTM surrogate model to predict high-resolution daily snow water equivalent in Norway, EGU General Assembly 2026, Vienna, Austria, 3–8 May 2026, EGU26-8502, <https://doi.org/10.5194/egusphere-egu26-8502>, 2026.
- Gómez-Escalonilla V, Baron H, Read T, Watson J, Rodríguez del Rosario M, Darch G, Keller V, Rickards N, Martínez-Santos P (2026). Spatial prediction of groundwater nitrate through machine learning and stakeholder co-creation: The Chalk Aquifer case in East Anglia, UK. *International Journal of River Basin Management*. DOI: 10.1080/15715124.2026.2706022.
- Martínez-Santos P, Gómez-Escalonilla V, Rodríguez del Rosario M, Díaz-Alcaide S, Martín-Loeches M, Montero E (2026). AI-driven groundwater mapping: systematic review and implications for practical uptake. *Applied Water Science*, DOI: 10.1007/s13201-026-02964-1.
- Miaari, S., Klotz, D., and Kollet, S. (2026). Application of Ensemble LSTM Transfer Learning for Water Table Depth Prediction and Uncertainty Assessment in Data-Scarce Regions, EGU General Assembly 2026, Vienna, Austria, 3–8 May 2026, EGU26-9892, <https://doi.org/10.5194/egusphere-egu26-9892>.

Annex 1: Presentations

A. Global Upscaling of Ensemble LSTM Using True Water Table Depth Observation

Global Upscaling of Ensemble LSTM Using True Water Table Depth Observations



STARS4Water workshop, 5 February 2026

Sami Miaari¹, Daniel Klotz², Stefan Kollet¹

¹Forschungszentrum Jülich

²Interdisciplinary Transformation University Austria

EU Grant Agreement
101059372



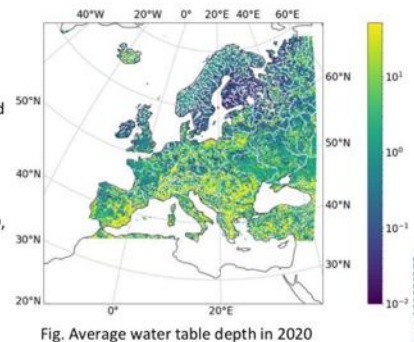
Motivation and objective

Motivation:

Transfer learning applications are a solution for WTD monitoring in locations with sparse and unevenly distributed observations; however, prediction quality assessment is challenging without observations.

Objective:

We leverage an ensemble of LSTM networks to predict WTD, using the ensemble spread to establish spread-skill relationships that quantify prediction quality in unobserved target locations.



2



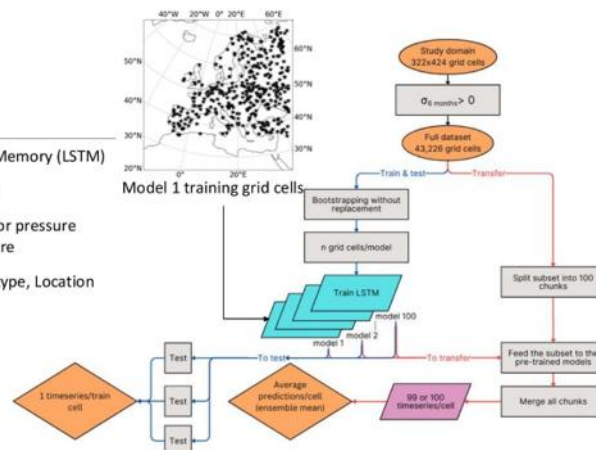
EU Grant Agreement 101059372

Method

Experimental setup

ML architecture	Long-Short Term Memory (LSTM)
Target variable	Water table depth
Dynamic input features	Precipitation, Vapor pressure deficit, Soil moisture
Static input features	Slopes (x&y), Soil type, Location (Lon&Lat)
Time step	Daily
Training period	2002-2015
Testing period	2017-2020
Lookback window	365 days

TSMF simulations

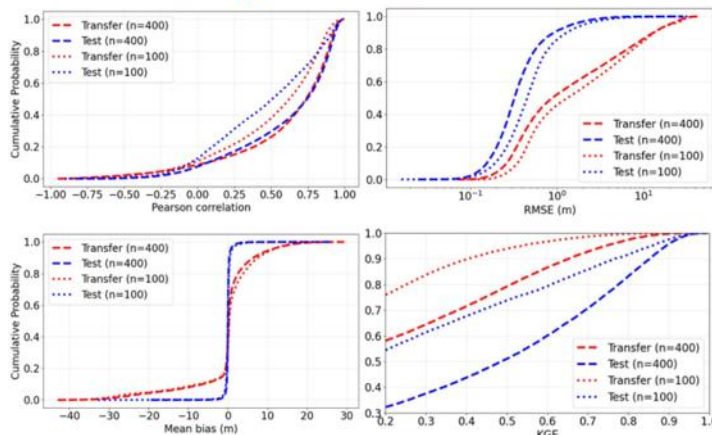


3



EU Grant

Model testing and ensemble transfer

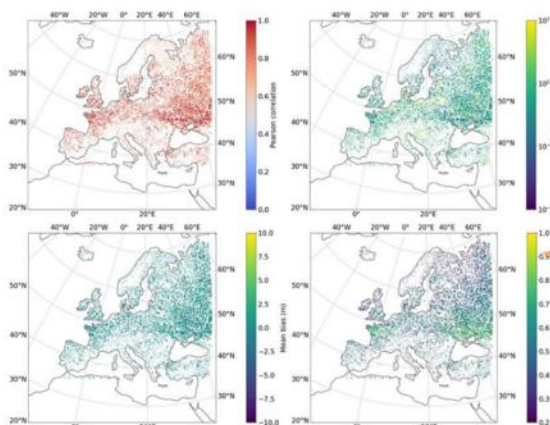


The n=400 outperforms n=100, improved global RMSE and Pearson correlation by 11.52% and 17.42%, respectively. Thus, spatial diversity in training (more grid cells) boosts model robustness and accuracy in transfer scenarios.

The ensemble yields low KGE values for many grid cells. Large negative values are due to the erratic nature of KGE against the low variance of WTD.



Performance metrics for KGE>0.2 in the EU domain



Grid cells with transfer KGE>0.2 for the ensemble of sample size n=400.

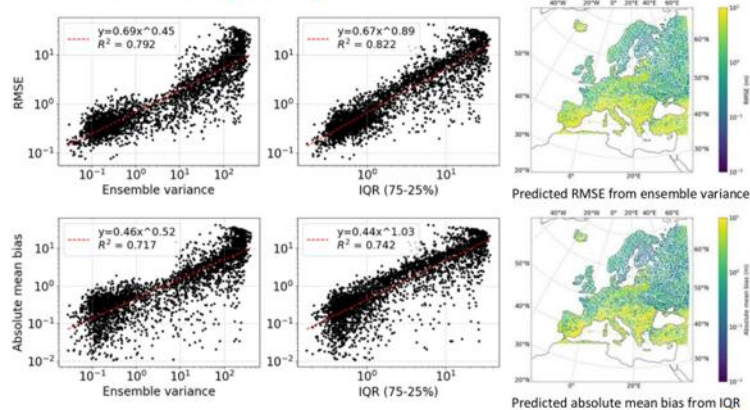
Strong generalization and transfer performance: The model demonstrates effective generalization and transferability across different spatial region

The ensemble provides reliable transfer predictions in specific regions.

- But how can we identify these regions ?
- What if we don't have observations ?



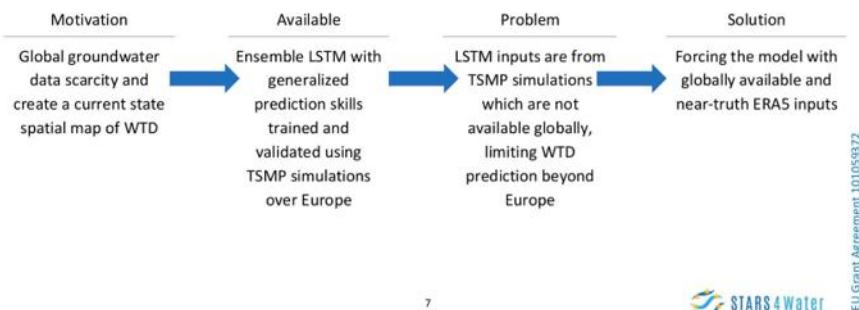
Prediction quality from ensemble statistics



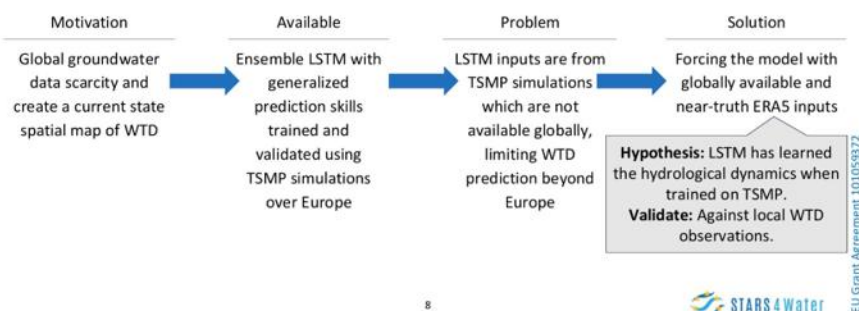
The ensemble spread (ensemble variance and interquartile range) shows a strong correlation with RMSE and absolute mean bias, suggesting that spread can serve as a useful proxy for estimating prediction quality without relying on observations.



Global Upscaling of Ensemble LSTM Using True Water Table Depth Observations

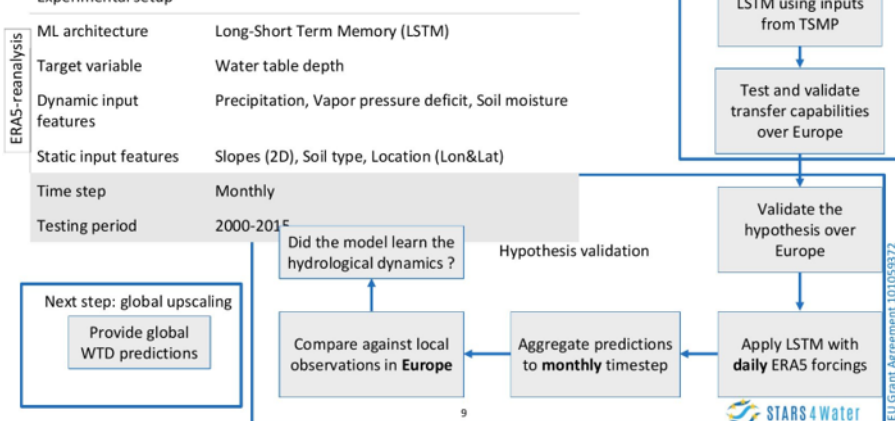


Global Upscaling of Ensemble LSTM Using True Water Table Depth Observations

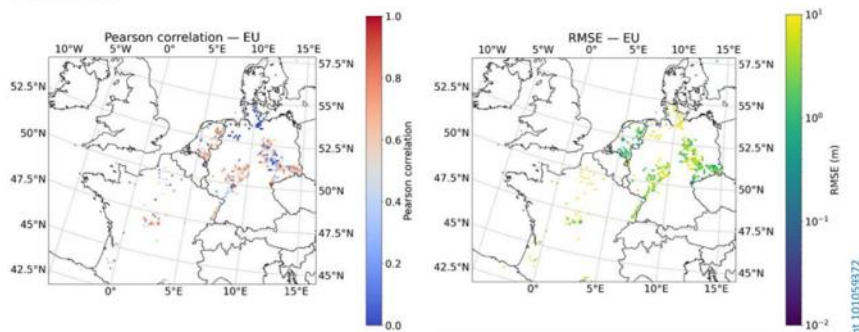


Method

Experimental setup



Results

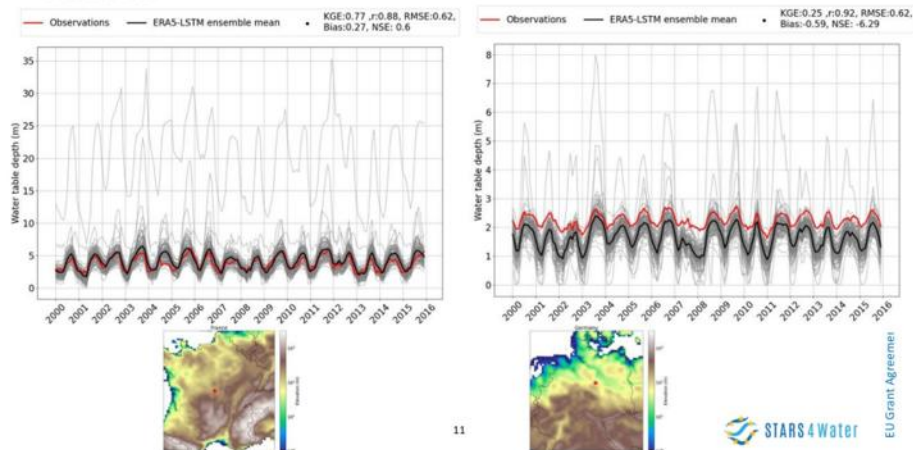


- Good prediction accuracy in a large number of grid cells implies that the model is able to understand the hydrological dynamics.
- The degradation in model performance arises from a distributional mismatch between the training inputs from TSMP dataset and the evaluation inputs from ERA5 dataset.



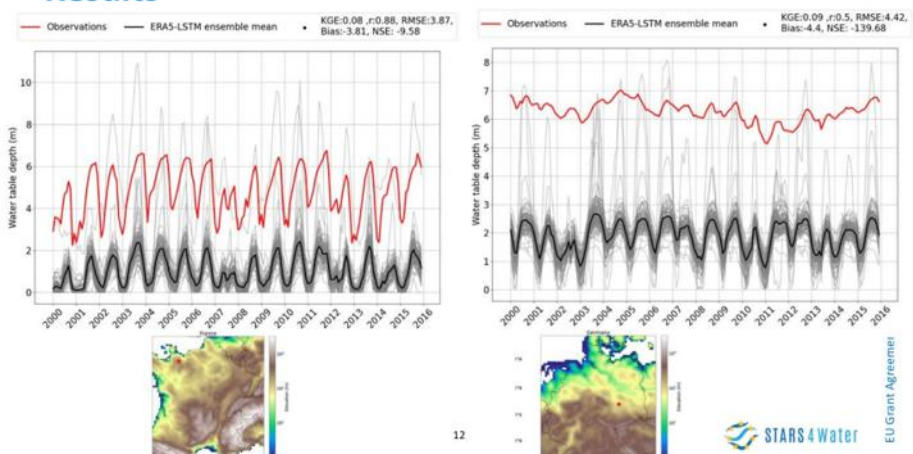
EU Grant Agreement 101059372

Results



EU Grant Agreement

Results



EU Grant Agreement

Conclusion

- The results indicate that the model learned the hydrological dynamics and can reproduce temporal patterns and magnitude of water table depth.
- Prediction biases are observed and are largely attributable to scale mismatches between gridded LSTM inputs (~12 km resolution) and point-scale water table depth observations.
- Model performance degrades in certain locations due to several factors such as the different model input distribution and bias.
- Next steps involve applying the pretrained ensemble LSTM globally using ERA5 inputs and evaluating predictions against independent local observations.
- The transferability of the derived spread-skill relationships will be assessed to enable global estimation of prediction quality in regions lacking observations.

13



EU Grant Agreement 101059372

Thank you for your attention

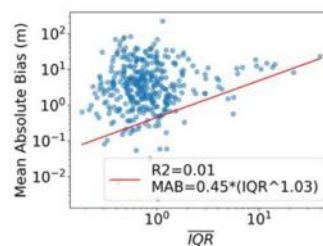
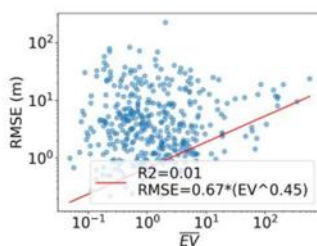


14



EU Grant Agreement 101059372

Additional slides - Spread-skill relationships from TSMP inputs

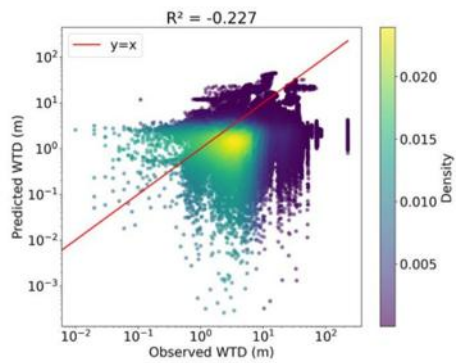


15



EU Grant Agreement 101059372

Additional slides - Results

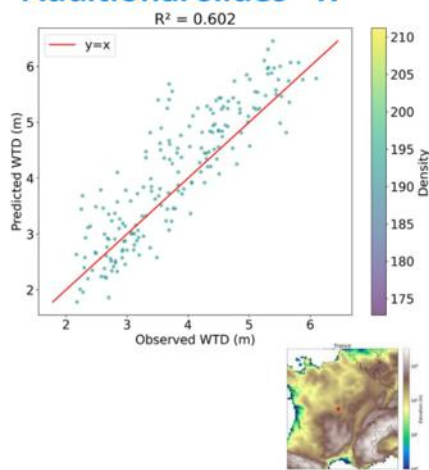


16

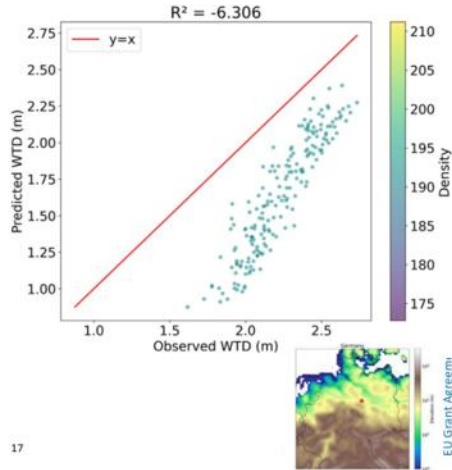


EU Grant Agreement 101059372

Additional slides - R^2

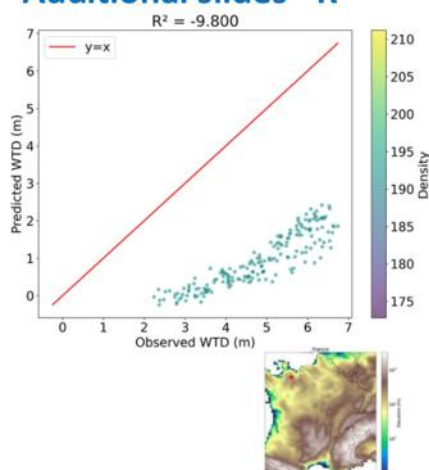


17

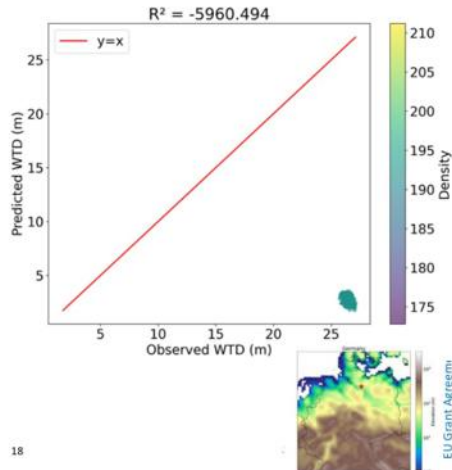


EU Grant Agreement

Additional slides - R^2

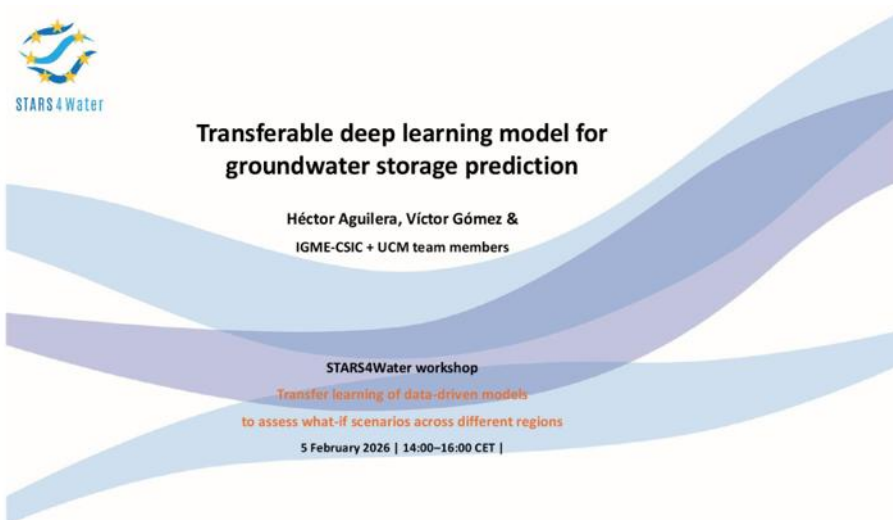


18



EU Grant Agreement

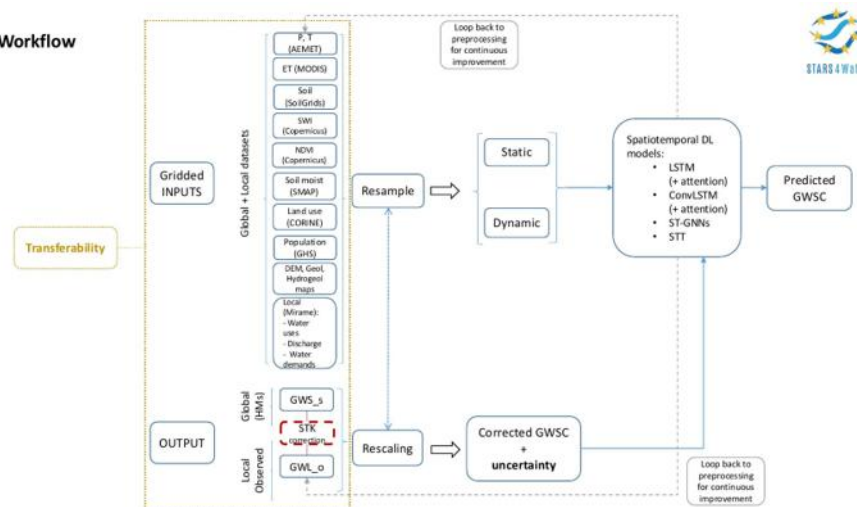
B. Transferable deep learning model for groundwater storage prediction



Objectives

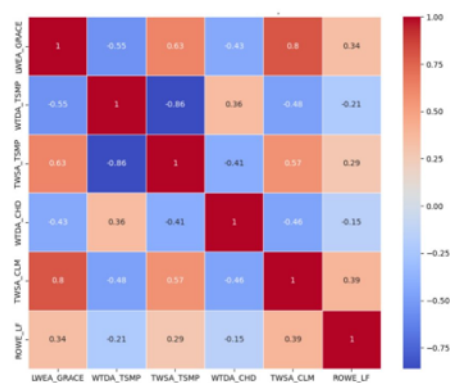
- Give answer to stakeholders's needs: refine estimation of groundwater resources for what/if management scenarios.
- Develop a robust DL surrogate model capable of reproducing spatiotemporal variations in groundwater storage (GWSC): Duero pilot.
- Make the model configuration transferable to other basins.

Workflow

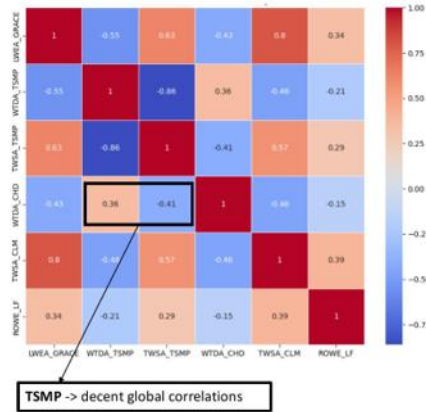


[illegible]

	Category	Item	Value	Unit	Notes
General	Category 1	Item 1	100	kg	Weight of Item 1
	Category 1	Item 2	200	kg	Weight of Item 2
	Category 1	Item 3	300	kg	Weight of Item 3
	Category 1	Item 4	400	kg	Weight of Item 4
	Category 1	Item 5	500	kg	Weight of Item 5
	Category 1	Item 6	600	kg	Weight of Item 6
	Category 1	Item 7	700	kg	Weight of Item 7
	Category 1	Item 8	800	kg	Weight of Item 8
	Category 1	Item 9	900	kg	Weight of Item 9
	Category 1	Item 10	1000	kg	Weight of Item 10
Special	Category 2	Item 11	1100	kg	Weight of Item 11
	Category 2	Item 12	1200	kg	Weight of Item 12
	Category 2	Item 13	1300	kg	Weight of Item 13
	Category 2	Item 14	1400	kg	Weight of Item 14
	Category 2	Item 15	1500	kg	Weight of Item 15
	Category 2	Item 16	1600	kg	Weight of Item 16
	Category 2	Item 17	1700	kg	Weight of Item 17
	Category 2	Item 18	1800	kg	Weight of Item 18
	Category 2	Item 19	1900	kg	Weight of Item 19
	Category 2	Item 20	2000	kg	Weight of Item 20

[illegible]

Selecting output target: comparing global data/models with local observations



Correction of TSM Total Water Storage

1. Spatiotemporal kriging

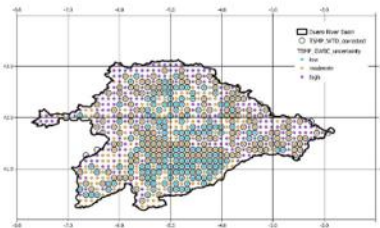
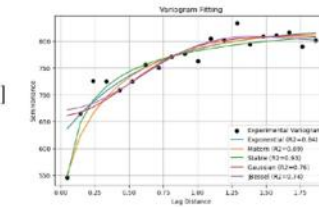
$$WTD_{TSM(corrected)} = WTD_{TSM} + \sum_{i=1}^N w_i [WTD_{0,i} - WTD_{TSM,i}]$$

2. Corrected GWSC targets

$$GWSC_{TSM(corrected)} = S_j \cdot \Delta WTD_{TSM(corrected)}$$

3. Uncertainty

- Low uncertainty: If $(R^2 > 0.7 \text{ \& } S < 0.3)$ & $GWSC_{TSM(corrected)}$ estimate available, $GWSC_{TSM(corrected)}$ assigned as target value.
- Moderate uncertainty: If $(R^2 < 0.7 \text{ or } S > 0.3)$ & $WTD_{TSM(corrected)}$ estimate available, simulated ΔWTD_{TSM} assigned as target value.
- High uncertainty: If $(R^2 < 0.7 \text{ or } S > 0.3)$ & $WTD_{TSM(corrected)}$ estimate not available, simulated ΔWTD_{TSM} assigned as target value.



Best DL model: Spatio-temporal Transformer (STT)

Model Overview

- Purpose:** Multi-step groundwater storage change (GWSC) forecasting in the Duero aquifer basin
- Architecture:** Transformer encoder with concatenated spatial embeddings
- Prediction horizon:** 12 months ahead
- Input window:** 48 months of historical data

Key Design Choices

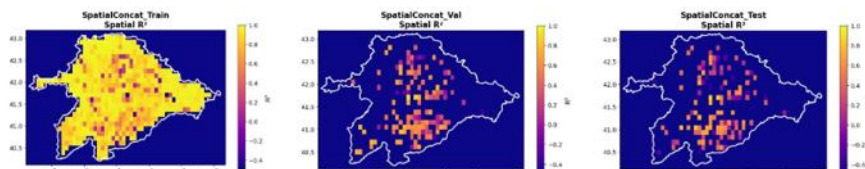
- Late fusion** - Spatial embeddings injected after transformer processing, preserving temporal learning while adding location awareness
- Attention pooling** - Learnable softmax-weighted aggregation dynamically emphasizes most informative timesteps
- Uncertainty-aware** - Observation confidence serves as both input feature and loss weight to handle unreliable measurements
- Yeo-Johnson transform** - Normalizes targets while handling negative values (groundwater depletion vs. recharge)
- Spatial cross-validation** - Shuffled spatial folds prevent data leakage from spatially autocorrelated samples



Performance results



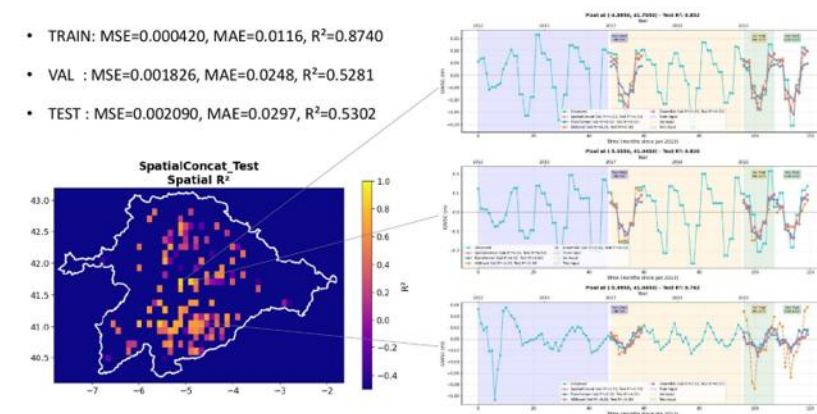
- TRAIN: MSE=0.000420, MAE=0.0116, $R^2=0.8740$
- VAL : MSE=0.001826, MAE=0.0248, $R^2=0.5281$
- TEST : MSE=0.002090, MAE=0.0297, $R^2=0.5302$



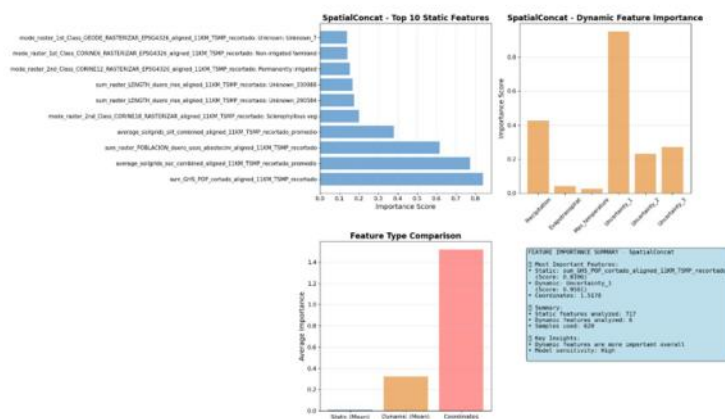
Performance results



- TRAIN: MSE=0.000420, MAE=0.0116, $R^2=0.8740$
- VAL : MSE=0.001826, MAE=0.0248, $R^2=0.5281$
- TEST : MSE=0.002090, MAE=0.0297, $R^2=0.5302$



Feature importance (permutation tests)



Transferability tests: Seine and Messara



Transferability to Other Basins

- ✓ **Transfers Well**
 - Transferable architecture
 - Covert MLP structure
 - Training procedure
 - Loss function design
 - Uncertainty mechanism
- ⚠ **Needs Re-fitting**
 - All scalars (total, dyn, left)
 - Category encoders
 - Hyperparameters (Optional)
 - Covert normalisation
- ▽ **Transfer Options**
 - Full retails: new region
 - Fine-tune: similar basins
 - Zero-shot: same basin, new period



Seine

- Issue -> poor correlations:
 - WTD_TSMP vs TWS_TSMP
 - WTD_TSMP vs WTD_Seine_Wells
- Alternative: look for other global model data

What/if scenarios



- Seasonal GWSC forecasts for groundwater allocation strategies (accuracy + spatio-temporal refinement)
- Future climate scenarios.
- Management scenarios (requires water use data): demands, agricultural land use.

Next steps



- Continue transfer assessments in Seine and Messara.
- Future water resources what/if scenarios in the Duero.
- Probabilistic uncertainty quantification.
- Try targeting GWLs instead of GWS.
- Operability to be discussed with stakeholders.

C. Application of deep-learning frameworks to estimate groundwater levels and daily snow water equivalent over Europe

STARS4Water



Application of deep-learning frameworks to estimate groundwater levels and daily snow water equivalent over Europe

Transfer learning of data-driven models to assess what-if scenarios across different regions

Leandro Avila¹, Stefan Kollet¹, Kolbjørn Engeland², Trine Jhar²

FZJ, NVE

05.02.2026



Introduction



- Implement tools based on data-driven models to predict and estimate hydrological variables.
- Most ML models used in hydrology (e.g forecasting water levels) are point-focused models.
- Forecasting hydrological variables for large areas using point-focused ML models can be impractical in some situations.
- Explore algorithms that can be more appropriate over a large area, such as ConvLSTM.

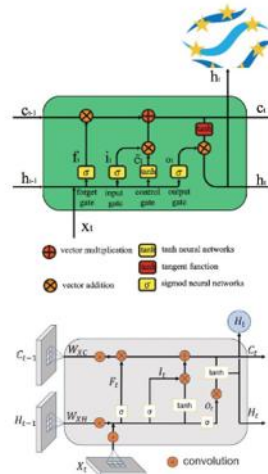
Objectives



- Implement ConvLSTM models that can effectively emulate well-established physical and conceptual hydrological models in different regions across Europe
- Transfer the implemented ConvLSTM to multiple regions to support water resources management.
- Evaluate “what-if” scenarios in the context of climate change

ConvLSTM network

- ConvLSTM model combines the strenghts of CNN and LSTM, to process both spatial and temporal pattern.
- ConvLSTM cell is similar to a LSTM cell but changing internal matrix multiplications with convolution operations.
- CNN-LSTM: A two-stage architecture**
 - A CNN first extracts spatial features from each frame.
 - A LSTM processes the sequences of those features accross time

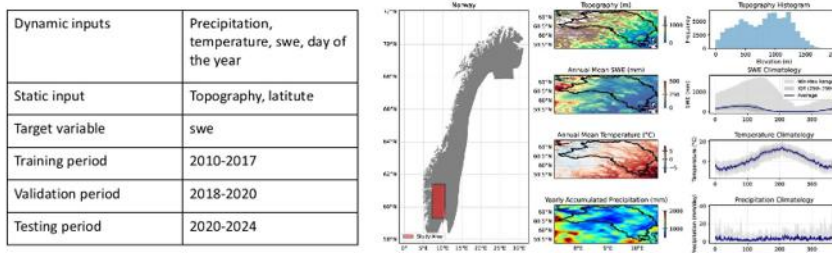


ConvLSTM model to predict SWE



► The SeNorge dataset and SeNorge snow model (1km) are used to train the ConvLSTM model

► Model can predict up to 7 days ahead, using an autoregressive approach

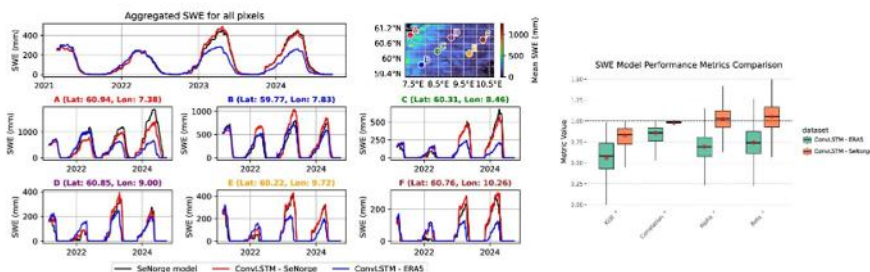


ConvLSTM model to predict SWE



► Evaluation using the two different forcing datasets: SeNorge and ERA5

► Results with ERA5 underestimate daily SWE in some regions. However, general performance is good

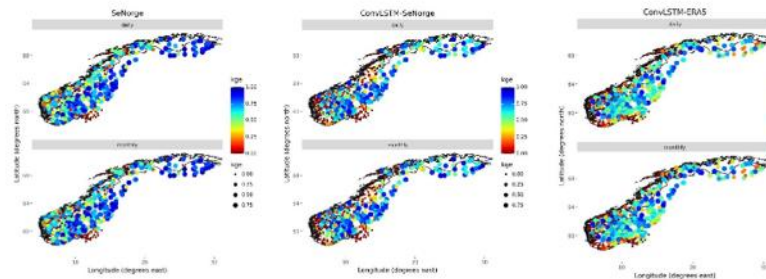


ConvLSTM model to predict SWE



Application of ConvLSTM model for Norway

Comparison with in-situ observations (NH-SWE dataset):
SeNorge outputs – ConvLSTM model forced with SeNorge – ConvLSTM forced with ERA5



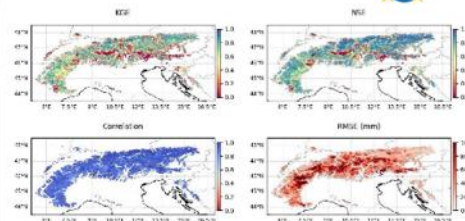
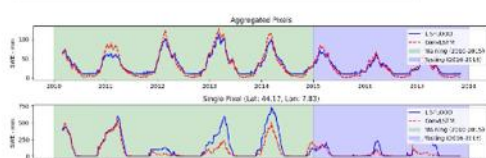
ConvLSTM model to predict SWE



► The proposed ConvLSTM model was transferred in the Alps region

► Lifslood model outputs (Task3.3) are used as input dataset to train the model in the Alps:

Spatial resolution: 0.0166 degrees
Temporal resolution: 6 hourly
Forcing: EMO1 (precipitation and temperature)

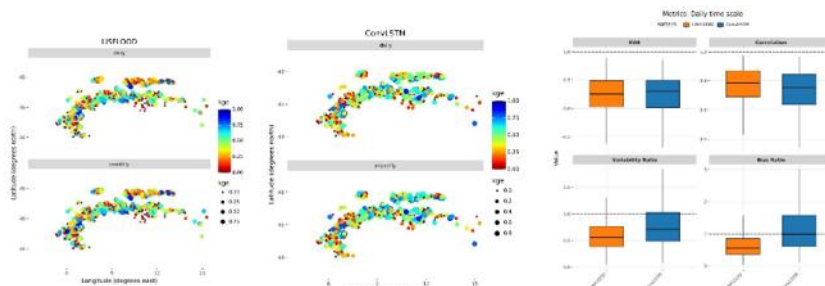


ConvLSTM model to predict SWE



► Comparison with in-situ observations (NH-SWE dataset)

► ConvLSTM outperforms LISFLOOD model at certain stations



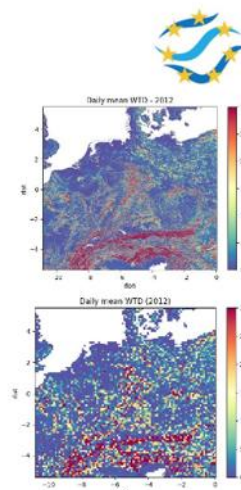
ConvLSTM model to estimate wtda

▶ Parflow-DE056 domain (600m x 600 m - 2000 x 2000 pixels)

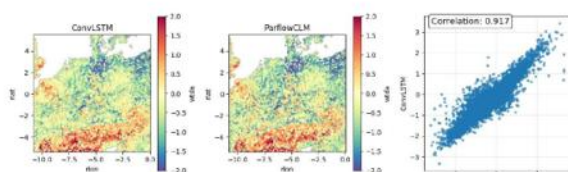
▶ Target resolution: 12km x 12km (100 x 100 pixels)

Note: The ConvLSTM is expected to be implemented for the original resolution

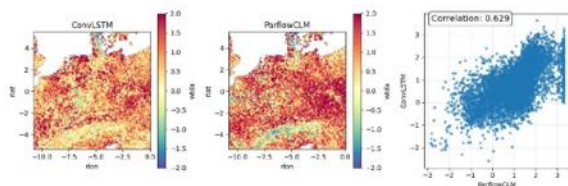
Dynamic inputs	Soil moisture, antecedent precipitation index
Static input	Topography, slope, soil type, CLYPPT, SNDPPT
Target variable	wtda
Training period	2012-2017
Validation period	2018-2021



ConvLSTM model to estimate wtda



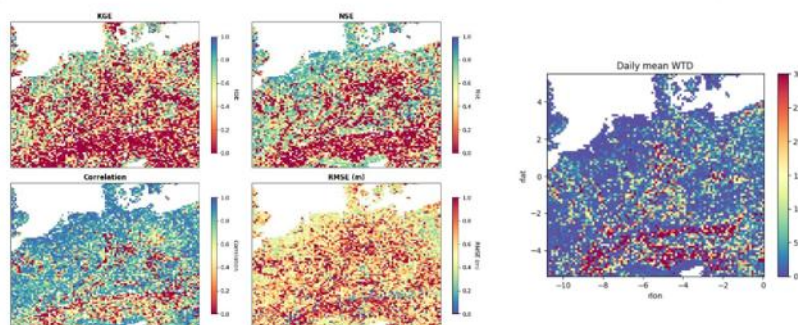
Training period



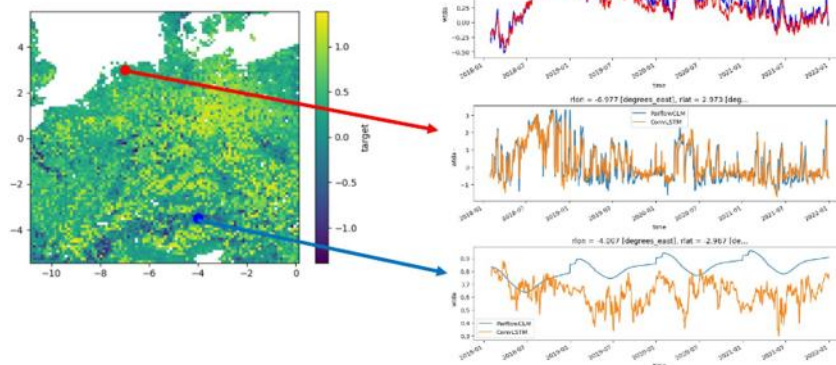
Validation period

ConvLSTM model to estimate wtda

The ConvLSTM model can well estimate WTDA in shallow groundwater regions. Lower performance in deep groundwater regions



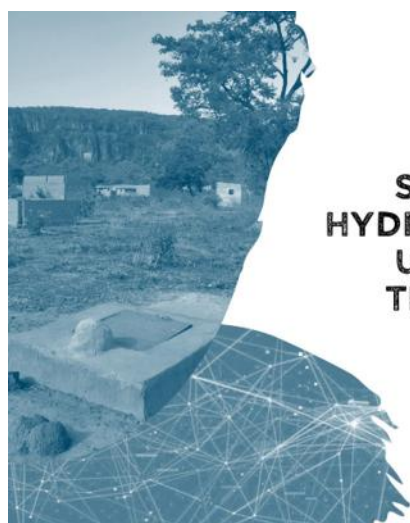
ConvLSTM model to estimate wtda



Conclusions and Next steps

- Transfer and validate the ConvLSTM to predict SWE outside Europe (e.g Rocky Moutains).
- Apply the ConvLSTM to estimate WTDA at a finner resolution and transfer it to other regions across Europe.
- Implement “what-if” scenarios for both models
 - For SWE: Simulate changes in precipitation and temperature patterns
 - For wtda: Include pumping scenarios to evaluate water uses and human impacts.

D. Spatial Prediction of Hydrogeological Variables using ML: Examples of transfer to different regions



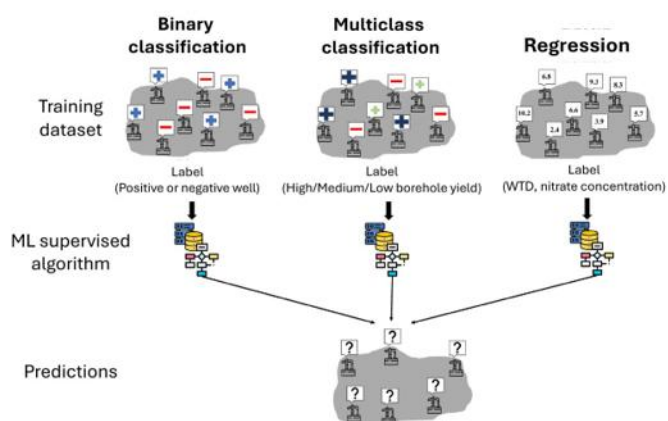
SPATIAL PREDICTION OF HYDROGEOLOGICAL VARIABLES USING ML: EXAMPLES OF TRANSFER TO DIFFERENT REGIONS

Víctor Gómez-Escalonilla Canales, Manuel Rodríguez del Rosario
& Pedro Martínez Santos

UCM, Spain

INTRODUCTION TO THE MODELLING APPROACH

MACHINE LEARNING APPROACHES



INTRODUCTION TO THE MODELLING APPROACH

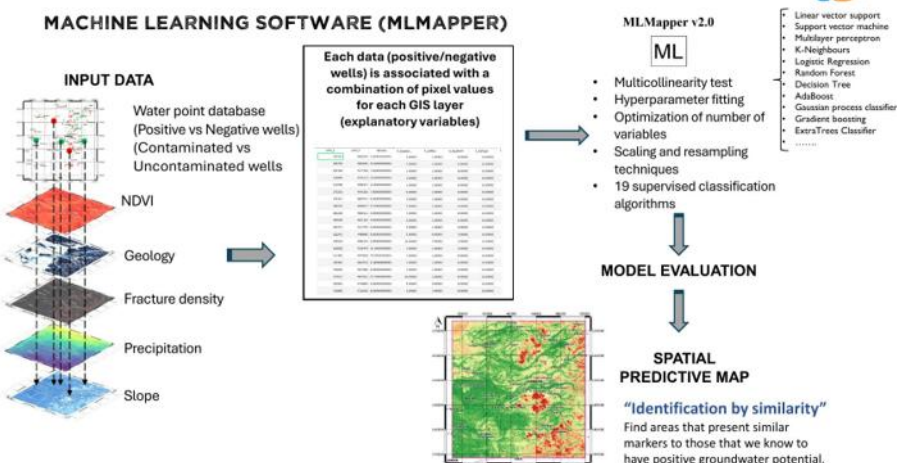
MACHINE LEARNING SOFTWARE (MLMAPPER)



- ❖ **MLMapper 2.0** is an evolution of the original MLMapper code developed by Martínez-Santos and Renard (2019).
- ❖ Programmed in **Python 3.7**
- ❖ MLMapper provides **spatially distributed predictions** for a target variable using:
 - ❖ Supervised classification or regression algorithms
 - ❖ A set of explanatory variables
 - ❖ A training set of points where the outcome of the target variable is known.

ML

INTRODUCTION TO THE MODELLING APPROACH

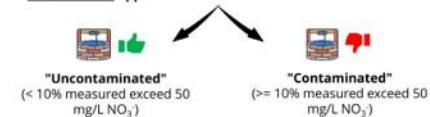


APPLICATIONS WITHIN THE STAR4WATER PROJECT

East Anglia (UK)

- The monitoring network consisted of around 150 wells with at least 10 readings for the period 2020-2024.

Classification approach



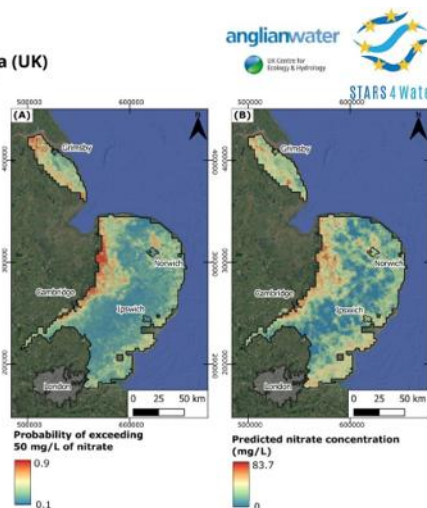
Regression approach

Median value (mg/L) of NO_3^-

- Performance:**
- Classification** - Accuracy rate ~ 90%
 - Regression** - R2 score ~ 0.70

Important variables:

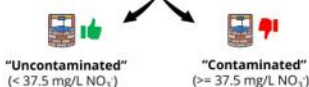
- Thickness of materials Q
- Depth of aquifer base
- CDP density
- Distance to agricultural areas



APPLICATIONS WITHIN THE STAR4WATER PROJECT

Duero basin (Spain)

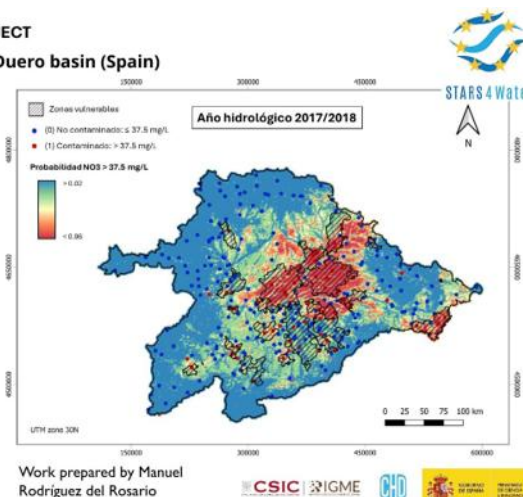
- Monitoring network consisted of around 290 points for the period 2017-2020.



Accuracy rate ~ 85%

Important variables:

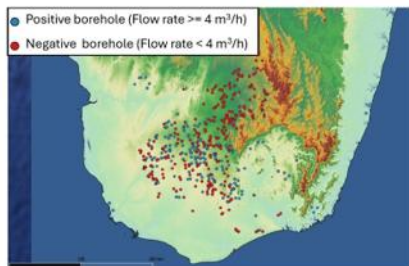
- Distance to surface water bodies
- Diffuse nitrogen pressures
- Geology
- Sand and coarse fragment content in the soil



TRANSFER TO OTHER REGIONS

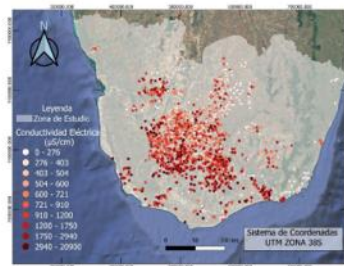
MADAGASCAR

Groundwater potential – Solar-based pumps



- 775 boreholes with estimated flow from pumping tests
- 17 Explanatory variables

Quality prediction – Electrical conductivity

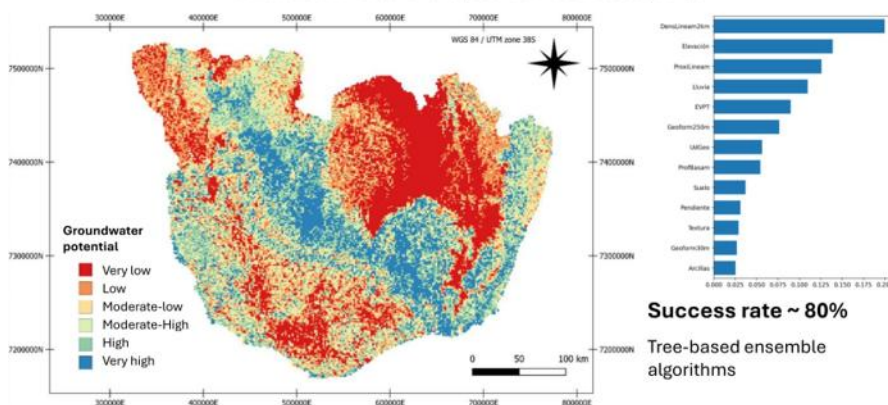


- 2314 boreholes with EC information
- 20 Explanatory variables

Climatic, Geological, Topographical, and Hydrogeological Conditions

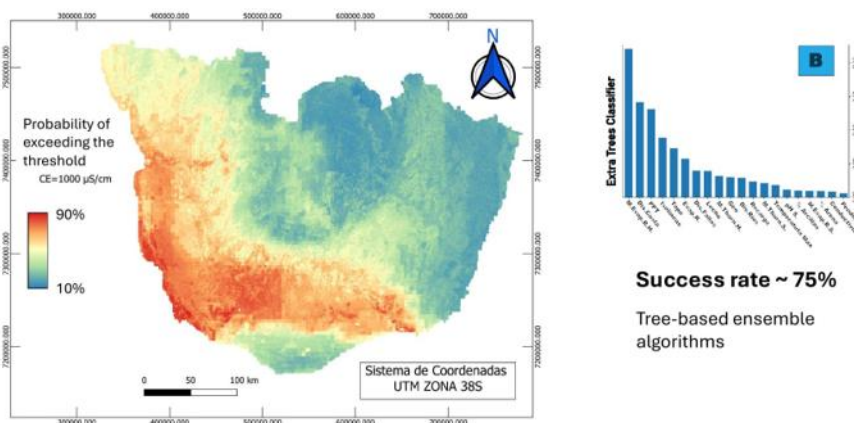
TRANSFER TO OTHER REGIONS

MADAGASCAR - Groundwater potential – Solar-based pumps

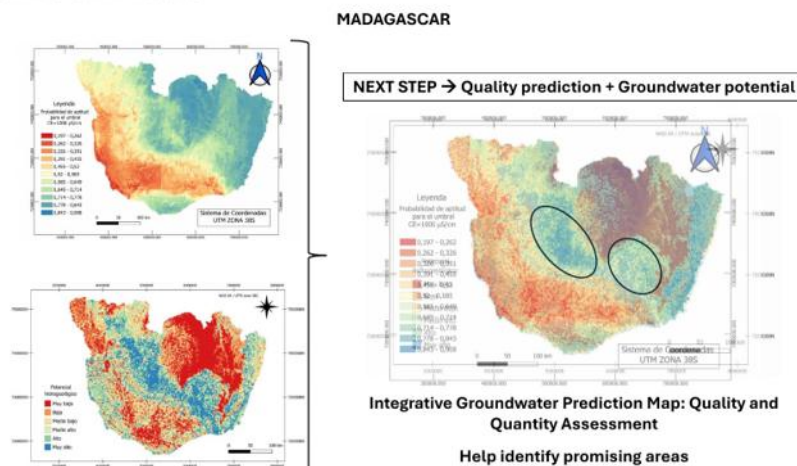


TRANSFER TO OTHER REGIONS

MADAGASCAR - Quality prediction – Electrical conductivity

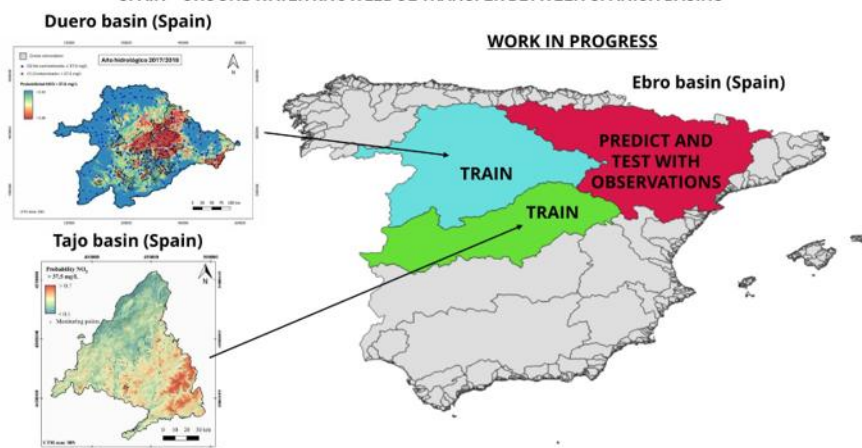


TRANSFER TO OTHER REGIONS



TRANSFER TO OTHER REGIONS

SPAIN – GROUNDWATER KNOWLEDGE TRANSFER BETWEEN SPANISH BASINS



CONCLUSIONS

- The application of ML models to **spatially predict hydrogeological variables** has demonstrated **high transferability** between different basins and regions.
- The **success of transferring these models** depends mainly on **the availability and quality of input data** in the target area.
- **The target variable and explanatory variables** must always be **adapted** to the context and specific objectives of the study.
- Regardless of the region or variables applied, **tree-based ensemble models** have been observed to **perform best**.