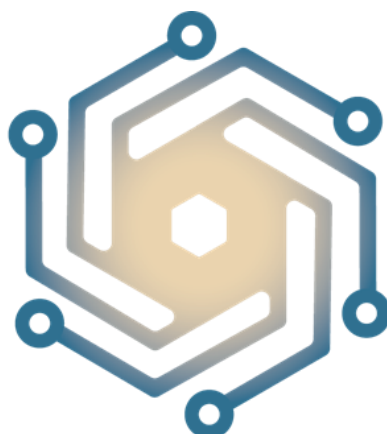




CLINT

CLIMATE INTELLIGENCE

AI-enhanced Climate Services for extreme impacts

Deliverable 6.3

April 2025



This project is part of the H2020 Programme supported by the European Union, having received funding from it under grant agreement No 101003876

Programme call:	Building a low-carbon, climate resilient future: climate action in support of the Paris Agreement (H2020-LC-CLA-2018-2019-2020)
Grant Agreement ID:	101003876
Project Title:	CLINT
Partners:	POLIMI (Project Coordinator), CMCC, HEREON, CSIC, SMHI, HKV, E3M, TCDF, DKRZ, IHE, ECMWF, UAH, JLU, OGC, UCM
Work-Package:	WP6: AI-enhanced climate services to foster adaptation at the pan-European scale
Deliverable #:	6.3
Deliverable Type:	Report
Contractual Date of Delivery:	30 April 2025 – extension included
Actual Date of Delivery:	30/04/2025
Title of Document:	AI-enhanced Climate Services for extreme impacts
Responsible partner:	JLU
Author(s):	Odysseas Vlachopoulos, Elena Xoplaki, Niklas Luther, Ilias Pechlivanidis, Yiheng Du, Kostas Kavvadias, Kristina Riepin
Content of this report:	The report integrates milestone MS37 and evaluates AI-enhanced CS for European EE impacts for the water, energy, and food sectors.
Availability:	This report is public.

Document revisions		
<i>Author</i>	<i>Revision content</i>	<i>Date</i>
Odysseas Vlachopoulos, Niklas Luther, Elena Xoplaki	Deliverable structure ready	08.09.2024
Yiheng Du, Ilias Pechlivanidis	Water section added	05.03.2025
Odysseas Vlachopoulos, Niklas Luther, Elena Xoplaki	Food section added	10.03.2025
Kristina Riepin, Kostas Kavvadias	Energy section added	12.03.2025
Odysseas Vlachopoulos, Elena Xoplaki, Niklas Luther, Ilias Pechlivanidis, Yiheng Du, Kostas Kavvadias, Kristina Riepin	All sections are filled in by the impact models and respective authors. Formatting of the document.	25.03.2025
Odysseas Vlachopoulos, Elena Xoplaki, Niklas Luther	Final formatting and submission for internal review.	31.03.2025
David Barriopedro (CSIC), Antonello Squintu (CMCC)	Internal review	14.04.2025
Elena Xoplaki, Odysseas Vlachopoulos, Yiheng Du, Kristina Riepin, Andrea Castelletti, Veronica Piuri	Revised deliverable available	22.04.2025

Disclaimer

Use of any knowledge, information or data contained in this document shall be at the user's sole risk. Neither the CLINT consortium nor any of its members, their officers, employees or agents shall be liable or responsible, in negligence or otherwise, for any loss, damage or expense whatever sustained by any person as a result of the use, in any manner or form, of any knowledge, information or data contained in this document, or due to any inaccuracy, omission or error therein contained.

The European Commission shall not in any way be liable or responsible for the use of any such knowledge, information or data, or of the consequences thereof.

This document does not represent the opinion of the European Union, and the European Union is not responsible for any use that might be made of it.

Copyright notice

© Copyright 2022 by the CLINT Consortium

This document contains information that is protected by copyright. All Rights Reserved. No part of this work covered by copyright hereon may be reproduced or used in any form or by any means without the permission of the copyright holders.

Table of Content

Table of Content	4
List of Figures	5
List of Tables	7
EXECUTIVE SUMMARY	8
LIST OF ACRONYMS.....	10
1. Introduction.....	12
1.1. Summary of previous work	12
1.2. Structure and objectives of this document	13
1.3. Connection to other work packages.....	14
2. Enhanced impact models description and operability.....	14
2.1. Water sector	14
2.2. Energy sector	17
2.3. Food sector – The surrogate crop growth model.....	19
3. Presentation of the AI-enhanced Climate Services for Extreme Impacts - Water sector 27	
3.1. Model evaluation	27
3.2. Enhanced simulation of hydrological signatures	28
3.3. Updated streamflow signature pattern	31
4. Presentation of the AI-enhanced Climate Services for Extreme Impacts - Energy sector 33	
4.1. AI-enhanced hydro inflows projections for the European energy sector	33
4.2. Long-term trends of hydropower generation variability in Europe	38
4.2.1 Hydropower variability and capacity factor assumptions for hydropower systems	38
5. Presentation of the AI-enhanced Climate Services for Extreme Impacts - Food sector 46	
5.1. Performance	46
5.2. Areas of Concern.....	51
6. Summary and conclusions.....	55
6.1. AI-enhanced climate services for the water sector	56
6.2. AI-enhanced climate services for the energy sector	56
6.3. AI-enhanced climate services for the food sector.....	56
Bibliography	57
[Chapter 2.1]	57
[Chapter 2.3]	58
[Chapter 4]	58
[Chapter 5]	59

List of Figures

Figure 1 - Surrogate model architecture.	22
Figure 2 - The histogram shows the frequency of values for the spearman correlation coefficient between the predictions of the test dataset and the actual ERA5-forced TWSO.	24
Figure 3 - Random samples (grid points) of crop growth timeline from DVS1 to DVS2 in terms of actual and predicted TWSO (kg ha-1).	25
Figure 4 - Performance distribution of AI-enhanced approaches at different gauged and ungauged conditions. For SMAE, the perfect performance is 0 on the x-axis, therefore the model with better performance will appear with higher curve position (e.g. closer to 0); while for NSE and logNSE, the perfect performance is 1 on the x-axis, the model with better performance will appear lower in the plot (e.g. closer to 1).	27
Figure 5 - Performance distribution of NSE of AI-enhanced approaches at different gauged and ungauged conditions in specific hydrological regimes.	28
Figure 6 - Spatial comparison of streamflow characteristics across Europe showing observations, E-HYPE, and AI-enhanced predictions (multi-basin approach), with runoff values in mm represented by the color scales.	29
Figure 7 - Scatter plots comparing observed versus predicted streamflow signatures (Mean, Q20, and Q90) for raw E-HYPE (left panels) and AI-enhanced (right panels) results. Results are shown for both training and testing periods, and for gauged and ungauged stations. Different colors represent distinct hydrological clusters (1-11). The dashed line indicates the 1:1 perfect prediction line.	31
Figure 8 - Spatial patterns of simulated extreme high streamflows (90th percentile; Q90) before and after the AI enhancement.	32
Figure 9 - Historical and projected inflows for Austria: models' ensemble and scenarios: RCP2.6, RCP4.5, RCP8.5 (Note: Graphs for the rest of hydro-dominant countries are available in the Appendix).	35
Figure 10 - Projected changes in annual ROR hydropower inflow variability (2020–2060): median vs. 95th percentile ensemble estimates.	36
Figure 11 - Projected changes in annual ROR hydropower inflow variability (2020–2060): median vs. 95th percentile ensemble estimates.	37
Figure 12 - Estimated ROR capacity factors: projections for 2050 for climate scenarios: RCP2.6, RCP4.5, RCP8.5.	40
Figure 13 – Power mix in 2030. Emissions neutrality scenario under the historical median weather years. Source: PRIMES-IEM.	42
Figure 14 - Power mix in 2050. Emissions neutrality scenario under the historical median weather years. Source: PRIMES-IEM.	42
Figure 15 – Comparing 3 climate scenarios for 2030: RCP2.6, RCP4.5, RCP8.5. Source: PRIMES-IEM.	43
Figure 16 – Relationship between hydropower generation and fossil run generation for 3 climate scenarios for 2030: RCP2.6, RCP4.5, RCP8.5 Source: PRIMES-IEM.	44
Figure 17 - Comparing climate scenarios for: RCP8.5 in 2030: median, lower 5% and upper 95%. Source: PRIMES-IEM.	45

Figure 18 – Comparing climate scenarios RCP4.5 in 2030 and 2050: median, lower 5% and upper 95%. Source: PRIMES-IEM.....	45
Figure 19 - Annual Pearson correlation plot for an ensemble of 25 SEAS5.1 reforecasts and ERA5-forced yields from 1993 to 2016.....	47
Figure 20 - Overlaid probability density functions (PDFs) of TWSO (kg/ha) simulations forced with ERA5 and the SEAS5.1 ensemble member 01 (SEAS5.1_ens01)	48
Figure 21 - Comparison of the ERA5 and SEAS5.1_ens01 TWSO simulations over the period 1993-2016: (a) Pearson correlation; (b) MAE; (c) RMSE; (d) Mean bias.	49
Figure 22 - Probability density functions (PDFs) of TWSO (kg/ha) AI-based surrogate simulations forced with ERA5 and (a) raw GFDL-ESM4 (r1i1p1f1) / (b) bias-adjusted GFDL-ESM4 (r1i1p1f1) for 1993-2014.	50
<i>Figure 23 - Comparison of the ERA5 and raw GFDL-ESM4 (r1i1p1f1) TWSO simulations over the period 1993-2014: (a) MAE; (b) RMSE; (c) Mean bias. Comparison of the ERA5 and bias adjusted GFDL-ESM4 (r1i1p1f1) TWSO simulations over the period 1993-2014: (d) MAE; (e) RMSE; (f) Mean bias.</i>	<i>51</i>
Figure 24 - Probabilistic Areas of Concern for each year of the 2017-2023 period based on 51 ensemble members of SEAS5.1 and the surrogate TWSO simulations.	53
Figure 25 - 10-yearly average AoC maps for GFDL-ESM4 under SSP5-8.5 based on the surrogate model; Left: raw projections; Right: bias adjusted projections.	55

List of Tables

Table 1 - The evaluation metrics used to quantify the model performance improvements assessing different characteristics of the streamflow time series. The interpretation gives common standards for model evaluation with the corresponding metric (Crochemore et al.).	16
Table 2 - Evaluation metrics for assessing climate-induced impacts on power generation and hydropower availability.	18
Table 3 - Evaluation metrics employed for the experiments of the food sector.	23
Table 4 - The CMIP6 ESMs that were used in the Food sector.	26
Table 5 - Run-of-river generation share and capacity in each country (2020). Source: ENTSO-E.	34
Table 6 - Change in potential hydro generation for RCP scenarios for selected countries for climate scenarios: RCP2.6, RCP4.5, RCP8.5.	38
Table 7 - CMIP6 models historical simulations used to simulate TWSO with the surrogate model for AoC generation, 1993-2014.	53

EXECUTIVE SUMMARY

Deliverable 6.3 presents the consolidated results on Artificial Intelligence and Machine Learning (AI/ML) enhanced climate services across Europe. The analysis covered historical, forecasted, and projected climate conditions, highlighting their implications for extreme impacts.

In the evolving field of climate research, the integration of AI/ML into climate services is revolutionizing our understanding and response strategies to climatic changes affecting the water, energy, and food sectors. This transformation depends significantly on data accessibility and integrity. This deliverable illustrates this through various methods and case studies, encompassing a wide range of environmental parameters such as temperature, river discharge, and precipitation.

This report expands from the preliminary findings (see Deliverable D6.2) and the selection of suitable AI/ML methodologies to address the climate data complexity and enhance its integration into modeling frameworks, to the developed cornerstones of the currently built impact models and climate services per sector.

Water Sector

We introduce here a Multi-basin Long-Short-Term-Memory (LSTM) framework that builds on the previously developed Single-basin approach (D6.2). Unlike the Single-basin model, which trains a separate LSTM for each station, the Multi-basin approach uses data from 2,072 monitoring stations to train a single model, enabling it to learn generalizable hydrological patterns across diverse catchments. The model integrates dynamic inputs from the pan-European E-HYPE hydrological model (simulated streamflow, temperature, and precipitation over a 3-day window) and static inputs (station-specific climatic, physiographic, hydrological, and anthropogenic attributes). To assess both temporal and spatial generalization, a dual cross-validation strategy was applied—splitting data by time and location and employing K-fold validation across stations. The results show that AI post-processing significantly reduces biases in streamflow predictions, especially for high flows, where the raw E-HYPE model typically underestimates runoff. The AI-enhanced model offers a more accurate and spatially detailed representation of streamflow, particularly in hydrologically complex areas like mountainous and coastal regions. Temporal and spatial validation confirms the robustness of these improvements. This approach not only boosts predictive accuracy but also enables reliable streamflow simulations in ungauged basins, enhancing the spatial resolution and reliability of hydrological assessments.

Energy Sector

This work investigates the role of climate variability in shaping the future of the European power system, using the PRIMES-IEM model integrated with hydrological data from the E-HYPE model under multiple climate scenarios (RCP2.6, RCP4.5, RCP8.5). Building on previous research, the analysis applies machine learning techniques to identify key river basins influencing national hydropower production and to assess projected changes in hydropower availability. This approach enables a detailed quantification of system-level responses as the energy system transitions toward a fully decarbonized power sector by 2050. The findings demonstrate that even in a net-zero system, hydropower remains a cornerstone of system flexibility, with a significant impact on storage requirements, cross-border electricity flows, and the deployment of backup generation capacity.

Importantly, the assessment captures both annual hydropower availability and the seasonal dynamics of inflow patterns, offering insights into the climate-informed design of the power system.

Food Sector

For the food sector we present the development, the methodology, the evaluation and the results for the AI surrogate model, emulating the crop growth model ECroPS, for various datasets from the Coupled Model Intercomparison Project Phase 6 (CMIP6) collections, the European Centre for Medium-Range Weather Forecasts (ECMWF) Seasonal forecast System 5.1 (SEAS5.1) and ECMWF ReAnalysis, version 5 (ERA5, Hersbach et al. 2020). Additionally, we determine Areas of Concern (AoC), showing the more vulnerable areas in the future using the projections of the CMIP6 models. We show that the newly developed AI surrogate model is robust in accurately emulating yields, while performing in a faster and more simplified manner, which is the backbone of operationalization, scaling and further development. In terms of assessing the historical representation of climate impacts on yield, we see that the ERA5 reanalysis comparison with SEAS5.1/CMIP6 forcings is successful in metrics and qualitative analyses. At the same time, we observe an objective overestimation, likely due to the lack of human interventions in terms of mitigation and prevention measures in the model.

LIST OF ACRONYMS

Abbreviations

AI	Artificial Intelligence
AoC	Areas of Concern
ANN	Artificial Neural Network
ATT_TWSO	Attainable Total Weight of Storage Organs
C3S	Copernicus Climate Change Service
CMIP6	Coupled Model Intercomparison Project Phase 6
CCS/U	Carbon Capture and Storage / Use
CNRM-CM6-1-HR	Centre National de Recherches Météorologiques - Climate Model version 6-1 – High Resolution
CSs	Climate Services
DKRZ	Deutsches Klimarechenzentrum (German Climate Computing Centre)
DOY	Day Of Year
DVS	Development Stage
EC-Earth3	European Community Earth System Model Third Generation
EC JRC	European Comission's Joint Research Center
ECMWF	European Centre for Medium-Range Weather Forecasts
ECroPS	Engine for Crop Parallelizable Simulations
EE	Extreme Events
ERA5	ECMWF Reanalysis, version 5
ETS	Emissions Trading System
ESM	Earth System Model
GFDL-ESM4	Geophysical Fluid Dynamics Laboratory - Earth System Model version 4
GLM	Generalised Linear Model
HadGEM3-GC31-MM	Hadley Centre Global Environment Model version 3 - Global Coupled 3.1 - Medium Mesh

HPC	High-Performance Computing
HydroGFD v3.2	Hydrological Global Forcing Data version 3.2
logNSE	Logarithmic Nash-Sutcliffe Efficiency
LSTM	Long-Short-Term-Memory
MAE	Mean Absolute Error
MBC	Multivariate Bias Correction
ML	Machine Learning
MPI-ESM1.2-HR	Max Planck Institute Earth System Model version 1.2 - High Resolution
MSE	Mean Squared Error
NorESM2-MM	Norwegian Earth System Model version 2 - Medium Mesh
NSE	Nash-Sutcliffe Efficiency
PCA	Principal Components Analysis
PRIMES – IEM	Price-Induced Market Equilibrium System Internal Energy Market (Model)
RCP	Representative Concentration Pathways
ReLU	Rectified Linear Unit
RFNBO	Renewable Fuels of Non-Biological Origin
RNN	Recurrent Neural Network
RMSE	Root Mean Square Error
RoR	Run of River
SEAS	Seasonal forecast System
SMAE	Scaled Mean Absolute Error
SSP	Shared Socioeconomic Pathways
TWSO	Total Weight of Storage Organs
XGBoost	eXtreme Gradient Boosting
WOFOST	WOrld FOod STudies
WP	Work Package

1. Introduction

1.1. Summary of previous work

The work presented in this document builds upon Deliverable 6.2 and highlights the transformative potential of Artificial Intelligence (AI)/Machine Learning (ML)-enhanced climate services (CSs). As climate change intensifies the frequency and severity of extreme events such as floods, droughts, and heatwaves, there is a growing need for accurate, actionable, and operational climate information. CSs play a vital role in translating complex climate data into tools that support adaptation, mitigation, and disaster risk management across sectors like water, energy, and food.

Modern CSs are designed to meet the needs of diverse user groups by providing tailored tools, ranging from forecasts and long-term projections to risk assessments, and economic analyses. These tools enable stakeholders to make informed decisions, to help increase resilience in the face of climate variability and change. However, significant challenges remain in scaling and operationalizing these services to meet sector-specific and regional demands. Issues such as data complexity and heterogeneous data sources, lack of standardization, and difficulties in translating scientific outputs into context-specific information continue to limit their effectiveness.

The integration of AI and ML into CSs offers promising solutions to these challenges. These technologies enhance the capacity to detect, predict, and attribute extreme events by improving the processing and interpretation of large, complex datasets. For example, ML methods can refine outputs from numerical weather prediction or impact models, making them more relevant for on-the-ground decision-making. AI-driven tools can also help bridge the gap between climate science and user needs by enabling more targeted and adaptive services.

Focusing on the water, energy, and food sectors, we highlight how climate data, processed through AI-enhanced methodologies, can be embedded into impact models tailored to each sector's requirements. These sectors demand different types of data and temporal resolutions: the water sector, for instance, depends on both long-term projections for planning and short-term forecasts for emergency response. Likewise, the energy and food sectors require customized climate information to optimize operations and ensure long-term sustainability.

Building on methodologies developed in earlier phases of the project and more specifically in Deliverable 6.2, this report illustrates AI-enhanced CSs at the pan-European scale. These innovations are designed to inform local case studies and help demonstrate the added value of AI in making CSs more usable, responsive, and impactful. By prioritizing operationalization and user-centered design, we aim to contribute to closing the gap between scientific research, practical applications, and climate-informed decision-making, fostering climate resilience across Europe.

The Water, Energy, and Food Sectors

The CLINT project targets three key sectors, namely water, energy, and food, each with distinct vulnerabilities to climate change and extreme weather events. For each sector, the project integrates climate data and advanced AI/ML methodologies into modeling frameworks to improve predictions and inform better management strategies.

Water Sector

In the water sector, the project employs the E-HYPE hydrological model to enhance predictions of runoff and water availability across Europe. The integration of AI/ML techniques, including statistical methods like Generalized Linear Models and machine learning methods such as Random Forest and Long Short-Term Memory networks, refines model outputs by addressing biases and improving predictions of extreme hydrological conditions. Post-processing techniques have significantly enhanced the accuracy of streamflow simulations, particularly in Central Europe. The Budyko framework further attributes runoff changes to climatic drivers like precipitation and evapotranspiration, offering insights into future water resource variability under different climate scenarios. These advancements provide decision-makers with robust tools for water resource planning and management, ensuring resilience to climate variability and extreme events.

Energy Sector

The energy sector relies on the PRIMES model to evaluate the impacts of climate-induced variations in temperature and river discharge on power demand and hydropower generation. AI-enhanced modeling links climate variability with energy production and consumption, highlighting the sector vulnerabilities to extreme realizations of future hydrologic conditions. Preliminary results demonstrate the potential for these tools to improve the resilience of energy systems by providing detailed insights into changes in hydropower energy availability. These findings are critical for developing adaptive strategies to ensure stable energy supplies and mitigate the impacts of extreme climate events on energy infrastructure.

Food Sector

For the food sector, the ECroPS crop growth model is utilized to analyze the effects of climate variability and climate extremes on agricultural yields. The model integrates detailed simulations of crop growth processes under varying environmental conditions, incorporating seasonal forecasts and high-resolution climate projections to assess vulnerabilities. The CS is built using AI-based surrogate modeling that aims to enhance the efficiency and scalability of such analyses, enabling resource-efficient and robust predictions of crop performance under different scenarios. Given the observed and projected increase of high-impact compound events like droughts and heatwaves, the project supports the development of adaptive agricultural practices to ensure food security and optimize productivity.

The transformative potential of AI and ML in advancing CSs forms the foundation of the work detailed in the upcoming chapters. These technologies have enhanced the ability to detect, predict, and attribute extreme events, offering stakeholders insights that are both precise and sector specific. The integration of advanced modeling techniques across water, energy, and food sectors demonstrates how AI/ML can address the complexities of climate variability, enabling a deeper understanding of its cascading impacts and supporting the development of resilient strategies.

1.2. Structure and objectives of this document

The overarching objective of this report is to describe and present the methodologies, data and results that are used in the three sector-oriented CSs for the European domain.

Each chapter explores how these methodologies have been applied to refine existing models and operationalize them for real-world use. The focus is on addressing specific challenges in each sector,

enhancing runoff simulations and hydrological predictions in the water sector, linking climate variability with energy demand and production in the energy sector, and assessing agricultural vulnerabilities to extreme weather in the food sector. These detailed analyses underscore the importance of tailoring CSs to meet diverse user needs, incorporating high-quality data, and fostering collaboration across sectors.

The chapters also illustrate the ways in which this integration of AI/ML lays a foundation for future advancements. From improving data integration and accessibility to scaling models for broader applications, the innovations highlighted set a benchmark for operationalizing CSs. By addressing sector-specific challenges and enabling adaptive strategies, these efforts establish CSs as indispensable tools for mitigating the impacts of accelerating climate change and building resilience across interconnected systems.

1.3. Connection to other work packages

This deliverable is complementary to the findings and methodologies described in other work packages (WPs), particularly WP2 and WP3, in some cases refining the AI/ML methods developed therein. Additionally, the datasets and web service requirements related to the CSs developed and utilized are managed through WP8. Finally, the methodologies presented in this deliverable at pan-European scale can be streamlined to activities under WP7 for local case studies. An example of a link that can be identified in terms of the CS developed for the food sector, the crop growth surrogate model, is the application at the greater Como lake hotspot area in Italy, where modelling of crops without irrigation, which is currently accommodated using water from the lake basin, could be examined in terms of future risks such as persisting droughts under climate change scenarios.

2. Enhanced impact models description and operability

2.1. Water sector

2.1.1. E-HYPE hydrological model

Here we only briefly describe the hydrological model used in the pan-European investigation. More details about the hydrological model setup can be found in Deliverable D6.2 “Preliminary report on AI-enhanced Climate Services for extreme impacts”.

The E-HYPE model v.3.1.3 covering the pan-European domain (8.8 million km²) was used to generate streamflow simulations. E-HYPE is a semi-distributed process-based model operating at a fine spatial resolution with 35 408 sub-basins and an average spatial resolution of 215 km² (Hundecha et al., 2016). The model itself has been used in various investigations involving seasonal predictions and assessment of the impacts of climate change on water resource management.

The model has been forced with meteorological input from the Hydrological Global Forcing Data version 3.2 (HydroGFD v3.2) product. HydroGFD is an observation-corrected reanalysis dataset providing historical meteorological information of precipitation and mean temperature at a 0.5° gridded resolution (Berg et al., 2021). HydroGFD was used to force the E-HYPE hydrological model to generate the reference simulation (here considered as 'observations'; pseudo-observations) across all 35 408 sub-basins.

2.1.2. AI-Based Streamflow Enhancement in Ungauged Conditions

Artificial intelligence (AI) has demonstrated strong potential for improving hydrological simulations by enhancing process-based models. In this study, we developed a regionalized streamflow post-processing framework using Long Short-Term Memory (LSTM) networks to correct residual errors in the process-based E-HYPE model.

LSTM networks, a specialized form of recurrent neural networks, are particularly well-suited for sequential data analysis (Hochreiter & Schmidhuber, 1997). Their ability to capture complex temporal dependencies and retain information over extended sequences makes them highly effective for hydrological modeling. LSTMs have been widely applied in hydrology due to their capacity to model the non-linear and non-stationary characteristics of hydrological time series (Kratzert et al., 2018; Lees et al., 2022; Li et al., 2021). By leveraging information from multiple time steps, LSTMs can effectively represent the intricate interactions between hydrological variables and their temporal evolution.

This study implements a Multi-basin LSTM framework, building upon the Single-basin approach previously introduced in D6.2. While the Single-basin trains distinct LSTMs for each station individually (Du & Pechlivanidis, 2024), the multi-basin approach trains a single LSTM using data from all stations simultaneously. This enables the model to learn common hydrological patterns across multiple catchments, improving its ability to generalize beyond individual locations.

The dataset includes observations from 2,072 monitoring stations, incorporating both dynamic and static variables. The dynamic inputs consist of simulated outputs from the E-HYPE model, specifically streamflow, temperature, and precipitation time series, with a look-back period of three days (i.e., the current day, as well as one and two days prior). Static inputs include environmental characteristics of each station, such as climatic indicators, physiographic attributes, hydrological regime descriptors, and measures of human impact on water systems.

To ensure a robust evaluation of the model's performance, a dual cross-validation strategy is applied across both temporal and spatial dimensions. Temporal validation involved an 80-20 split of the dataset, where 80% of the data are used for training and 20% for testing, assessing the model's generalization across different time periods. Spatial validation follows a similar 80-20 split at the station level, with K-fold cross-validation employed to evaluate the model's ability to generalize across different geographical regions. This comprehensive validation strategy ensures that the model is not only temporally robust but also transferable across diverse hydrological settings.

By integrating both dynamic and static features, the LSTM model effectively captures the interplay between temporal variations and spatial heterogeneity in hydrological processes. The inclusion of K-fold cross-validation specifically addresses the challenge of spatial transferability, ensuring that the model remains applicable across a wide range of hydrological conditions (Willmott & Matsuura, 2005).

2.1.3. Model Evaluation

Evaluation metrics

To evaluate the added value from post-processing, three evaluation metrics are used to assess the potential improvements regarding errors in total volume, high and low streamflow extremes (*Table 1*), as represented by the Mean Absolute Error (MAE; Willmott & Matsuura, 2005), Nash-Sutcliffe Efficiency (NSE; Nash & Sutcliffe, 1970) and its logarithmic form (logNSE; Lamontagne et al., 2020), respectively. In particular, the Scaled Mean Absolute Error (SMAE) is applied to adjust MAE in relation to the average streamflow observed at each station, thus allowing the comparison of MAE values across stations that have varying streamflow magnitudes.

The performance of the multi-basin LSTM is evaluated through comparison with both the raw E-HYPE simulations and the individually trained Single-basin LSTMs. This comparative assessment helps quantify the potential benefits of the regionalized approach against both the original hydrological model outputs and station-specific machine learning models.

Table 1 - The evaluation metrics used to quantify the model performance improvements assessing different characteristics of the streamflow time series. The interpretation gives common standards for model evaluation with the corresponding metric (Crochemore et al.).

Characteristic of the streamflow signal	Evaluation metric (Abbreviation)	Equation	Interpretation
Total Volume	Mean Absolute Error (MAE); Scaled Mean Absolute Error (SMAE)	$MAE = \frac{\sum_{t=1}^T y_o^t - y_m^t }{T}$ $SMAE = \frac{MAE}{\bar{y}_o}$	Focusing on the volumetric biases Very good [0 – 0.3] Good [0.3 - 0.4] Fair [0.4 - 0.5] Poor [0.5 - 0.6] Very poor [0.6 - 0.7] Unsatisfactory [0.7 - Inf]
High streamflow extreme	Nash-Sutcliffe Efficiency (NSE)	$1 - \frac{\sum_{t=1}^T (y_o^t - y_m^t)^2}{\sum_{t=1}^T (y_o^t - \bar{y}_o)^2}$	Focusing on high extremes Very good [0.7 – 1] Good [0.5 - 0.7] Fair [0.2 - 0.5] Poor [0 - 0.2] Very poor [-0.5 - 0] Unsatisfactory [-Inf - -0.5]
Low streamflow extreme	Logarithmic Nash-Sutcliffe Efficiency (logNSE)	$1 - \frac{\sum_{t=1}^T (\log(y_o^t) - \log(y_m^t))^2}{\sum_{t=1}^T (\log(y_o^t) - \log(\bar{y}_o))^2}$	Focusing on low extremes Very good [0.7 – 1] Good [0.5 - 0.7] Fair [0.2 - 0.5] Poor [0 - 0.2] Very poor [-0.5 - 0] Unsatisfactory [-Inf - -0.5]
y_o^t and y_m^t denotes the observation and model simulation at each timestep t , respectively, where t ranges from 1 to T .			

Evaluation of Hydrological Signatures

The evaluation of hydrological extremes focused on three key flow statistics (Du et al., 2024) to characterize the range of flow conditions across the study area:

The 20th percentile flow (Q20) represents low flow conditions, providing insight into periods of reduced water availability and potential drought conditions. This quantile is particularly relevant for

understanding water resource management during dry periods and assessing environmental flow requirements.

The 90th percentile flow (Q90) captures high flow conditions, offering information about potential flood events and periods of abundant water resources. This value is crucial for flood risk assessment and infrastructure planning.

The mean flow provides a central tendency measure that characterizes the typical flow conditions in each catchment. This index serves as a baseline for understanding the overall water availability and helps contextualize the extreme flow conditions represented by Q20 and Q90.

Together, these three statistics provide a comprehensive characterization of the flow regime, enabling the assessment of both water scarcity and abundance across the pan-European region. This analysis framework allows for the identification of areas prone to hydrological extremes and supports informed decision-making in water resource management.

2.1.4. Model Application in the Ungauged Conditions

Following successful training and validation, the LSTM model was deployed across all catchments in the pan-European region to update the hydrological signature patterns. This comprehensive application enabled a systematic assessment of hydrological characteristics across diverse geographical and climatic conditions throughout Europe.

2.2. Energy sector

2.2.1. PRIMES-IEM energy model

This deliverable briefly introduces the PRIMES-IEM energy model, which is used for EU-level energy system analysis. A detailed description of the model setup can be found in Deliverable D6.2 “Preliminary report on AI-enhanced Climate Services for extreme impacts”.

The PRIMES model is a large-scale applied energy system model designed to project long-term energy system evolution and restructuring on both the supply and demand sides. It operates at both country and EU levels, integrating sector-specific modules for power generation, industry, transportation, residential, and services. The model incorporates behavioral dynamics and discrete choice theory to capture decision-making processes across different energy sectors (E3-Modeling, 2018).

The model ensures continuity between historical Eurostat energy data and future projections, covering the EU 27 countries. It operates in five-year increments with projections extending to 2100.

The PRIMES-IEM variant, used in this study, offers a flexible time-step framework, supporting up to 8760 hourly time steps per year. It provides a detailed representation of European power and heat generation, integrating supply, demand, and energy exchanges within a stylized network. The model simulates long-term electricity and heat demand across key sectors—industry, residential, services, and transportation—and assesses changes in the power mix under different climate and energy policies, economic, and technological conditions.

PRIMES-IEM includes an advanced representation of power generation technologies, including future energy solutions such as carbon capture, storage, and utilization (CCS/U), synthetic fuels, and renewable fuels of non-biological origin (RFNBO). It inherits the core characteristics of PRIMES, ensuring compatibility with EU policy instruments and climate targets, making it a valuable tool for analyzing the impact of climate change on the European electricity market.

2.2.2 AI-Based hydro generation Profiles

Run-of-river (ROR) hydropower and reservoirs are weather-dependent energy sources that rely on river discharge to generate electricity. Understanding the relationship between inflow and generation from ROR and lakes is crucial for assessing future hydropower potential, particularly in the context of climate change. ML techniques have increasingly been applied to predict weather-dependent electricity generation, offering a non-parametric approach to model the complex relationship between river discharge and hydro generation.

Building on previous work, this study applies the datasets developed with the ML method (XGBoost model) to analyze climate driven effects on hydropower generation at a European scale. The approach expands upon the methodology introduced in Deliverable D6.2, using ML to analyze hydrological data and project ROR and reservoirs' generation patterns for individual countries. The generated time series are then integrated into the PRIMES-IEM energy model to assess future energy under different projections of climate.

The dataset used in this study consists of daily river discharge estimates at the sub-basin level across Europe, derived from the E-HYPE (Europe Hydrological Predictions for the Environment) model. Future projections of hydro generation are available for the period 2025–2100, considering three Representative Concentration Pathways (RCPs; RCP2.6, RCP4.5, and RCP8.5). The uncertainty in these projections is assessed using the EURO-CORDEX climate model ensemble. To explore extreme hydro generation patterns, the median and the 95th percentile range of the ensemble are analyzed, ensuring consistency of generation patterns on a daily basis.

In this deliverable, we present the results of the ML-based generation projections and their integration into PRIMES-IEM. The analysis provides insights into the potential evolution of hydropower across different climate scenarios, supporting energy system planning and policy development at the European level.

2.2.3 Evaluation metrics for the scenarios in the Energy Sector

Table 2 - Evaluation metrics for assessing climate-induced impacts on power generation and hydropower availability.

Description of the variable	Evaluation metric	[unit]	Interpretation
Annual hydropower generation	Annual sum of hydropower generation per country, change in power generation by plant type	[GWh / year]	Evaluates the dependence of power generation on hydropower
Fossil generation	Annual sum of fossil - fueled power generation per country, change in power generation by plant type	[GWh / year]	Evaluates the dependence of power generation on fossil fuels
System costs	Annual total system costs for power generation (investment, energy purchase, network costs)	[EUR/year]	Balancing, flexibility and fuel costs included in the metric
Emissions	Change in annual emissions from power generation sector (5-year steps)	[MtCO ₂]	Emissions from the power generation sector

2.3. Food sector – The surrogate crop growth model

2.3.1. ECroPS crop growth model

Crop growth models are fundamental tools for understanding how climate and management practices influence agricultural productivity. Models like ECroPS (<https://github.com/ec-jrc/ecrops/>, Toreti et al. 2022) and WOFOST (de Wit et al. 2019; Ceglar et al. 2019) integrate a variety of biological, physiological, and environmental processes to simulate crop development over time. They offer valuable insights into the effects of temperature, precipitation, soil moisture, and nutrient availability on crop performance, which helps farmers, policymakers, and agribusinesses make informed decisions and risk assessments.

ECroPS, which is built on the WOFOST framework, simulates essential physiological processes such as photosynthesis, respiration, and transpiration. It monitors the entire crop life cycle through thermal time and tracks key growth stages. The model consists of interconnected modules that integrate water and nutrient balance calculations, and predict biomass and yield, providing time series of crop development under various conditions.

Considering climate change, ECroPS can be critical in assessing agricultural vulnerability and supporting adaptation/mitigation strategies. It simulates various production scenarios, including potential, water-limited, and nutrient-limited conditions, and accounts for spatial variability across the greater European domain. This enables the prediction of crop responses to extreme weather events and aids stakeholders in developing adaptive measures. For further details refer to deliverable D6.2.

The overall pipeline of the surrogate modeling commences from the deployment and running of the crop growth model itself. The ECroPS version 1.5.0 is installed and set in the High-Performance Computing (HPC) environment at the Deutsches Klimarechenzentrum (DKRZ) in order to utilize the hardware resources of the HPC and the parallelization/distribution capacities built both within the HPC and ECroPS itself. This system ensures scalability and the automated repetition of the experiments in a parallel workflow.

ECroPS integrates multiple datasets to assess crop performance under varying climatic conditions (see Deliverable D6.2 for details and schematics). The process workflow involves the acquisition of essential input data, including year, crop type, location (longitude and latitude), sowing date, weather variables, soil characteristics (i.e., field capacity, wilting point, saturation concentrations and soil maximum thickness for root depth), and crop-specific growth parameters. The static soil and crop-specific gridded datasets required to run ECroPS at a spatial resolution of 0.25° are provided by EC JRC.

In order to illustrate the added value of the AI surrogate model we focus on grain maize. Grain maize is a crucial crop in the EU, primarily used for animal feed, biofuel production, and industrial applications. In 2023, the EU produced approximately 61.0 million tons of grain maize and corn-cob-mix, marking a 15.2% increase from the drought-hit 2022 season, despite a 6.1% decrease in harvested area to 8.3 million hectares (Eurostat, retrieved 15/04/2025). ECroPS simulates maize growth dynamics across the timeline, providing biomass accumulation, soil moisture variations, and the impact of climatic factors on crop development. The weather data stem from the ERA5 reanalysis, consisting of daily minimum and maximum temperature, precipitation, radiation, wind speed, and relative humidity, and are preprocessed to extract the relevant spatial and temporal subsets for

simulation (per latitude and longitude and timestep). More specifically, the simulation covers the 1993 to 2023 period in order to align with the reforecast period of the seasonal forecasts (see Chapter 5). The spatial bounding box for grain maize within the greater European domain, as defined by the EC JRC for grain maize production, includes longitudes from 17.19° W to 48.99° E and latitudes from 32.76° N to 56.01° N (see also Deliverable D6.2).

ECroPS simulates each grid cell individually, determining the crop growth based on a sowing date. The model is executed in parallel for each grid cell, spanning the entire growing season, simulating crop growth, yield potential, and water balance dynamics. The model simulates phenology, biomass accumulation, and soil moisture balance, incorporating predefined workflows. Finally, following the simulations for attainable production, a post-processing pipeline is employed to refine the output. The simulation results are transformed into standardized NetCDF files, aligning geographic coordinates with reference datasets.

2.3.2. AI-based surrogate modelling

ECroPS, like most complex crop growth models, is highly demanding in computational resources and the calculations are quite time consuming. As a result, the democratization and wider accessibility of such models is currently quite difficult both in terms of deployment complexity and in resources. This is one major aspect that is currently addressed, with respect to the service-oriented goals in CLINT.

We have successfully developed an AI-based surrogate model to complement traditional mechanistic approaches such as ECroPS and WOFOST (see also D6.2). The newly developed surrogate model harnesses the deep learning capacities to directly capture the nonlinear and dynamic relationships between environmental variables and crop responses, overcoming the complexity and high computational demands that limit the broader use of conventional crop models.

The surrogate model leverages advanced computational techniques, including parallel GPU processing and distributed computing, to quickly analyze the input datasets, train the model and then swiftly predict. This capacity enables real-time decision support and risk assessment in agricultural management. Our surrogate model replicates the behavior of resource-intensive simulations while drastically reducing computational costs. This advancement allows the easy integration of crop growth modeling into operational decision support systems, allowing stakeholders to make better-informed decisions and effectively adapt to environmental uncertainties.

The surrogate model is based on Recurrent Neural Networks (RNNs), a class of neural networks specifically designed for processing sequential data. Unlike traditional feedforward networks, RNNs have internal feedback loops that retain and utilize information from previous time steps, making them particularly well-suited for tasks such as time series forecasting, language modeling, and the simulation of complex dynamical systems. This ability to capture temporal dependencies has established RNNs as indispensable tools across many fields where dynamic modeling is essential.

Building on the basic RNN architecture, LSTM networks were developed to overcome limitations of standard RNNs (such as the vanishing gradient problem) by employing gating mechanisms to regulate the flow of information. An evolution of this concept is the nested LSTM, in which LSTM cells are embedded within other LSTM cells to create a hierarchical memory structure. This approach enables the network to capture multiple layers of temporal abstraction, modeling both fine-grained details and broader trends in complex data sequences. The idea of incorporating hierarchical or nested

memory structures is exemplified by the work of Chung et al. (2017), who introduced hierarchical multiscale recurrent neural networks, laying the groundwork for these advanced architectures.

In terms of training (90%) and testing (10%) of the surrogate model, the feature space consists of a number of features both raw and engineered. Each sample from the datasets corresponds to a specific location and contains daily weather observations as independent variables across multiple years with the yield simulation from ECroPS being the dependent variable. The model uses three meteorological variables as its primary features: Temperature (maximum & minimum), total precipitation and engineered features including temporally lagged weather variables and day of the year (DOY).

A key aspect of this approach is the use of lagged features, which provide historical weather context, allowing the model to learn from past trends and improve prediction accuracy. The generation of lagged features provides substantial inferential torque relying mainly on three major rationales: (i) weather persistence, (ii) delayed crop responses to the cumulative impact of precipitation and temperature on crops, and (iii) improved pattern recognition capacity learning the effects of recurrent weather patterns, such as a dry spell followed by rainfall. For each day in the dataset, past weather conditions from the previous 1 to 5 days are included, which is expected to enhance the predictive power by allowing the model to link past weather fluctuations with future trends.

Overall, the feature space consists of 19 features: the 3 weather input variables, 5 lag features per weather variable and the DOY. The output dependent variable is the grain maize yield in terms of total weight of storage organs (TWSO).

2.3.3. The Architecture

Time-series AI models like RNNs require structured input sequences to capture meaningful temporal representations. In our model, fixed-length 6-day batches are created to infer short-term variability and long-term dependencies across multiple batches. This batching methodology is based on the rationale that each sample contains a meaningful time window, allowing the model to identify local fluctuations in the independent variables from the input feature space while maintaining the ability to track broader trends over time. By processing non-overlapping sequences, the model avoids redundant information while ensuring that patterns within each batch are independent, leading to more accurate learning. The 6-day window reflects a balance between preserving memory efficiency and providing the network with enough contextual information to infer relevant weather-driven dynamics, specifically for crop development.

This batching process is critical for the nested RNN architecture, as it enables the inner LSTMs to focus on intrinsic patterns within each batch, while the outer LSTMs look at the bigger picture, connecting sequences over time. This time-series partitioning strategy allows the model to efficiently process sequential dependencies across time scales and enhances the model's ability to recognize the underlying patterns.

The model is built using a hierarchical recurrent structure designed to process time-dependent data. The overall architecture and the components are shown in *Figure 1*. The surrogate model consists of two levels of LSTM networks: an inner TimeDistributed LSTM responsible for capturing short-term patterns within each 6-day batch and an outer LSTM that processes the sequence of batch outputs to capture long-term dependencies. The inner LSTMs focus on short-term variability, condensing patterns within each batch into a summarized representation, while the outer LSTMs track transitions

and relationships between consecutive time windows, enabling the model to identify temporally varying effects of weather on crop development, ultimately providing time-series of yield prediction.

The final fully connected layer, which is a TimeDistributed-wrapped dense layer, maintains the sequential nature of the predictions, while a Rectified Linear Unit (ReLU) activation function guarantees non-negative predictions, aligning with the needs of the target variable. Finally, the use of dropout layers confines potential overfitting.

The surrogate model was trained using the Adam (Kingma and Ba 2014) optimizer. All LSTM layers consist of 128 units, the dropout layer has a setting of 0.3. Additionally, a Huber loss function (Huber 1964) is employed, ensuring a more stable training than e.g. regular mean absolute error (MAE), using a loss delta of 0.5.

All aspects of the modelling are implemented in Python ≥ 3.7 , using the Tensorflow Keras framework (Chollet 2015; Abadi et al. 2016).

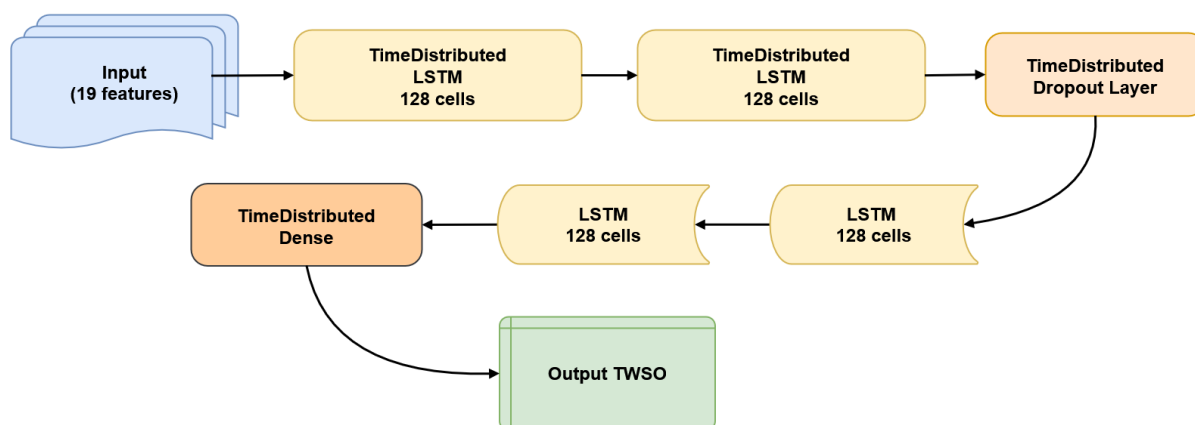


Figure 1 - Surrogate model architecture.

2.3.4. Model evaluation

The AI surrogate model is trained on ERA5-forced yearly crop yields from ECroPS for the years 1993 to 2023. More specifically, the crop yield target variable is the Attainable Total Weight of Storage Organs (ATT_TWSO), which quantifies water-limited scenarios without human intervention for harvestable yield. The temporal resolution is daily while on the spatial domain the model is agnostic, and like ECroPS it performs per grid cell. ATT_TWSO is measured in kilograms per hectare (kg/ha), representing the total weight of storage organs such as grains per hectare.

The model evaluation lies in a two stage process: Firstly, the actual training and testing process, using the ECroPS ATT_TWSO output (using ERA5 and the static EC JRC gridded datasets) to train and test the prediction pipeline. Secondly, once the AI surrogate is trained, and since ECroPS cannot be utilized with SEAS5.1 or CMIP6 weather data due to the different resolution of the input data (soil characteristics, crop-specific data and agromanagement data), the evaluation is performed between the ERA5-forced benchmark AI model output and the target AI surrogate yield (using SEAS5.1 or CMIP6).

A series of metrics are used to provide visual and numerical aspects:

- (i) Pearson correlation chart: yearly correlation values (one value per year) between the ERA5-based TWSO and SEAS5.1-forced TWSO simulations;
- (ii) correlation map: correlation per grid cell between two datasets for all the years;
- (iii) probability density functions (PDFs) overlaying: visual similarity metric which indicates whether the distributions share similar central tendencies, spreads, and overall shapes;
- (iv) MAE, RMSE and mean bias maps calculating the agreement between the two TWSO datasets over all the years and for each grid cell;
- (vii) yearly MAE, RMSE and mean bias for each cell and year.

A summary of the performance metrics employed for the assessment of the RNN prediction (y_m), with respect to a target variable (y_o) is presented in *Table 2* and *Table 3*.

Table 3 - Evaluation metrics employed for the experiments of the food sector.

Metric	Description	[unit]	Equation	Interpretation
Root Mean Square error (RMSE)	Average deviation based on the squared difference between predictions and the target variable. Penalizes larger errors more strongly, making it sensitive to outliers.	kg/ha	$\frac{1}{T} \sum_{t=1}^T (y_o^t - y_m^t)^2$	Lower RMSE indicates better prediction.
Mean absolute error (MAE)	Average deviation based on the absolute difference between predictions and the target variable. More robust to outliers due to the use of absolute differences.	kg/ha	$\frac{1}{T} \sum_{t=1}^T y_o^t - y_m^t $	Lower MAE indicates better prediction.
Pearson correlation	Measures linear dependencies between two samples.	Unitless	$\frac{\sum_{t=1}^T (y_o^t - \overline{y_o})(y_m^t - \overline{y_m})}{\sqrt{\sum_{t=1}^T (y_o^t - \overline{y_o})^2 \sum_{t=1}^T (y_m^t - \overline{y_m})^2}}$	Values close to 1 indicate strong positive linear relationship, and values close to -1 strong negative linear relationship.
Spearman correlation	Rank-based correlation coefficient measuring monotonic relationships.	Unitless	Same as Pearson, but with ranked data $r(x_i)$ and $r(y_i)$	Values close to 1 indicate a strong monotonic association of the ranked variables, and values close to -1 indicate a strong negative monotonic

				association of the ranked variables.
Mean Bias	Average raw deviation between predictions and the target variable. Includes the sign of the deviation and can be positive or negative	kg/ha	$\frac{1}{T} \sum_{o=1}^T y_o^t - y_m^{t2}$	Values close to zero indicate low bias.

The training process includes only data from cells that actually produce yields after they reach Development Stage (DVS) 1 to DVS2. The Development Stage is a key phenological indicator in crop growth modeling, representing the progression of a crop through its life cycle from sowing to maturity. In ECroPS, DVS1 marks the transition from vegetative growth to the reproductive phase, when the crop begins flowering and shifts energy from leaf and root development to grain production. As the plant continues to develop, it moves toward DVS2, the stage of physiological maturity. At this point, grain filling is complete, and the crop is ready for harvesting.

The overall results of the comparison of the test dataset (excluded from the training phase) with the corresponding ground truth from ERA5-driven ECroPS simulations, are shown in *Figure 2* and *Figure 3*. Both figures show that the model performance is very good, with the emulator capturing accurately both the evolution of the crop and the actual time of harvest reaching DVS2.

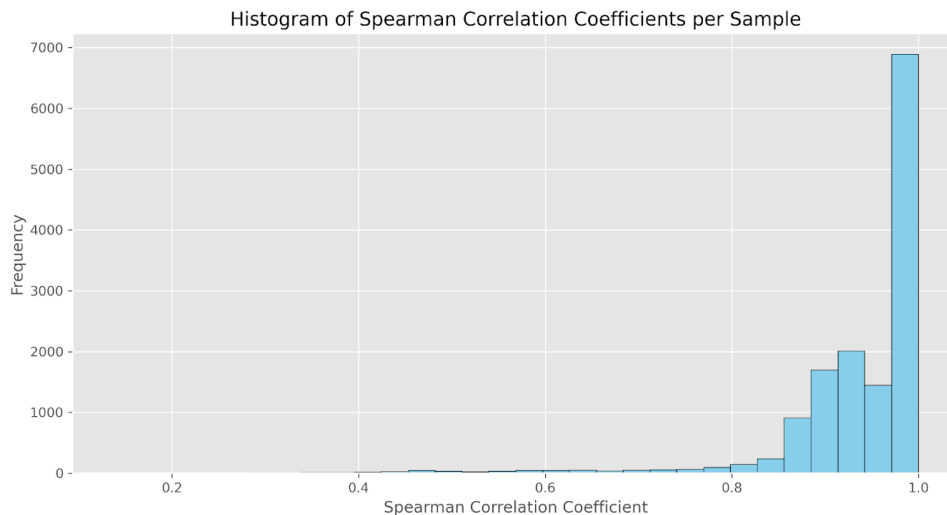


Figure 2 - The histogram shows the frequency of values for the spearman correlation coefficient between the predictions of the test dataset and the actual ERA5-forced TWSO.

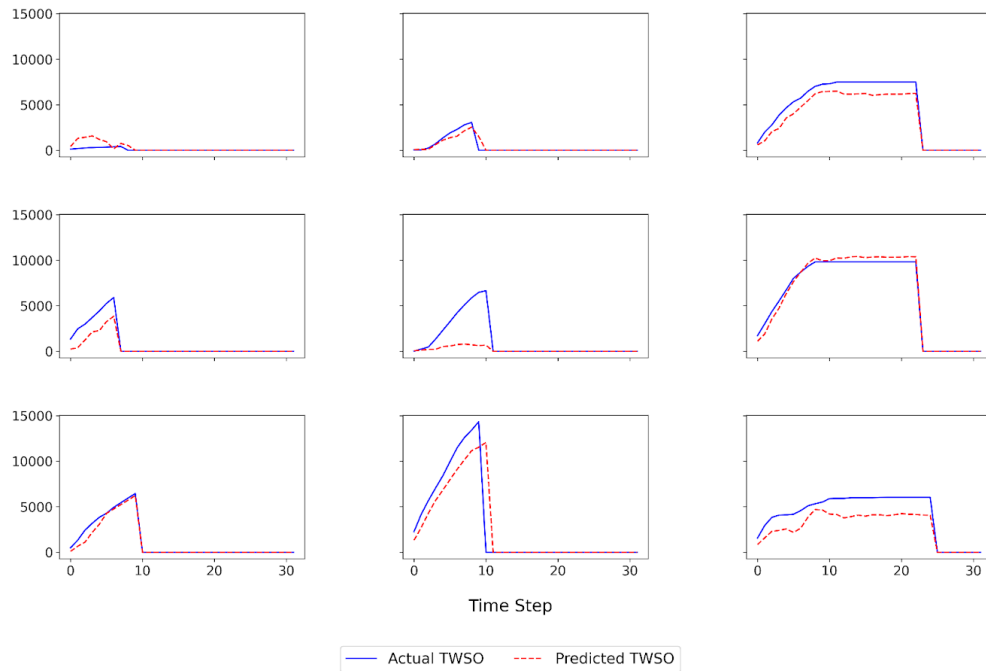


Figure 3 - Random samples (grid points) of crop growth timeline from DVS1 to DVS2 in terms of actual and predicted TWSO (kg ha⁻¹).

2.3.5. Model deployment

The surrogate model exhibits great computational performance: the prediction time for a single year for one grid cell amounts to a total of approximately 0.008 seconds CPU time in DKRZ's Levante HPC. Comparatively, the ECroPS CPU time for one year and one grid cell is approximately 70 seconds, 4 orders of magnitude slower than the surrogate model. Due to its efficiency, the model composes an efficient service to digest climate datasets and thus provide yield predictions in an operation mode.

The surrogate model was deployed for ERA5, six CMIP6 Earth System Models (ESMs) (*Table 4*) and the ECMWF SEAS5.1 seasonal forecasting system (Johnson et al., 2019). The CMIP6 ESMs were bias adjusted using ERA5 reference data (1940-2014) and the quantile delta mapping methodology (Cannon et al. 2015) using the MBC R-package. In terms of spatial extent, the EC JRC data for maize cultivation in the European domain is bound by a box with the following edge coordinates: Latitude: 32.7° N to 56.0° N; Longitude: 17.2° W to 49.0° E.

Table 4 - The CMIP6 ESMs that were used in the Food sector.

Model name	Historical realizations	Approximate nominal resolution in degrees (latitude, longitude)	Reference
CNRM-CM6-1-HR	'r1i1p1f2'	0.5°, 0.5°	(Voldoire et al. 2019)
EC-Earth3	'r11i1p1f1', 'r13i1p1f1', 'r15i1p1f1', 'r1i1p1f1', 'r3i1p1f1', 'r4i1p1f1', 'r6i1p1f1', 'r9i1p1f1'	0.7°, 0.7°	(Döscher et al. 2022)
GFDL-ESM4	'r1i1p1f1'	1°, 1.25°	(Dunne et al. 2020)
HadGEM3-GC31-MM	'r1i1p1f3'	0.56°, 0.8°	(Ridley et al. 2019)
MPI-ESM1-2-HR	'r10i1p1f1', 'r1i1p1f1', 'r2i1p1f1', 'r3i1p1f1', 'r4i1p1f1', 'r5i1p1f1', 'r6i1p1f1', 'r7i1p1f1', 'r8i1p1f1', 'r9i1p1f1'	1°, 1°	(Müller et al. 2018)
NorESM2-MM	'r1i1p1f1'	1°, 1.25°	(Seland et al. 2020)

2.3.6. Areas of Concern

The AI surrogate model TWSO output is additionally used to determine AoC. AoC for agricultural production are regions within the European domain that are prone to substantial impact and require targeted attention. EC JRC applies this term to areas facing challenges such as yield reductions and crop losses, water deficits and droughts (Ridoutt et al. 2016; Seguini et al. 2019). By marking these areas, AoC can be viewed as service-oriented forecasts that help focus research, preparedness and guide mitigation actions.

To figure out which areas need attention, an EC JRC council of experts considers multiple sources of information such as ground observations, satellite imagery and weather data among others, fusing them in a comprehensive way. The EC JRC creates scientific reports with AoC, transforming this information into bulletin updates for decision-makers, interested stakeholders as well as the public.

In CLINT, we introduce the AoC concept to build a similar AI-based product with the potential of extending the product to various lead times, from 1 to 7 months into the future as per the SEAS5.1 lead times implementation. AI-based AoC enhance the potential of an operational system providing CSs into the next seasons for different crops. AoC are derived from impact classification at the crop yield surrogate model output. AoC are herein defined as those grid cells exhibiting a 5% or greater loss with respect to the mean yields of a reference period.

3. Presentation of the AI-enhanced Climate Services for Extreme Impacts - Water sector

3.1. Model evaluation

Figure 4 shows the performance distribution of AI-enhanced approaches under different gauged and ungauged conditions. The figure presents cumulative distributions of performance metrics across the stations. The analysis considers two dimensions: Space (referring to station locations) and Time (representing temporal periods), with *Train* indicating training data and *Test* indicating testing data. For gauged stations, where *Space: Train, Time: Test* represents the testing period, both single-basin and multi-basin approaches show comparable baseline performance during training. In temporal validation, the multi-basin approach exhibits marginally superior performance compared to the single-model approach. This enhancement may be attributed to either increased sample size or the model's ability to learn complex patterns when data from multiple stations are pooled together during training. In ungauged scenarios (*Space: Test, Time: Test*), where the single-basin approach calculated on each gauged station serves as a reference, the multi-model approach shows modest but notable improvements in performance with respect to raw E-HYPE.

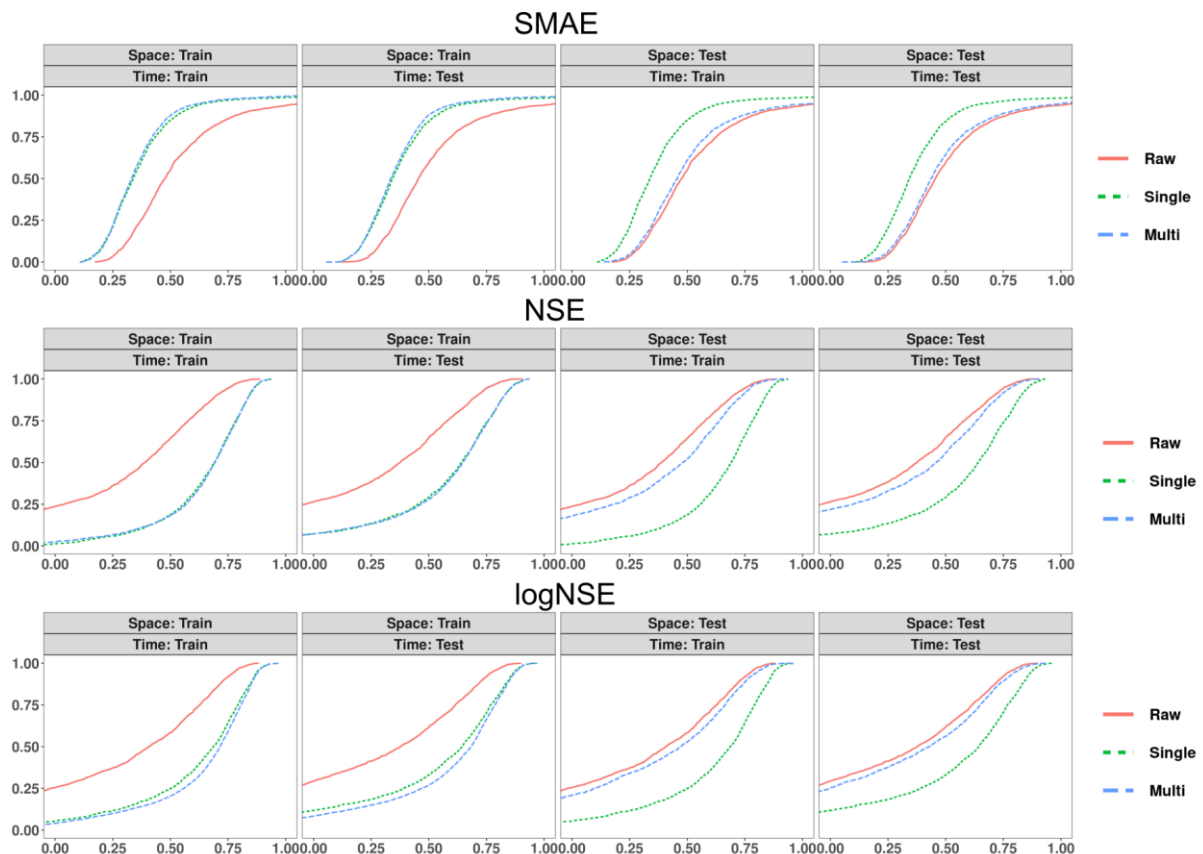


Figure 4 - Performance distribution of AI-enhanced approaches at different gauged and ungauged conditions. For SMAE, the perfect performance is 0 on the x-axis, therefore the model with better performance will appear with higher curve position (e.g. closer to 0); while for NSE and logNSE, the perfect performance is 1 on the x-axis, the model with better performance will appear lower in the plot (e.g. closer to 1).

Further analysis of different hydrological regimes reveals varying degrees of effectiveness across different basin types, categorized by their hydrological signatures, as shown in *Figure 5*. Particularly noteworthy is the performance in snow-dominated, baseflow-controlled catchments, where the multi-basin approach achieves considerable improvement in ungauged basins. This suggests the model's capability to effectively capture residual patterns in these catchment types. In precipitation-driven catchments with rapid response characteristics, the Nash-Sutcliffe Efficiency (NSE) shows minimal improvement for ungauged stations. This limited improvement is likely due to the inherent complexity and high variability of quick-response systems, which make their behavior patterns challenging to capture and generalize.

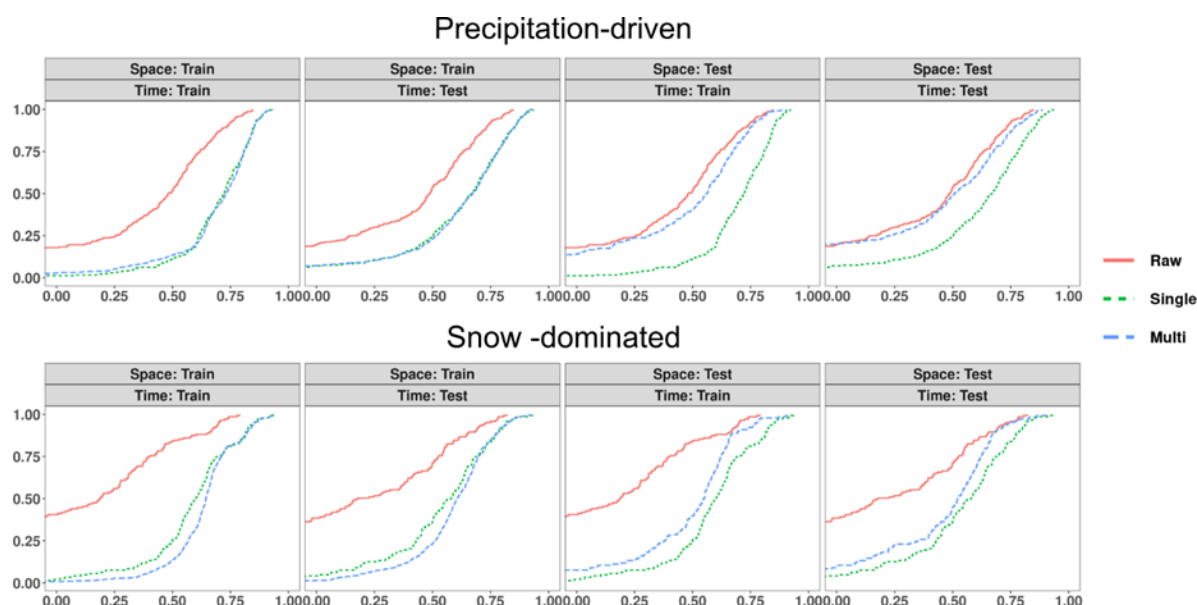


Figure 5 - Performance distribution of NSE of AI-enhanced approaches at different gauged and ungauged conditions in specific hydrological regimes.

3.2. Enhanced simulation of hydrological signatures

To assess model performance in capturing hydrological signatures, particularly high and low streamflow extremes, we compare AI-enhanced results with observations and raw E-HYPE outputs. Here only the multi-basin approach is used for AI enhancement, as the Single-basin approach lacks applicability to ungauged stations later.

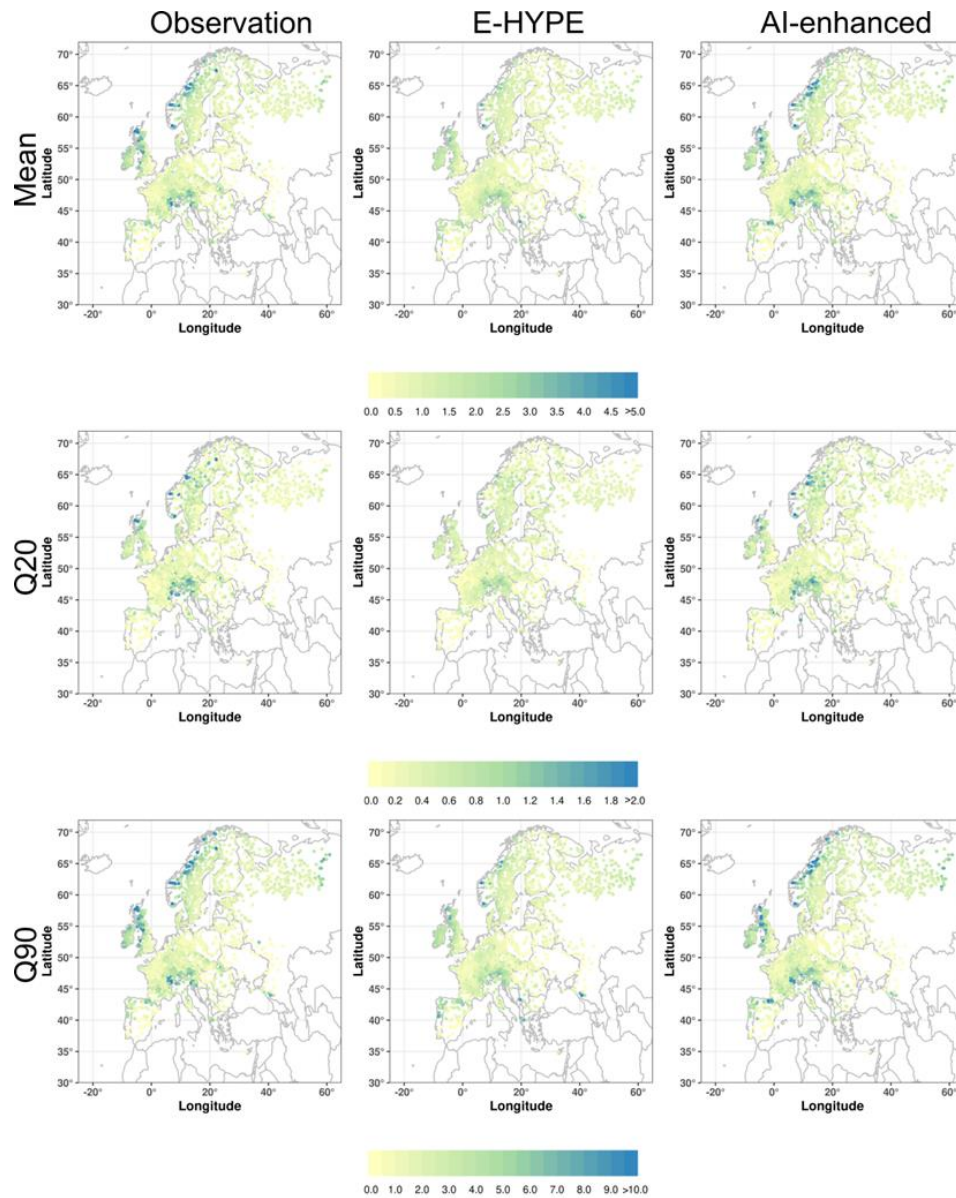


Figure 6 - Spatial comparison of streamflow characteristics across Europe showing observations, E-HYPE, and AI-enhanced predictions (multi-basin approach), with runoff values in mm represented by the color scales.

The spatial patterns of mean flow and flow quantiles (Q20 and Q90) across European catchments demonstrate the comparative performance of the E-HYPE model and our AI-enhanced approach against observation data (*Figure 6*). Both models effectively capture the general spatial distribution of flow characteristics of the observations, with notable patterns observed in northern Europe and along western coastal regions.

The mean flow distribution (top row) reveals that both modeling approaches successfully reproduce the observed spatial patterns. The AI-enhanced approach shows particular skill in representing the higher flow regions, especially in Scandinavian countries and along the Atlantic coast, where values frequently exceed 4.5 mm/day. These areas typically correspond to regions with significant precipitation and snowmelt contributions. Low flow conditions, represented by Q20 (middle row),

show more subtle spatial variations. The AI-enhanced model maintains consistency with observational patterns, particularly in capturing the gradual transitions between flow regimes. High flow characteristics Q90 (bottom row), demonstrate the models' capabilities in representing extreme flow conditions. The AI-enhanced approach shows notable improvement in capturing the intensity and spatial extent of high-flow regions. This enhanced performance is most evident in coastal Norway and the Alpine region, where complex terrain and diverse precipitation patterns influence flow regimes.

In summary, AI-enhanced approach shows superior performance in representing local variations and extreme values, suggesting that the AI enhancement effectively captures additional spatial and temporal dependencies that may be oversimplified in the E-HYPE model.

Figure 7 presents a scatterplot comparison of model signatures before and after AI enhancement, illustrating improved alignment with observations following post-processing. For mean flows, the AI-enhanced model shows a closer match to observations compared to the raw E-HYPE output, with reduced scatter and better adherence to the 1:1 line. This improvement is consistent across both training and testing periods and extends to ungauged stations. The enhancement is particularly evident for high flows (Q90), where the raw E-HYPE model systematically underestimates streamflow. The AI-enhanced version significantly reduces this bias, capturing flow variability more accurately. For low flows (Q20), both models perform similarly, though the AI-enhanced approach provides slightly better predictions, especially for gauged stations. Colors of the points denote different hydrological regimes of that station, with further details on the cluster classifications can be found in Pechlivanidis et al. (2020). The performance improvements remain robust across both temporal validation (testing period) and spatial validation (ungauged stations), confirming that the AI enhancement improves predictive accuracy rather than simply overfitting to training data.

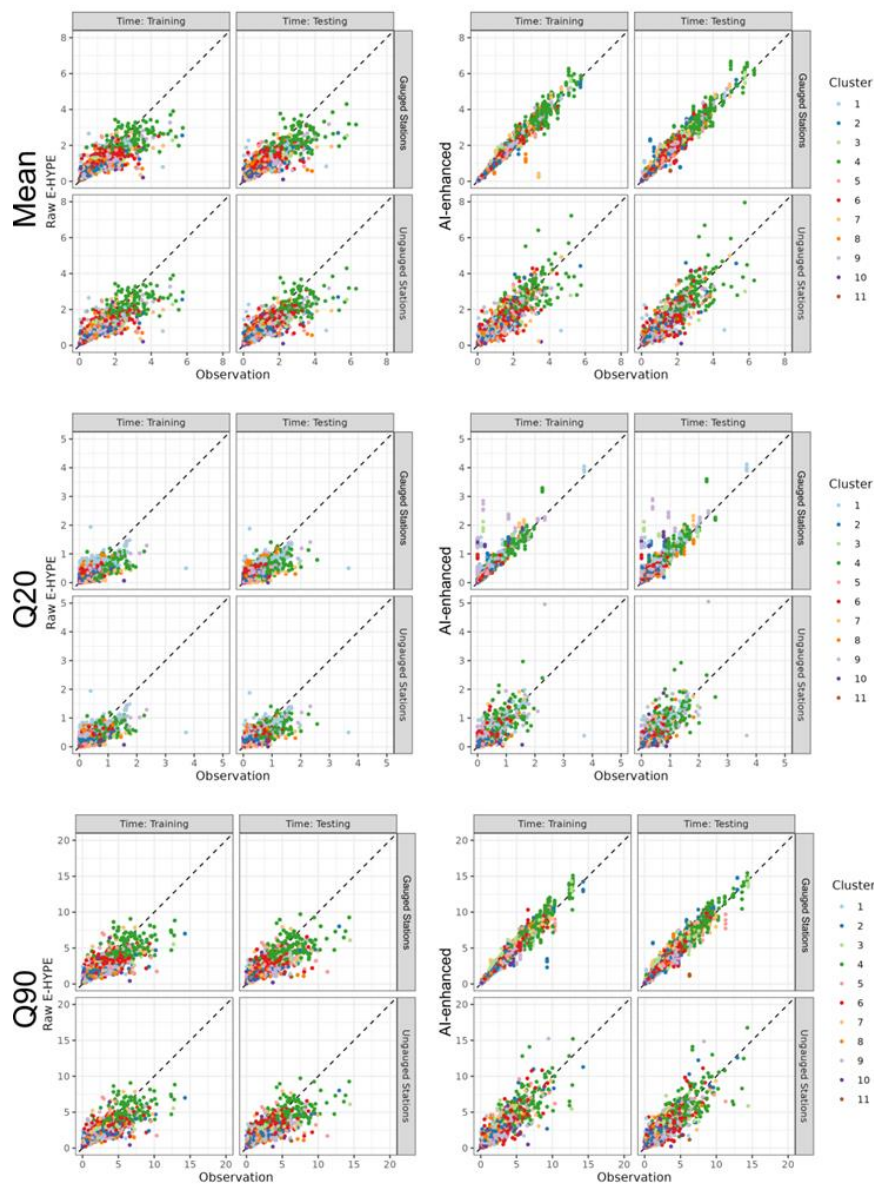


Figure 7 - Scatter plots comparing observed versus predicted streamflow signatures (Mean, Q20, and Q90) for raw E-HYPE (left panels) and AI-enhanced (right panels) results. Results are shown for both training and testing periods, and for gauged and ungauged stations. Different colors represent distinct hydrological clusters (1-11). The dashed line indicates the 1:1 perfect prediction line.

3.3. Updated streamflow signature pattern

Based on the results above, the AI-enhanced multi-basin approach is applied to all ungauged catchments across the pan-European area, generating updated streamflow simulations. The AI-enhanced streamflow signature patterns are presented in *Figure 8*, offering insights into the spatial distribution of the streamflow.

A comparison between the raw E-HYPE runoff estimates and the AI-enhanced runoff data (Q90) reveals significant differences in spatial patterns and variability. The E-HYPE model (top panel) appears to produce a more smoothed representation, potentially missing finer-scale hydrological features. In

contrast, the AI-enhanced model (bottom panel) provides a more detailed and heterogeneous distribution, particularly in hydrologically complex regions such as mountainous areas, northern Europe, and major river basins. One key improvement is the enhancement of high-runoff regions, where the AI-processed data reveals more pronounced localized hotspots, particularly in high-altitude and coastal areas. This refinement allows for a more accurate representation of extremes, improving the ability to capture small-scale variability.

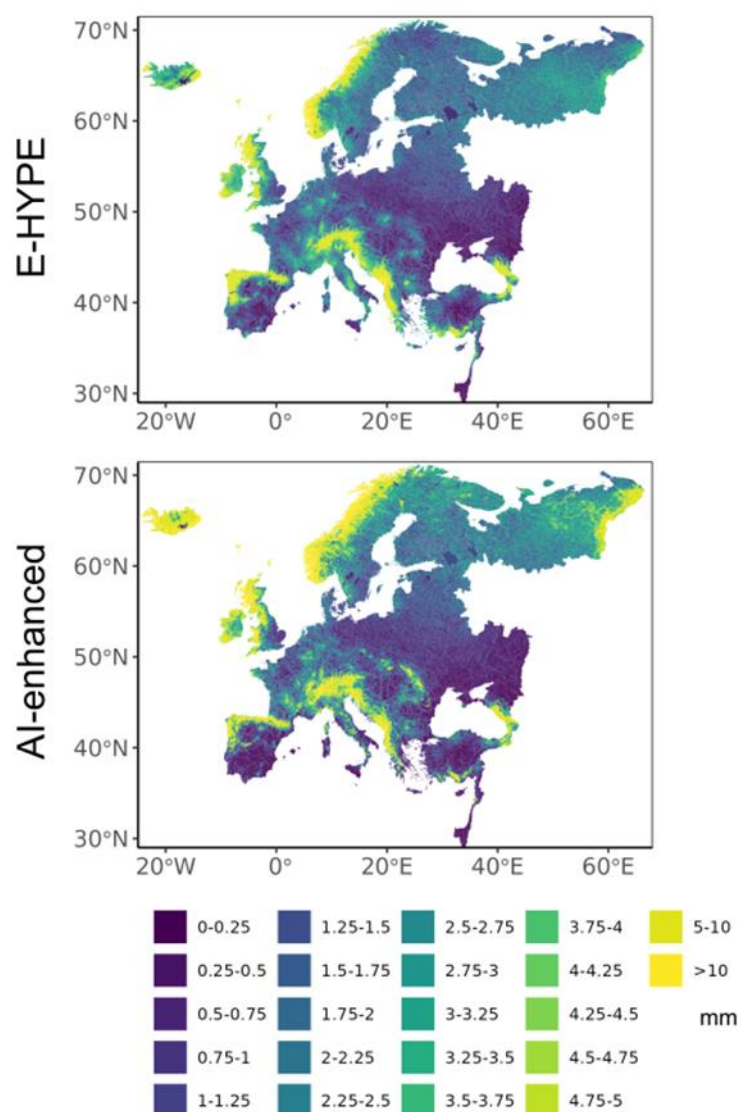


Figure 8 - Spatial patterns of simulated extreme high streamflows (90th percentile; Q90) before and after the AI enhancement.

This study introduces a multi-basin post-processing approach that combines data from multiple river systems to train a single, regionalized model. Unlike previous methods that develop separate models for each gauged basin in D6.2, this approach integrates static basin characteristics, e.g. climatic conditions, physiographical attributes, and hydrological regimes, into the model inputs. This enables the method to capture common hydrological patterns across basins and extend its applicability to

ungauged or data-scarce locations by transferring knowledge from hydrologically similar gauged systems.

A key contribution of this work is its potential to improve forecast skills in operational applications. Hydrological forecast skill depends on two main factors: the initialization of hydrological states and the forcing with bias-adjusted meteorological forecasts. Biases in both components can compromise forecast quality, as meteorological skill declines with lead time and errors propagate from the initial hydrological conditions. By applying the proposed post-processing framework on reforecasts and observations at each lead time, lead time-specific correction factors can be generated and then can be applied to new forecasts, to improve accuracy throughout the forecast horizon. These advancements align with global efforts, such as the IAHS HELPING scientific decade and the UN's Early Warnings for All initiative, which aim to reduce the impacts of hydrological hazards. Achieving these goals depends on accurate forecasts, highlighting the need to operationalize enhanced modelling frameworks.

4. Presentation of the AI-enhanced Climate Services for Extreme Impacts - Energy sector

4.1. AI-enhanced hydro inflows projections for the European energy sector

This chapter analyzes seasonal run-of-river (ROR) inflow patterns for a historical baseline (2015–2024) compared to projected medium-term (2025–2035) and long-term (2045–2055) periods under three climate scenarios: RCP2.6, RCP4.5, and RCP8.5. Key aspects discussed include analysis of changes in ROR inflow potential, such as alterations in seasonal variability between the medium-term and long-term climate projections. We discuss potential impacts on hydropower generation reliability. Finally, we present the results of the PRIMES-IEM model run for the long-term (2045–2055) periods under three climate scenarios: RCP2.6, RCP4.5, and RCP8.5.

The analysis estimates the climate variability at different time scales on the European power system, contributing to the EU climate neutrality target by 2050 using the PRIMES IEM model. We demonstrate the projections for the power generation sector, taking into account climate variability under current and planned policies at the European scale to demonstrate the impact of extreme events in the energy sector. We analyze the potential vulnerabilities arising in specific periods of the year due to projected climate variability. We discuss the opportunity to mitigate the impact of climate-related events on the European power system that undergoes the transition under the plan of the European Union to reduce greenhouse gas emissions and reach emissions neutrality by 2050 (European Commission, 2019).

We analyze hydropower generation: changes in hydropower generation at the national or regional level induced by climate change. The analysis focuses on the ROR-dominant European countries and presents data and findings on the country's level. We simulate European countries and focus our assessment on the ROR-dominant power systems under 3 different climatic scenarios (RCP2.6, RCP4.5, and RCP8.5) available from 9 GCM models' ensemble. The analysis focuses on the impact of annual and interannual variability hydropower under different realizations of climate projections.

4.1.1. Overview of changes in the hydro inflows: historical and climate projections

The availability of water for hydropower is crucial for assessing risks in a future decarbonized energy system under varying climate conditions. Decarbonizing the power sector requires a more flexible energy system, relying on options such as long-term storage, batteries, and pumped storage.

Maximizing hydropower resources, particularly in hydro-dominant countries, can enhance system reliability. We discuss the effect of changes in annual and intra-annual inflow patterns, affecting the run of river power generation and generation from river based reservoir capacities.

In our previous work (D6.2), we identified European countries where hydropower plays a dominant role in the power generation mix: Austria, Croatia, Finland, France, Germany, Italy, Latvia, Portugal, Romania, and Spain (*Table 5*). **Errore. L'origine riferimento non è stata trovata.** Machine learning (ML) techniques used to project future inflows performed best in countries where ROR contributed over 5% of total generation or exceeded 0.5 GW of installed capacity in 2020.

Fitting ML models for countries where hydropower is produced in small quantities proved more challenging, often leading to overfitting during training. The model struggled to identify meaningful sub-basin combinations to represent the limited number of generators in the dataset, resulting in low accuracy. This is because small hydropower generators are usually not driven by weather patterns but mostly by economic signals and system's flexibility needs.

Table 5 - Run-of-river generation share and capacity in each country (2020). Source: ENTSO-E.

Country	ROR % of overall generation (2020)	ROR capacity (GW, 2020)
Austria	27.6%	5.25
Croatia	20.3%	0.93
Finland	11.4%	1.85
France	1.4%	1.71
Germany	2.4%	4.01
Italy	7.0%	3.12
Latvia	48.2%	1.59
Portugal	13.9%	3.05
Romania	15.5%	3.62
Spain	6.4%	4.51

Note: In Sweden, run-of-river hydropower also plays a significant role in the national power mix. However, the incompleteness of time series data in the ENTSO-E database poses a limitation for effectively training the ML models used in the analysis and is therefore not included in the results.

Future projections of hydro inflows for each country were generated using a supervised machine learning regression model XGBoost¹. We found the relationship between river basins and the ROR annual generation profiles. These inflows represent the availability of water resources for hydropower generation. ROR generation patterns have been projected for the period 2025–2100 under three climate scenarios: RCP2.6, RCP4.5, and RCP8.5. Simulations were conducted using nine EURO-CORDEX models' ensemble. The following paragraphs describe the next steps towards constructing the hydropower generation profiles for run of river and lakes. We provide the descriptive analysis of changes in the hydropower inflows for the selected countries and discuss observed seasonal changes

¹A detailed description of the method is provided in Deliverable D6.2 and is not discussed here for the sake of conciseness. In the D6.2 we analyzed performance of XGBoost and Neural Networks models. The XGBoost was found to be more balanced in its prediction in relation to the ground truth and was chosen to construct the annual generation profiles.

in the projected hydropower inflows. We follow up with the analysis of climate-enhanced scenarios for the European power generation system under 2050 emissions neutrality target.

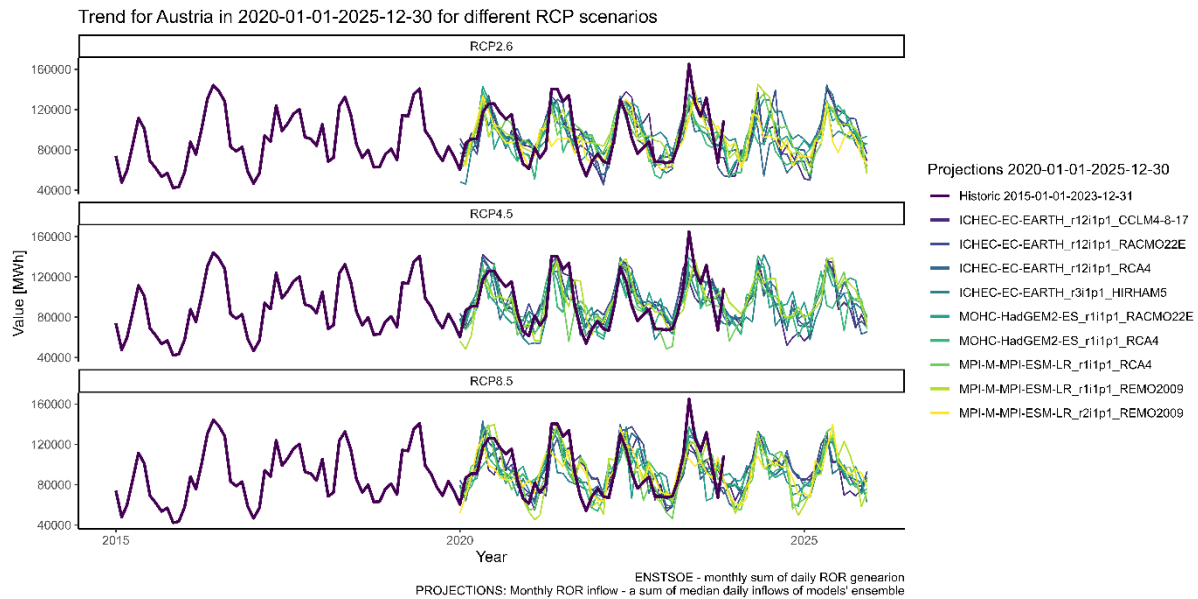


Figure 9 - Historical and projected inflows for Austria: models' ensemble and scenarios: RCP2.6, RCP4.5, RCP8.5 (Note: Graphs for the rest of hydro-dominant countries are available in the Appendix).

Figure 9 illustrates the projected ROR inflows for Austria, showing historical inflow variability alongside ML-projected inflows for each RCP scenario and the models' ensemble. The variability in ROR generation patterns for each considered country is assessed by analyzing projections for each RCP scenario and variation in the models' ensemble of the annual hydropower availability. Figure 10 illustrates the percentage deviation in annual hydropower inflow from the historical average (2015–2023) for the period 2020–2060 using the AI-enhanced annual inflow projections. For projections, annual inflows are estimated for the climate models' ensemble mean and 95th percentile. The comparison of the median and the 95th percentile ensemble projections, highlights the range of potential ROR inflow variations within the analyzed period. While the analysis of ensemble's median projection suggests moderate fluctuations with respect to the long-term historical average, the upper bound (95th percentile) reveals the possibility of extreme hydropower inflow reductions for the RCP8.5 scenario, particularly in southern European countries.

These findings emphasize that the management of hydropower availability and the plans for future resilience need region-specific adaptation strategies and consider climate scenarios and model ensembles to manage hydropower availability under future climate conditions. In particular we observe that:

- In the RCP2.6 scenario, hydropower inflow variability remains relatively moderate, with noticeable deviations in Spain, Portugal, and Latvia, where periods of increased and decreased inflows alternate. Similar changes are observed across the selected hydropower dominant countries, and we illustrate this in the example of Austria in section 0.
- Under RCP4.5, variability intensifies, particularly in Romania and Spain, where inflows frequently deviate by more than $\pm 10\%$ from the historical average.

- The RCP8.5 scenario exhibits the highest fluctuations, with several regions—including Spain, Portugal, and Latvia—experiencing prolonged periods of reduced inflows (below -15%) and occasional extreme increases (above +15%).

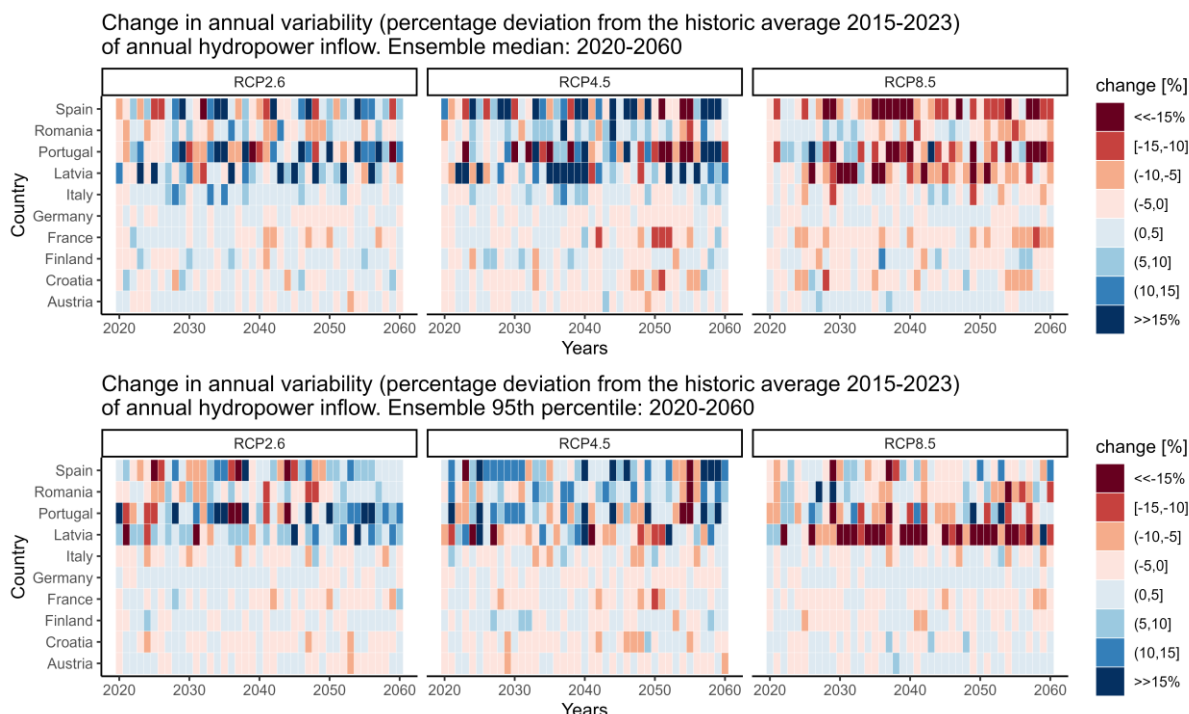


Figure 10 - Projected changes in annual ROR hydropower inflow variability (2020–2060): median vs. 95th percentile ensemble estimates.

4.1.2 Implications for hydropower inflow variability: Case Austria

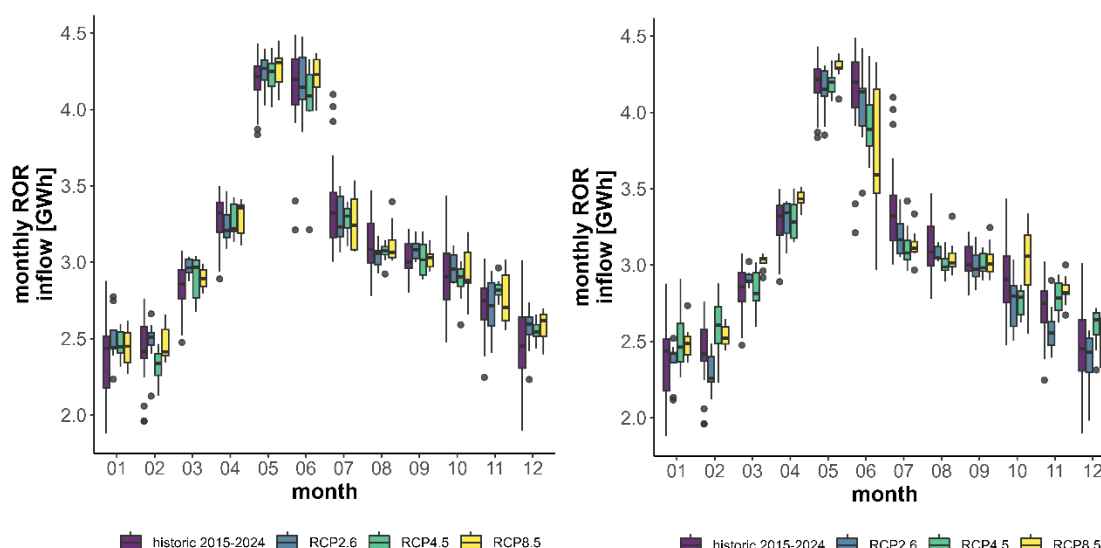
This section presents an overview of the ROR inflow variability on the hydropower generation profiles on the example of Austria. The historical period (2015–2024) shows a pronounced seasonal cycle in a ROR-dominant country - Austria. Winter months (Dec–Feb) exhibit the lowest flows of the year, reflecting typical alpine conditions where precipitation is stored as snow and natural runoff is minimal. February is generally the lowest point of the year in terms of ROR inflow. Flows begin to increase in early spring (March–April) as temperatures rise and snowmelt commences.

In the medium-term future (2025–2035), projected under RCP2.6, RCP4.5, and RCP8.5, Austria's seasonal inflow pattern still resembles the historical cycle, but with subtle shifts emerging. Under all three scenarios, winter and spring inflows tend to increase slightly relative to the 2015–2024 historical average, while summer inflows show slight decreases, especially under higher-emission scenario (RCP8.5). The overall shape of the seasonal distribution remains, but the timing and magnitude of flows' low and peak points begin to adjust:

- Winter (Dec–Feb): By 2025–2035, winter ROR inflows are higher than in the historical period, particularly under RCP4.5 and RCP8.5. Winter inflows show a substantial increase by mid-century: in RCP8.5 (2045–2055), monthly flows in December, January, and February are considerably higher than the historical level.

- Spring (Mar–May): Spring inflows rise noticeably in the 2025–2035 period across all RCP scenarios, with the effect growing from RCP2.6 to RCP8.5. Spring flows are projected to become much larger, with peak inflow arriving earlier by 2045–2055, especially under RCP8.5: May inflow exceeds June as the month with the highest inflow in the historical period, representing the new annual peak.
- Summer (Jun–Aug): During early summer, June in 2025–2035 still tends to be a high-flow month. Mid-summer inflows are starting to decline, possibly due to earlier depletion of the snowpack and drier summer conditions. Under RCP4.5, July–August declines are present but modest, and under RCP2.6 they are minimal (summer flows remain close to historical levels). By 2045–2055 The decline in mid-summer flows becomes very pronounced for July and August. The median June inflow under RCP8.5 is *lower* than the historical median June (and also lower than May’s in this scenario).
- Autumn (Sep–Nov): September and October inflows in 2025–2035 remain relatively low, similar to the historical baseline. By mid-century, autumn flows remain relatively low, with RCP8.5 introducing a slight extension of the summer dryness into early autumn.

In summary, the medium-term projections show that Austria’s peak inflow timing is beginning to shift earlier (from June to May). Under RCP2.6, the peak remains centered on June (only a marginal advance). Under RCP4.5, peak flows are distributed between May and June. By 2045–2055, Austria’s peak runoff clearly shifts from early summer to late spring in the projections.



Note: Historic period 2015–2023 and ensemble median monthly ROR production. Graphs for the rest of ROR-dominant countries are available in the Appendix.

Figure 11 - Projected changes in annual ROR hydropower inflow variability (2020–2060): median vs. 95th percentile ensemble estimates.

Deviations of the occurrence of the peak inflow are modest in 2025–2035 but become more significant in 2045–2055. If electricity demand patterns do not align with this (for instance, if summer demand is higher due to cooling needs), there could be a mismatch in timing leading potentially to periods of deficit of ROR generation within each year. If Austria or neighboring regions (e.g. Germany) face peak

electricity demand in July-August, the hydropower supply might be contributing less to the power balance with river flows below historical levels. The earlier spring peak may also overlap with periods when reservoirs are full due to snowmelt, limiting the potential to store electricity during surplus of ROR generation.

4.2. Long-term trends of hydropower generation variability in Europe

4.2.1 Hydropower variability and capacity factor assumptions for hydropower systems

As a next step, the analysis investigates trends in 5-year changes in ROR hydropower inflows for selected European countries identified as ROR-dominant. Historical inflow data (2015–2023) is compared to projected future inflows under three climate scenarios (RCP2.6, RCP4.5, and RCP8.5) to identify long-term trends in hydropower availability, *Table 6*. The use of 5-year intervals aligns this inflow analysis with the time resolution of the PRIMES-IEM model horizon, facilitating integrated assessments of climate impacts on future hydropower availability.

Table 6 - Change in potential hydro generation for RCP scenarios for selected countries for climate scenarios: RCP2.6, RCP4.5, RCP8.5.

		Average annual inflow (historical period)	Change in annual five-year average ROR inflow compared to historical period*					
		[MWh]	[%]	[%]	[%]	[%]	[%]	[%]
Country	RCP	2015-2023	2025	2030	2035	2040	2045	2050
Austria	RCP2.6	1,121,087	1.6	1.1	0.5	1.5	-0.3	-0.4
	RCP4.5	1,138,844	-3.1	0.4	-1.3	-2.4	-3.9	-4.6
	RCP8.5	1,119,161	2.6	1.2	2.2	1.4	1.0	1.4
Croatia	RC2.6	72,470	0.0	-1.3	-1.7	0.0	-0.2	1.1
	RCP4.5	73,100	-2.8	-3.1	-2.1	-2.5	-5.2	-8.9
	RCP8.5	74,198	-2.4	-2.9	-2.3	-3.6	-1.5	-0.4
Finland	RC2.6	573,184	-1.6	0.8	0.7	-0.8	0.2	2.6
	RCP4.5	568,171	3.0	2.8	2.7	0.6	-1.4	2.3
	RCP8.5	568,136	-1.3	-1.8	3.3	0.2	0.5	-4.1
France	RC2.6	1,804,906	1.6	0.9	-1.6	-2.2	-2.0	-7.8
	RCP4.5	1,805,637	0.2	-1.0	0.2	-0.3	-3.6	-10.2
	RCP8.5	1,857,247	-4.3	-1.2	-1.0	-3.5	-1.9	-3.9
Germany	RC2.6	646,567	0.8	0.6	0.7	0.0	-2.4	-2.6
	RCP4.5	649,689	-0.4	0.0	-0.3	-0.5	-1.3	-4.9
	RCP8.5	643,821	0.8	0.4	1.2	0.4	2.4	0.3
Italy	RC2.6	1,482,194	4.7	1.2	3.7	0.5	0.4	3.1
	RCP4.5	1,450,163	2.5	2.7	5.3	3.2	1.2	-0.2
	RCP8.5	1,503,412	-3.4	-1.2	0.3	0.5	0.4	-1.6
Latvia	RC2.6	135,289	-1.6	-5.7	3.4	-0.4	1.9	-5.5
	RCP4.5	138,804	-3.5	-3.9	15.3	-2.6	-6.4	1.1
	RCP8.5	147,013	-8.6	-8.5	-7.0	-11.6	-9.7	-9.4
Portugal	RC2.6	216,929	1.7	-1.1	-5.3	-3.6	2.4	15.5
	RCP4.5	213,074	11.5	-4.2	8.2	7.9	9.3	-9.2

	RCP8.5	230,711	-5.6	-0.6	-17.6	-5.5	-2.7	1.9
Romania	RC2.6	392,457	-0.3	0.8	3.6	-3.6	-3.8	7.5
	RCP4.5	384,860	4.2	4.7	7.2	8.1	6.5	0.3
	RCP8.5	408,160	1.1	5.1	-4.5	-1.2	1.8	-1.9
Spain	RC2.6	263,630	5.4	6.2	4.1	-1.2	3.6	8.3
	RCP4.5	256,356	11.1	6.5	8.7	10.1	13.7	-0.9
	RCP8.5	291,093	-8.6	-2.2	-17.2	-4.4	-4.7	1.1
*Changes in annual run-of-river (ROR) inflows relative to the historical period are presented as five-year averages, corresponding to the modeled time steps.								

Country-specific details for ROR hydropower inflow variability in changing climate conditions:

Austria: Under RCP2.6, slight positive changes in the medium term are projected (+2.6%), stabilizing towards mid-century. In contrast, under RCP4.5 and RCP8.5, inflows are projected to decline by approximately 4.6% by mid-century, reflecting increased drought conditions.

Croatia: Moderate negative inflow changes are consistent across all RCP scenarios, worsening significantly under RCP4.5 (-8.9%) by 2055.

Finland: Under RCP4.5 and RCP8.5, inflows initially rise moderately but show some fluctuations and declines by mid-century.

France: Notably reduced inflows projected under RCP8.5, especially pronounced by 2050 (-10.2%).

Germany: Stable to slightly negative inflow changes appear in medium to long-term periods, becoming most significant under RCP4.5 and RCP8.5 (-4.9%).

Italy: Short-term increases under RCP4.5 and RCP8.5, followed by stabilization or slight declines approaching mid-century.

Latvia: Significant positive variability appears in the medium-term under RCP4.5 (+15.3%) but diminishes substantially by 2055.

Portugal: Fluctuations occur, with notable increases under RCP4.5 in the medium term (+11.5%) followed by significant declines in long-term scenarios, particularly under RCP8.5 (-17.6%).

Romania: Consistent moderate-to-strong positive changes in the medium term under RCP4.5 (+8.1%) contrast sharply with significant negative trends under RCP8.5 in mid-century (-4.5%).

Spain: Negative changes dominate, especially under RCP8.5 scenario by 2050 (-17.2%), indicating substantial future vulnerability in ROR inflows.

With the aim to focus on the 2050 horizon, we estimate capacity factors for (the 5-year period related to) this year under the climate scenarios for each ROR-dominant country². These factors serve as an input for interconnected PRIMES-IEM modeling, assessing the reliability and operational flexibility of hydropower under projected climate-driven inflow changes. Capacity factors for ROR in the identified

² From now on we indicate the 5 year period related to 2050 as '2050'.

countries for the year 2050 have been developed using ensemble median projections across the RCP scenarios. The resulting hourly capacity factors illustrate the anticipated operational performance of hydropower plants, allowing integration into broader energy system models. Notably, these projections indicate an effect of climate scenarios on ROR generation potential.

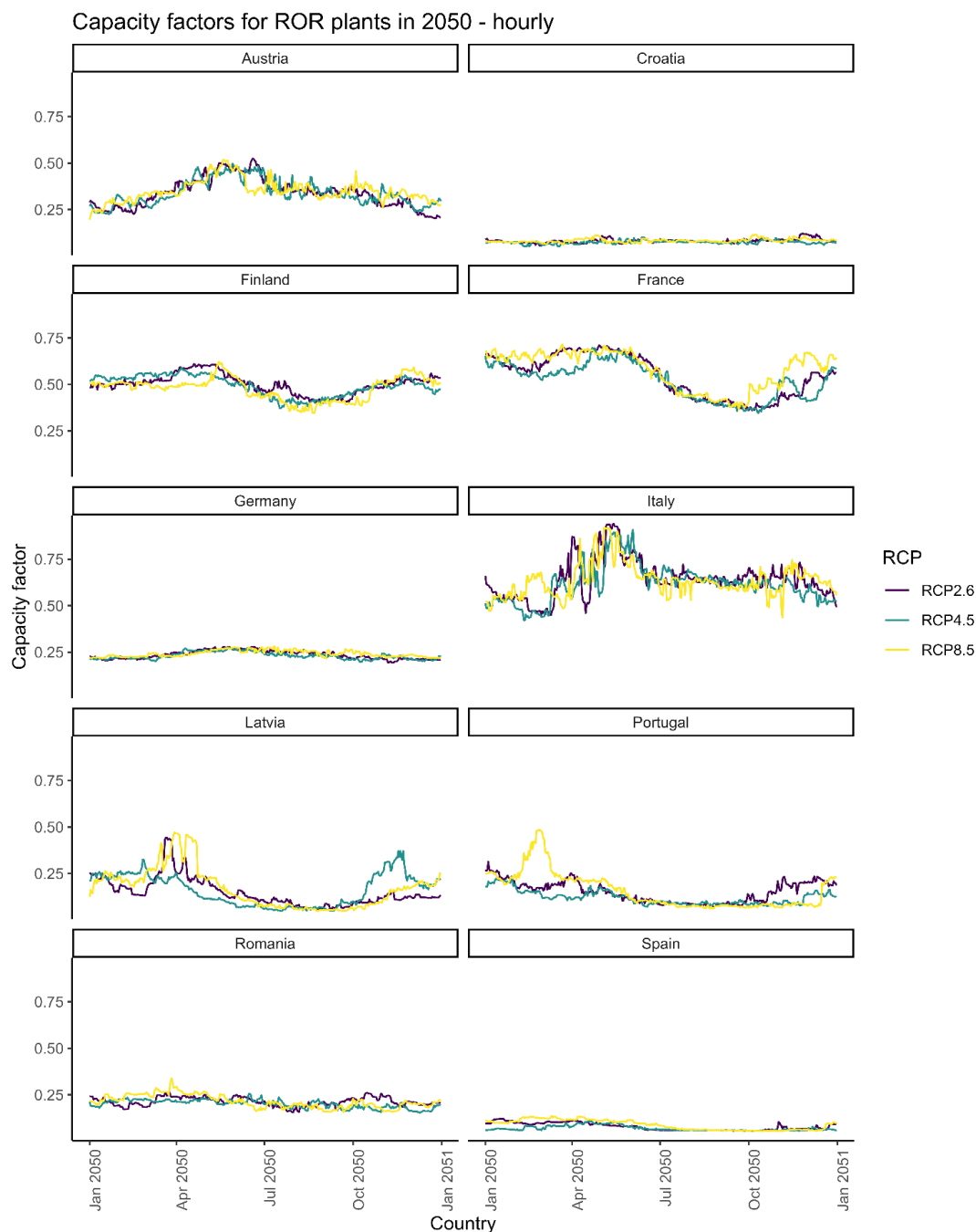


Figure 12 - Estimated ROR capacity factors: projections for 2050 for climate scenarios: RCP2.6, RCP4.5, RCP8.5.

4.2.2 Impact of climate variability on hydropower generation and cross-cutting effect across European power generation sector under emissions neutrality target

Scenario design

In order to assess the cross-cutting effect of hydropower variability, we simulate the future European power system that is aligned with long term EU energy and climate targets (i.e. reaching the climate neutrality target by 2050). We focus on the years 2030 and 2050 to analyze the behavior of the European power system under different projected climate conditions. Projections on the inflows for 3 RCP scenarios are available from the E-HYPE model: RCP2.6, RCP4.5, RCP8.5. For each RCP scenario we analyze the 9 EURO-CORDEX models ensemble's median. To assess the impact of extreme climate realizations, we model the 5th and 95th percentile ensemble projections of hydropower inflows under the RCP8.5 and RCP4.5 scenarios.

The main focus of the analysis is to show difference in the response to climate events with the decarbonized energy system described by the climate neutrality scenario, meeting the EU energy and climate targets in 2030 and the emissions neutrality target in 2050. *Figure 13* and *Figure 14* show the power mix in 2030 and 2050 respectively. At this point of the close future, decarbonization of the power generations sector is driven by the EU Renewable Energy Directive³. Long term emissions reduction in energy in energy sectors ensured by EU emissions trading system (ETS) and 2050 EU emissions neutrality target. National coal and nuclear phase out policies announced by the European member states form the power mix by 2030 and beyond. The 2050 emissions neutrality target drives further reduction in fossil run power generation capacities, increasing the share of intermittent renewable generation (solar and wind) in the power mix.

The following paragraphs present the results of a model-based scenario analysis of the European power system along a decarbonization pathway. The discussion focuses on key system-level parameters to highlight the cross-cutting effects of hydropower variability. The main evaluation metrics, summarized in *Table 2*, include: annual hydropower and fossil generation, total system costs, fuel purchase costs, and total GHG EU emissions. Additionally, we examine changes in net electricity imports between modelled regions to illustrate the impact of hydropower variability on cross-border power exchanges and system balancing needs. In the following chapters, the analysis of climate-related hydrological variability encompasses both ROR generation and generation from lakes, representing a mix of ROR and large reservoir operations.

As shown in *Figure 1*, the analysis follows five-year time steps, in line with the model's temporal resolution. Furthermore, we assess climate impacts using a set of RCP scenarios, incorporating the ensemble median as well as the 5th and 95th percentile projections to capture a broad range of possible climate outcomes. From a scenario design perspective, the analysis is guided by the overarching objective of achieving climate neutrality by 2050 and is structured accordingly to reflect this long-term target.

³ To the time of the scenario design, 2030 policy landscape was designed under the framework of 2021 legislative proposals for Fit for 55 policy package: 38-40% renewable share in the Final energy consumption: Proposal Renewable Energy Directive COM (2021/557).

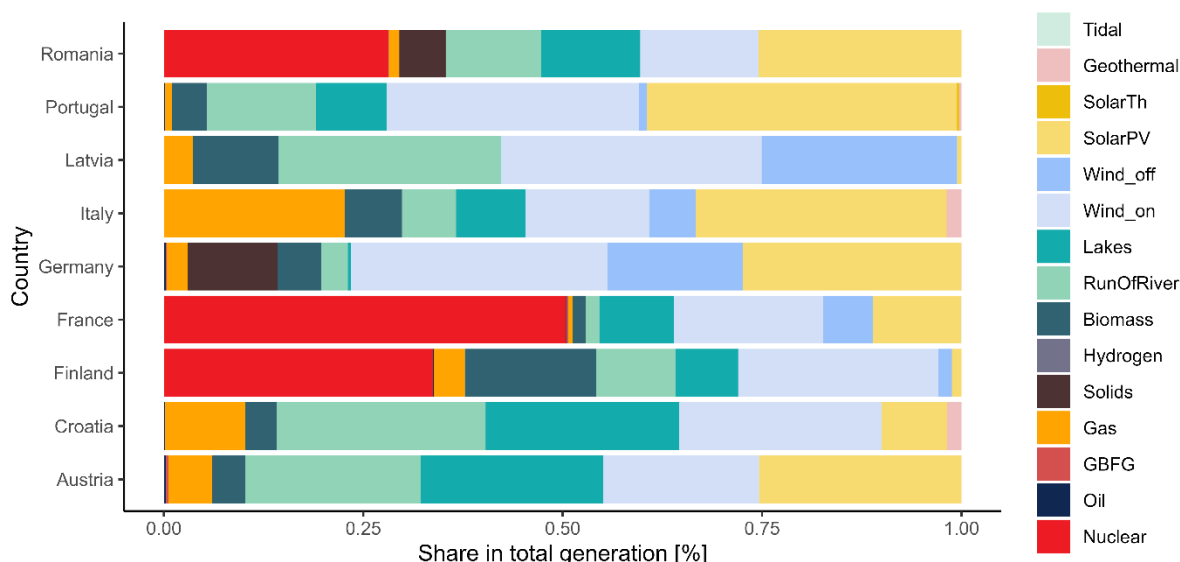


Figure 13 – Power mix in 2030. Emissions neutrality scenario under the historical median weather years. Source: PRIMES-IEM.

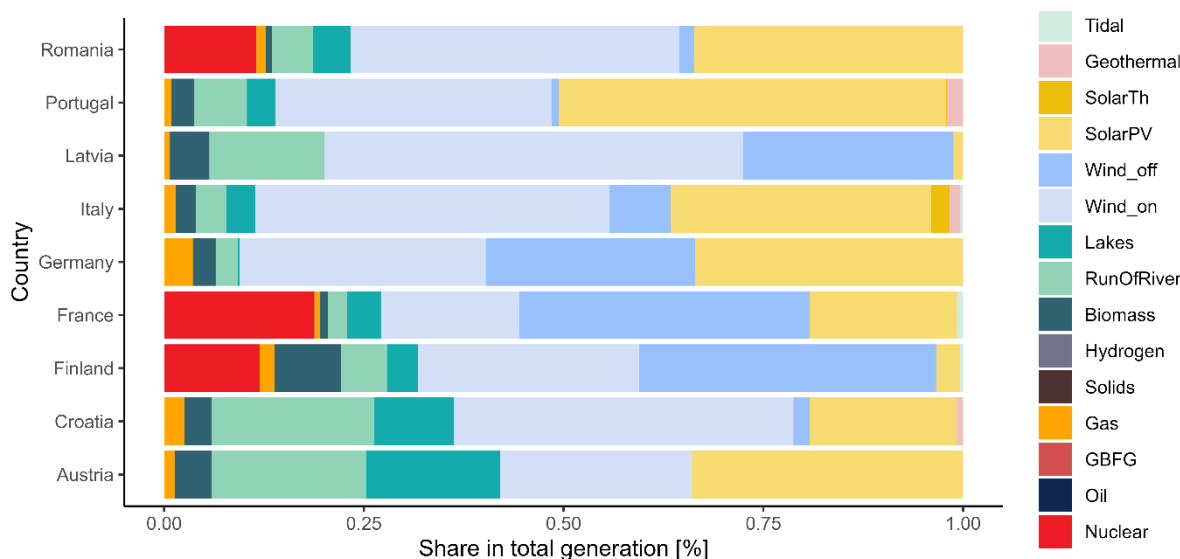


Figure 14 - Power mix in 2050. Emissions neutrality scenario under the historical median weather years. Source: PRIMES-IEM

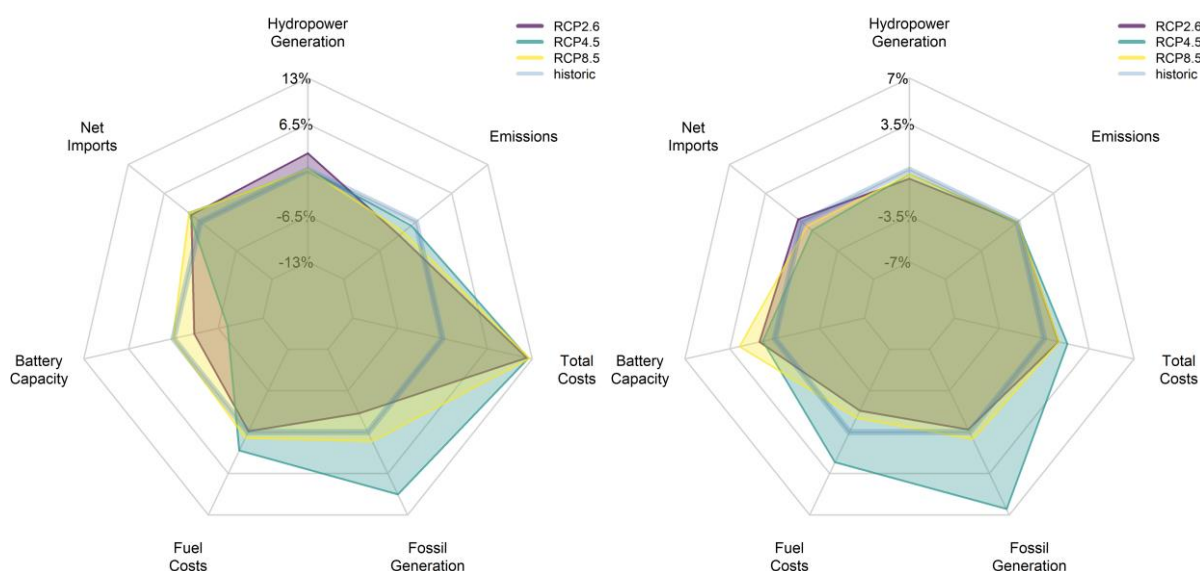
Climate-enhanced scenarios – projections of ensemble median

As fossil-based power generation declines by 2050, the power system becomes increasingly reliant on flexibility options to balance electricity supply and demand. These include various storage technologies, cross-border electricity exchanges, and the utilization of power generation capacities with fast ramp-up capabilities.

With relatively small variations in annual hydropower generation projected for 2030, and seasonal shifts in inflow peaks (as discussed in Section 4.2.1), the energy system must identify cost-optimal solutions to maintain balance under changing hydropower availability due to future climate conditions across European regions.

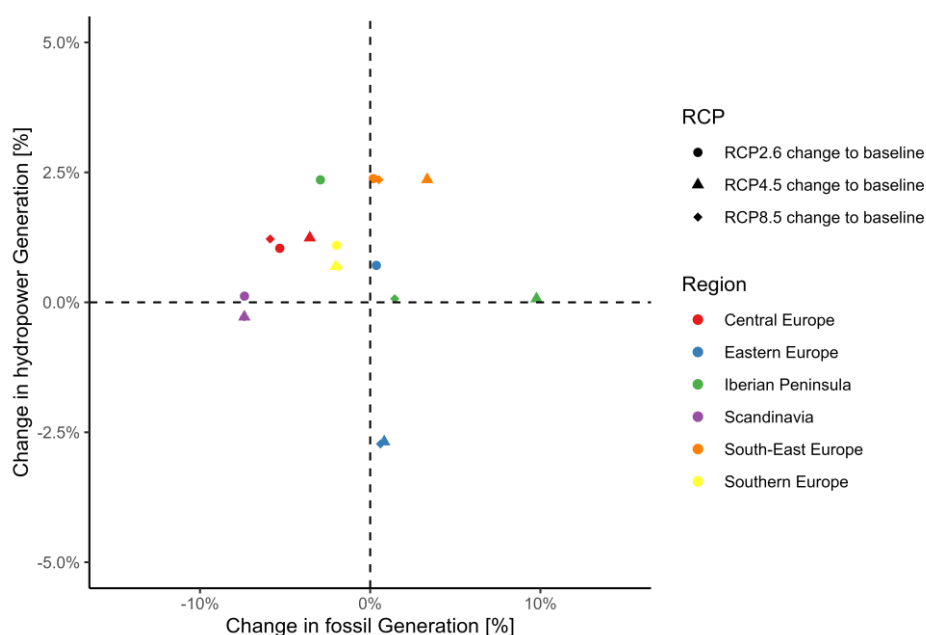
The system's response differs between 2030 and 2050, reflecting evolving infrastructure and constraints. In 2030 (see *Figure 15, left*), the system compensates for hydropower variability primarily through remaining gas-fired power generation, leading to increased fuel costs and higher total power generation costs, including ETS-related payments. Under the RCP2.6 scenario, higher annual hydropower generation compared to the long-term historical median results in lower fossil fuel use, reduced emissions, and lower system-wide costs. However, to manage seasonal hydropower shortfalls, the system relies more heavily on electricity imports from neighboring regions, reflected in higher net imports.

By contrast, in 2050, (see *Figure 15, right*), with more advanced system flexibility and reduced fossil fuel capacity, the optimization strategy shifts further toward renewables integration, storage, and cross-border coordination, underlining the critical role of climate-informed energy planning.



Note for this and following Figures: Hydropower includes power generation from run-of-river power plants and reservoirs. Fossil generators include natural gas, oil, and coal. The axes represent the percentage difference of the climate neutrality scenario for the historical median weather years and RCP scenarios. Results are presented for the aggregated region France, Portugal, Spain.

Figure 15 – Comparing 3 climate scenarios for 2030: RCP2.6, RCP4.5, RCP8.5. Source: PRIMES-IEM.



Central Europe: Austria, Belgium, Luxembourg, Netherlands, Germany. Eastern Europe: Czech Republic, Estonia, Latvia, Lithuania, Poland, Slovakia. Eastern Europe: Czech Republic, Estonia, Latvia, Lithuania, Poland, Slovakia. Scandinavia: Denmark, Finland, Sweden, Ireland. South-East Europe: Bulgaria, Croatia, Hungary, Romania, Slovenia. Southern Europe: Cyprus, Greece, Italy, Malta. Iberian Peninsula: France, Portugal, Spain.

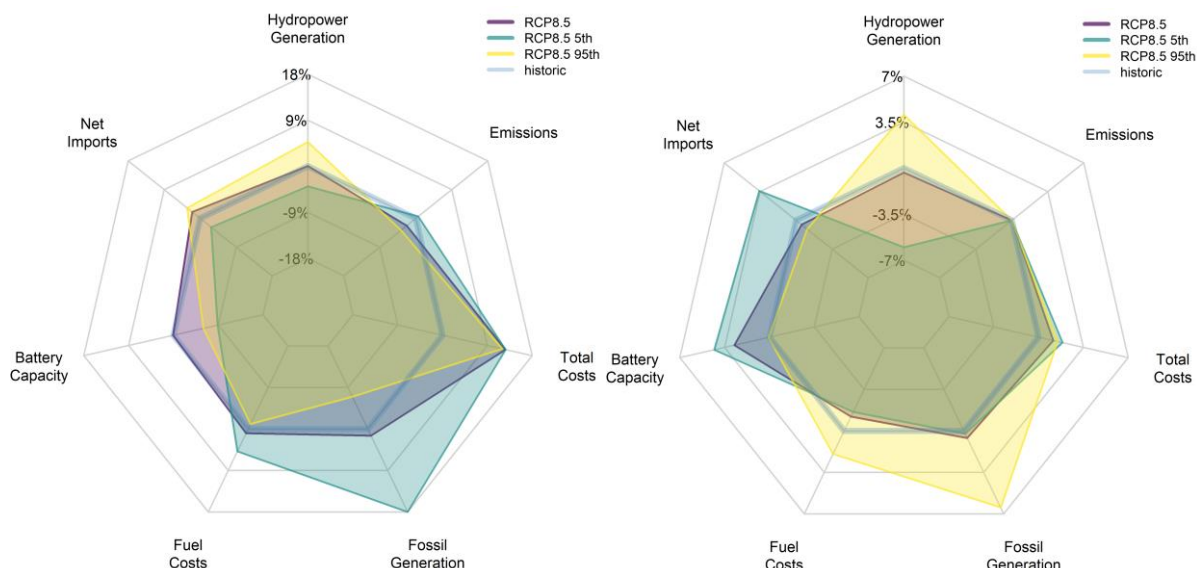
Figure 16 – Relationship between hydropower generation and fossil run generation for 3 climate scenarios for 2030: RCP2.6, RCP4.5, RCP8.5 Source: PRIMES-IEM.

Changes in hydropower generation—based on the median of the model ensemble projections by 2030 under the RCP climate scenarios—have a noticeable impact on fossil-based power generation, which at that point consists mainly of gas-fired capacities. These changes enable lower fossil-fired power generation (*Figure 16*), contributing to lower emissions and fuel purchase costs.

By 2050, however, the influence of hydropower variability on emissions and fossil generation becomes marginal, as fossil-based generation is nearly eliminated from the EU power system. At this stage, the system's flexibility is largely ensured through various storage options and interconnection capacity between the member states and regions.

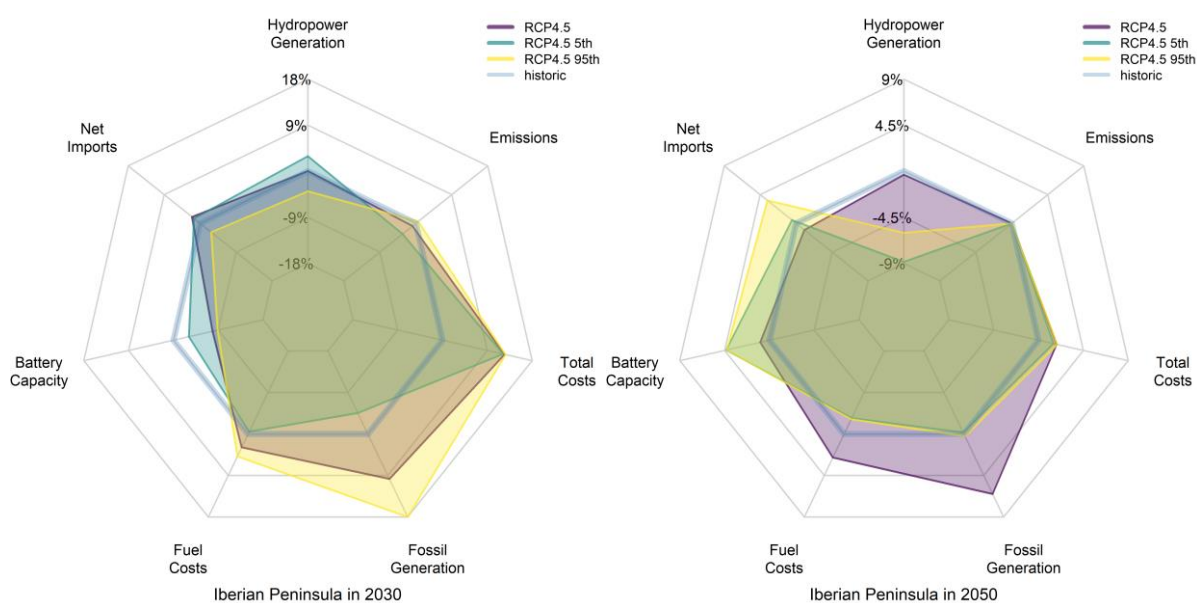
Climate scenarios – projections of lower and upper percentile of models' ensembles

To capture the full range of potential impacts, the ensemble's median projections can be complemented by an analysis of the 5th and 95th percentiles, representing less likely but more extreme climate realizations. The following paragraphs provide a detailed assessment of the power generation system a tightly interconnected regional cluster comprising France, Portugal, and Spain. Within this region, Portugal and France have been identified as ROR dominant countries, with the most reliable inflow projections. The analysis focuses on the structure and response of the power system under two climate scenarios—RCP4.5 and RCP8.5—including their 5th and 95th percentile realizations, which capture the range of potential climate impacts on hydropower availability and broader system performance.



Plots for other countries are available in the appendix. Plot is shown for the interconnected regional cluster comprising France, Portugal, and Spain. Plots for selected EU regions are provided in the Appendix.

Figure 17 - Comparing climate scenarios for: RCP8.5 in 2030: median, lower 5% and upper 95%. Source: PRIMES-IEM.



Plots for other countries are available in the appendix. Plot is shown for the interconnected regional cluster comprising France, Portugal, and Spain. Plots for selected EU regions are provided in the Appendix.

Figure 18 – Comparing climate scenarios RCP4.5 in 2030 and 2050: median, lower 5% and upper 95%. Source: PRIMES-IEM.

The analysis of the differences between lower and higher inflows projections under different climate scenarios on the interconnected regional power system, reveals that hydropower generation variability strongly influences system performance in 2030. In scenarios with higher hydropower inflows—such as the 95th percentile projections under RCP4.5 and RCP8.5—we observe reduced reliance on fossil fuel generation, leading to lower emissions and fuel costs (Figure 17 and Figure 18).

These scenarios also show reduced battery storage capacity and net imports, indicating that abundant hydropower can contribute to lower need for other flexibility mechanisms. Conversely, in low inflow scenarios (5th percentile), limited hydropower availability forces the system to heavily rely on gas-fired generation and storage, resulting in increased emissions and total costs.

By 2050, when the system undergoes the transformation towards decarbonization of the power generation sector, see *Figure 18*, the differences between climate projections become less pronounced, particularly in emissions and fossil use. However, hydropower variability continues to affect system dynamics. Higher inflow scenarios still lead to slightly lower battery capacity and imports, while low inflow scenarios necessitate greater use of flexibility options available. These trends illustrate how, even in a deeply decarbonized system, climate-driven hydrological variability continues to shape infrastructure needs and cost structures—especially through its influence on seasonal balancing discussed in Section 4.1.2.

The comparison between 2030 and 2050 highlights a transition in system flexibility: from fossil-based backup in the near term to a greater reliance on storage and interconnectors in the long term. By 2050, with natural gas fired power generation nearly absent, the system becomes reliant on deployment of batteries and cross-border electricity exchanges to manage seasonal and climate-related fluctuations. These findings underscore the importance of integrating climate resilience into long-term power system planning.

Even in fully decarbonized scenarios, climate variability—particularly shifts in hydropower inflows—can significantly affect system costs, storage needs, and power exchanges between the regions and countries. Integrating climate risk into energy planning is essential to ensure the stability and affordability of future power systems, under the risk of climate variability.

Using percentile-based hydropower inflow projections (5th–95th) under different RCP scenarios reveals how extreme, less likely climate outcomes could shape the ongoing transformation to the decarbonized power generation system. This approach enables more resilient, cost-effective infrastructure investment decisions—especially in regions with strong seasonal hydro dependencies.

5. Presentation of the AI-enhanced Climate Services for Extreme Impacts - Food sector

5.1. Performance

We assess the AI model performance of the observation-level quality of the SEAS5.1-forced and CMIP6-forced surrogate yield outputs by comparing those with the ERA5-forced surrogate yield outputs. Hereafter, we show example metric outputs that are used, with the same process applied to all considered datasets.

5.1.1. SEAS5.1

Crop yield simulations with SEAS5.1 are generated using the surrogate model both for the reforecast period (1993-2016) comprising 25 ensemble members and the forecast period (2017-2023) comprising 51 ensemble members, focusing on the first lead time component. This choice is due to the known heavy deterioration of model accuracy after the one-month lead time.

Figure 19 - shows the year-by-year Pearson correlation plot between all SEAS5.1 ensemble members and ERA5-forced yields from 1993 to 2016. In most years, correlations lie above 0.5, often clustering between 0.6 and 0.8. This indicates that, on average, the first-time lead SEAS5.1 ensembles capture a reasonable fraction of the year-to-year ERA5 TWSO variability. Some years (e.g. the mid-1990s) show better agreement than other years (e.g., early 2010s) greater uncertainty among members. Temporal variations in the correlation can denote internal climate variability, model biases, or sensitivity to initial- conditions.

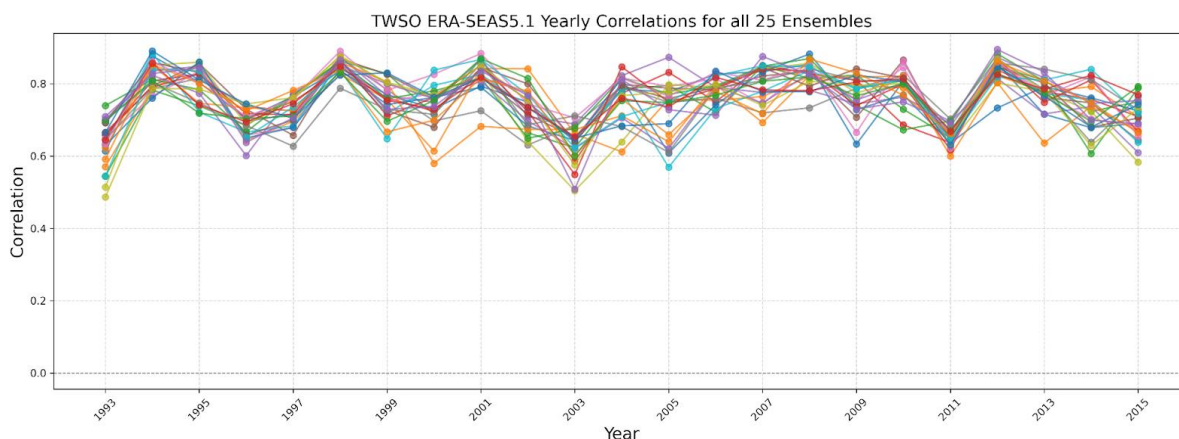


Figure 19 - Annual Pearson correlation plot for an ensemble of 25 SEAS5.1 reforecasts and ERA5-forced yields from 1993 to 2016.

Figure 20 and *Figure 21* represent the output metrics for the reforecast period using the SEAS5.1 ensemble member 01 (SEAS5.1_ens01).

Figure 20 shows two pronounced peaks in the simulated (1993-2016) crop yield distribution (one near low TWSO values and another around 12,000 kg/ha), whereas the ERA5 has broader peaks around 2,000–3,000 kg/ha and 10,000–11,000 kg/ha. This suggests that the SEAS5.1_ens01 is producing more frequent very-high TWSO occurrences compared to ERA5. In terms of the degree of overlap, the curves overlap substantially e.g., near 10,000–11,000 kg/ha.

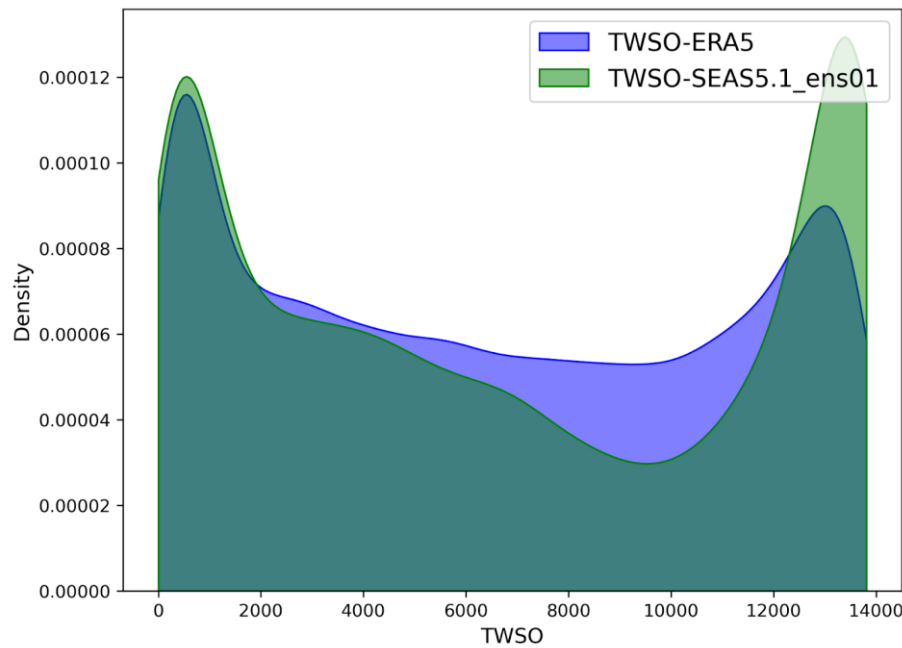


Figure 20 - Overlaid probability density functions (PDFs) of TWSO (kg/ha) simulations forced with ERA5 and the SEAS5.1 ensemble member 01 (SEAS5.1_ens01)

Figure 21 a displays the correlation between ERA5 and the SEAS5.1_ens01 TWSO. Overall, much of the domain shows moderate to strong positive correlations suggesting that this ensemble member generally captures the patterns in ERA5-forced TWSO reasonably well. Some localized patches of lower correlation indicate where the ensemble diverges from ERA5, possibly due to local-scale processes or/and poorer performance of the SEAS5.1 ensemble.

The Mediterranean coastline and the mountainous areas such as the Alps show consistently higher errors (e.g. very high and very low mean bias values in these areas), whereas the western Mediterranean near the Iberian Peninsula and mainland Europe show smaller errors and better agreement with ERA5 (Figure 21b, c and d). This suggests that in regions with complex topography and heterogeneous surface processes such as the land-sea interaction, the SEAS5.1 has greater difficulty reproducing accurately observed TWSO. Although regions such as the Alps and other mountainous areas generally yield misaligned results, they can be disregarded as most of these regions are not used for cultivation.

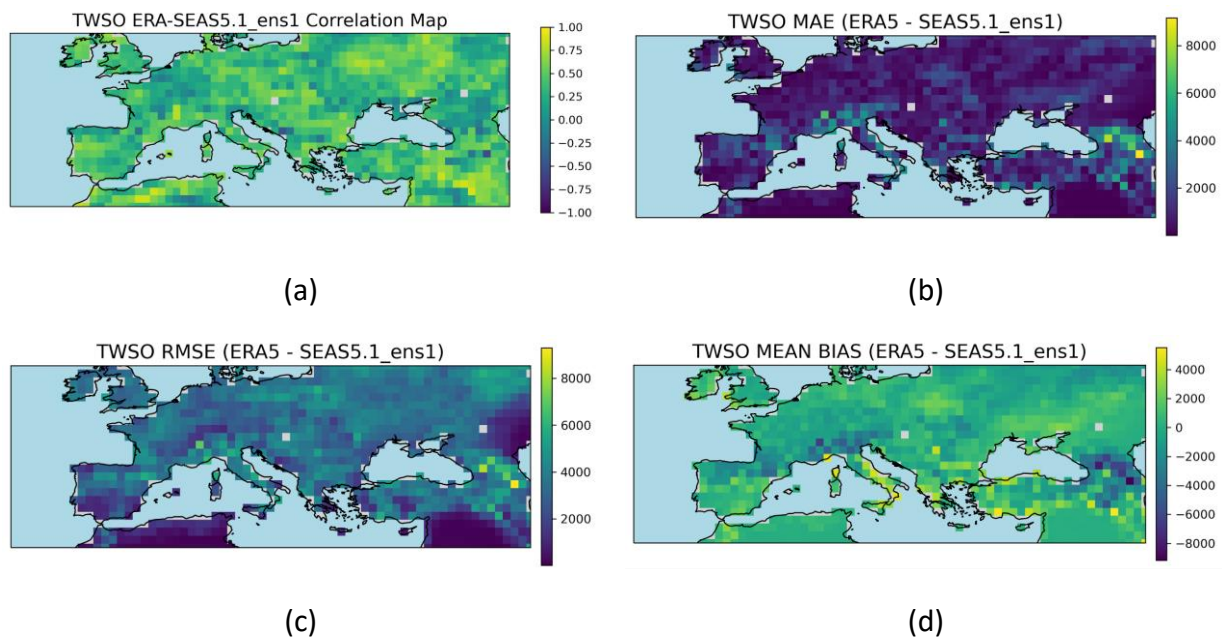


Figure 21 - Comparison of the ERA5 and SEAS5.1_ens01 TWSO simulations over the period 1993-2016: (a) Pearson correlation; (b) MAE; (c) RMSE; (d) Mean bias.

The assessment of annual metrics through the analyzed period confirms spatial variations in model performance, with reduced error over the Mediterranean and larger discrepancies in northern and eastern parts of the continent. In terms of mean bias, yield simulations exhibit persistent (almost systematic) overestimation in some northern and central regions, whereas the Mediterranean generally experiences a more mixed, occasionally negative, bias.

The combined analysis of MAE, RMSE, and mean bias suggests that yield estimation discrepancies are greatest and most consistently noticeable across Northern and Eastern Europe, with Central Europe exhibiting intermediate, variable patterns. Conversely, the Mediterranean region generally shows lower deviations.

5.1.2. CMIP6

In respect to the CMIP6 models, we ran the surrogate model and generated yield outputs for the historical simulations shown in Figure 22. The considered period is from 1993 to 2014, which is the same period used for SEAS5.1. This is a reasonable period for the monitoring of agricultural activity, since it is not affected by extreme underlying diversity in practices and production schemes. In *Figure 22* and *Figure 23* we present sample metrics.

As an example, *Figure 22a* shows the PDF of ERA5-forced and GFDL-ESM4-forced yields. The distributions of TWSO values match reasonably, except for the highest-ranking yields. The ERA5 distribution peaks at lower yields (around 0–3000 kg/ha), whereas the GFDL-ESM4 model strongly peaks at high yields (around 12,000–14,000 kg/ha), indicating a tendency towards higher yield estimations compared to ERA5 reanalysis.

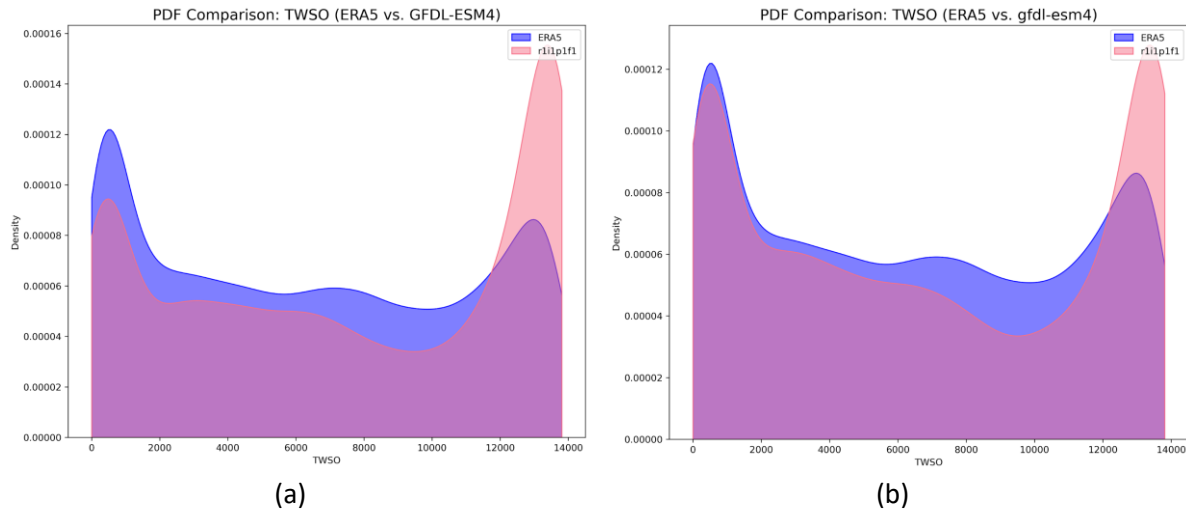


Figure 22 - Probability density functions (PDFs) of TWSO (kg/ha) AI-based surrogate simulations forced with ERA5 and (a) raw GFDL-ESM4 (r1i1p1f1) / (b) bias-adjusted GFDL-ESM4 (r1i1p1f1) for 1993-2014.

The performance in terms of PDFs similarity is evidently better in the bias adjusted version of that CMIP6 model than in the raw data as seen in *Figure 22b*.

The panels of *Figure 23* illustrate the performance metrics of GFDL-ESM4 with respect to ERA5-forced TWSOs. High RMSE and MAE values consistently appear over the Mediterranean and certain eastern European regions. These errors correspond to negative mean biases, especially in southern and southeastern areas, revealing an overestimation of yield by the CMIP6-forced yields. Conversely, northern and central European areas show comparatively small differences, although localized discrepancies do appear. The left panels of *Figure 23* shows the raw (non-bias corrected) error maps, which show comparatively larger departures than the bias adjusted CMIP6-forced TWSO outputs as depicted in the right panels.

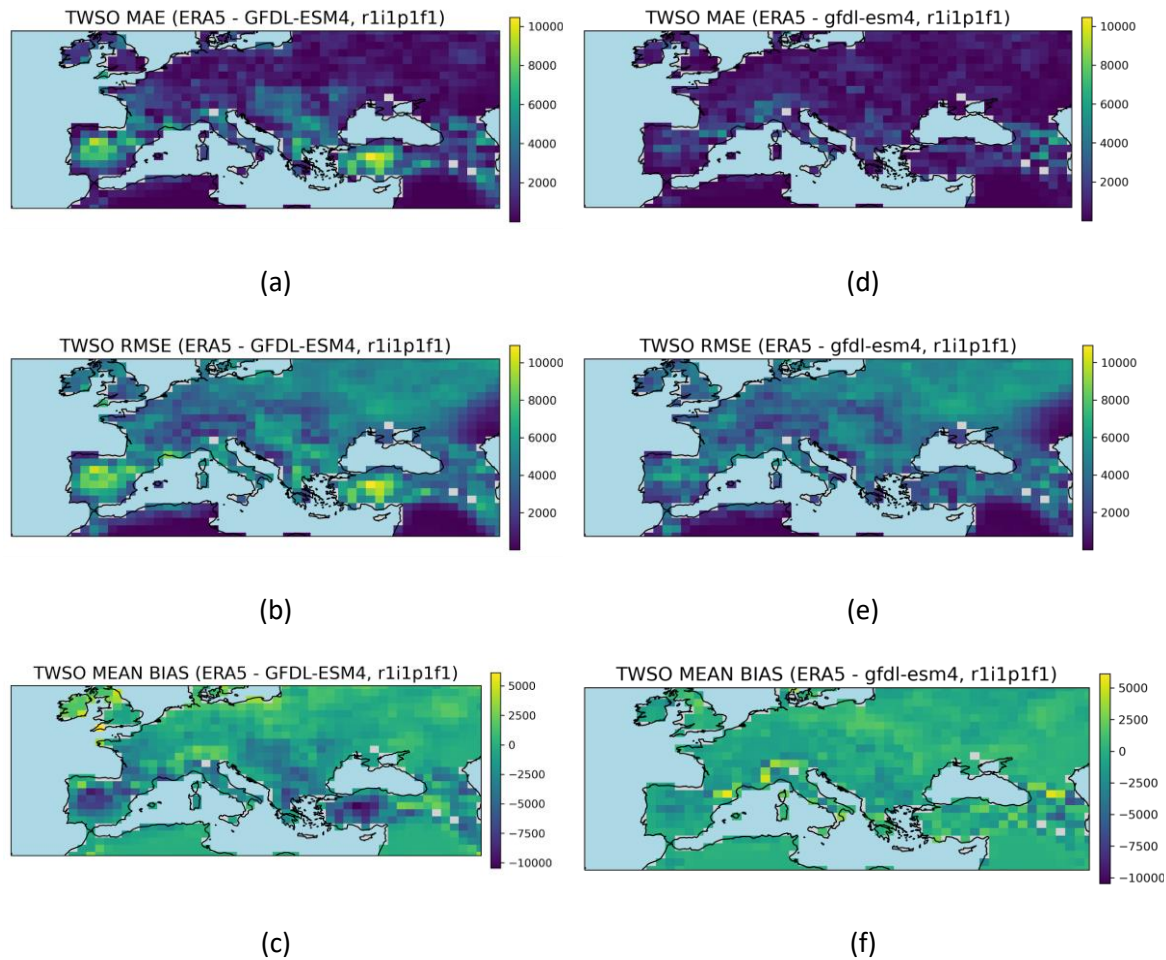


Figure 23 - Comparison of the ERA5 and raw GFDL-ESM4 (r1i1p1f1) TWSO simulations over the period 1993-2014: (a) MAE; (b) RMSE; (c) Mean bias. Comparison of the ERA5 and bias adjusted GFDL-ESM4 (r1i1p1f1) TWSO simulations over the period 1993-2014: (d) MAE; (e) RMSE; (f) Mean bias.

5.2. Areas of Concern

5.2.1. SEAS5.1 – probabilistic AOC

Detecting the AoC for the surrogate SEAS5.1-forced yields is based on a probabilistic pipeline. The full ensemble of reforecasts (1993 – 2016) is used as the reference yields. First, the procedure computes the mean over time and the ensemble members of TWSO from this reference dataset. Deviations are then defined as relative anomalies, as a percentage change of the reference value, with only negative anomalies retained. Dynamic thresholds are derived by computing the lower percentile (0.33 quantile) and the upper percentile (0.66 quantile) of the relative anomalies over the reference period and all ensemble members. The relative anomaly of the forecast (2017 – 2023) is then calculated using the same reference period, and again only negative anomalies are selected.

For each ensemble member in the forecast, the anomaly is classified into one of three categories: *above-normal* if the forecast anomaly is greater than the upper tercile, *normal* if it falls between the lower and upper terciles, and *below-normal* if it is less than the lower tercile. The percentage of ensemble members in each category is then calculated to determine the associated probabilities.

Finally, we develop decision rules to be applied with the aim of assigning a “most probable forecast” category for each grid cell and forecast time. These rules are: (i) *inconclusive*, when the probabilities of below-normal and above-normal are equal, (ii) *above-normal*, when the probability above-normal exceeds the rest, (iii) *normal to above-normal*, when the probabilities of above-normal and normal are equal, (iv) *normal*, when the probability of normal exceeds the others, and (v) *below-normal*, when the probability below-normal exceeds both above-normal and normal.

The results in terms of annual AoC forecasts for each year of the 2017-2023 period are shown in *Figure 24*. During the period 2017–2023, Europe experienced substantial variability in agricultural conditions linked to climatic patterns, as depicted by the probabilistic Areas of Concern derived from seasonal forecasts. In 2017, southern European regions, notably the Iberian Peninsula and Italy exhibited predominantly below-normal conditions, contrasting with central and northern Europe's mixed patterns. The year 2018 exhibits widespread below-normal conditions across central and northern Europe, including France, Germany, and much of eastern Europe, suggesting significant agricultural stress, while parts of southern Europe, notably the Iberian Peninsula, experienced more favorable conditions. In 2019, below-normal conditions intensified, particularly affecting the Iberian Peninsula, north-eastern Europe and central Europe. The year 2020 presented a fragmented distribution of impacts, with persistent below-normal conditions in southern Spain and southeastern regions such as Greece, whereas central Europe displayed a mix of normal, above-normal, and inconclusive categories, emphasizing regional variability and uncertainty. By 2021, below-normal conditions expanded across eastern Europe, particularly impacting the Balkans, while western Europe, including western France and the Iberian Peninsula, showed notable improvements. The year 2022 again stood out, with extensive below-normal conditions dominating much of western, central, and southeastern Europe, including Spain, France, Germany, Italy, and Greece, signaling widespread agricultural stress. Finally, 2023 is characterized by highly fragmented conditions and spatial variability, with stress conditions affecting mainly central Europe. Overall, from 2017 to 2023, Europe’s agricultural regions encountered substantial climatic variability, with pronounced episodes of widespread agricultural stress, particularly in 2018 and 2022. Southern Europe consistently exhibited greater vulnerability, emphasizing the necessity of targeted agricultural management and preparedness strategies in these regions.

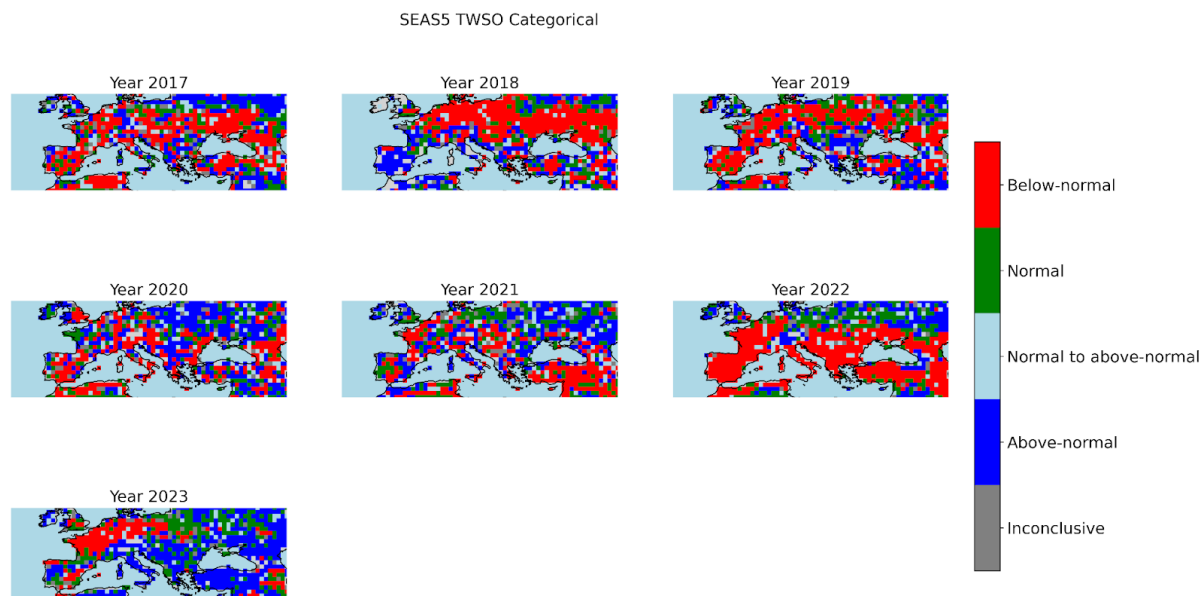


Figure 24 - Probabilistic Areas of Concern for each year of the 2017-2023 period based on 51 ensemble members of SEAS5.1 and the surrogate TWSO simulations.

5.2.2. CMIP6

For the CMIP6 models, we performed the AoC analysis for different realizations of two selected Shared Socioeconomic Pathways (SSP), namely SSP3-7.0 and SSP5-8.5 as shown in *Table 7*. Model runs are required to extend up to year 2050 and have a corresponding historical realization which is used to generate the reference yield dataset. Current national pledges under the Paris Agreement aim to limit warming to below 2 °C, ideally 1.5 °C. Even with full implementation of current policies, the global temperature is likely headed towards 2.4–2.8 °C by 2100, which is closer to SSP3-7.0 than SSP5-8.5 (Intergovernmental Panel on Climate Change, 2023; Climate Action Tracker: 2024 warming projection update).

Table 7 - CMIP6 models historical simulations used to simulate TWSO with the surrogate model for AoC generation, 1993-2014.

Model name	SSP3-7.0	SSP5-8.5
CNRM-CM6-1-HR	-	'r1i1p1f2'
EC-Earth3	'r11i1p1f1', 'r13i1p1f1', 'r15i1p1f1', 'r1i1p1f1', 'r6i1p1f1', 'r9i1p1f1'	'r11i1p1f1', 'r13i1p1f1', 'r15i1p1f1', 'r1i1p1f1', 'r3i1p1f1', 'r4i1p1f1', 'r6i1p1f1', 'r9i1p1f1'
GFDL-ESM4	'r1i1p1f1'	'r1i1p1f1'
HadGEM3-GC31-MM	-	'r1i1p1f3'
MPI-ESM1-2-HR	'r10i1p1f1', 'r1i1p1f1', 'r2i1p1f1', 'r3i1p1f1', 'r4i1p1f1', 'r5i1p1f1', 'r6i1p1f1', 'r7i1p1f1', 'r8i1p1f1', 'r9i1p1f1'	'r1i1p1f1', 'r2i1p1f1'
NorESM2-MM	'r1i1p1f1'	'r1i1p1f1'

An example of 10-yearly averages of AoC for GFDL's ESM for SSP5-8.5 (*Figure 25*) illustrates the concept of the AoC service based on the surrogate model. One strong model assumption for ECroPS and the surrogate model is the lack of any factor other than the climatological information provided by the ESM in any part of the process, such as land cover and land use, irrigation, adaptation/mitigation measures. This assumption is also the rationale behind limiting the yield projections and AoC to the year 2050.

During the first decade (2015–2024), areas of agricultural concern identified by the raw projections are predominantly limited to the southern Mediterranean region, notably southern Spain, Italy and Greece, while central and northern Europe remain largely stable. Conversely, bias-corrected projections reveal a significantly broader extent of stress, affecting also central Europe, particularly France, and eastern Europe.

In the following decade (2025–2034), non-bias corrected data indicate a moderate northward expansion of impacted areas, especially across southern and southeastern Europe, with growing intensity in Spain, southern Italy, and Greece. Scattered stress emerges in central Europe. The bias-corrected scenario, however, demonstrates a more pronounced geographic expansion of concern, highlighting extensive impact across central and eastern Europe, including major agricultural regions in France, Ukraine, and the Balkans.

By the period 2035–2044, non-bias corrected scenarios show similar climate-induced agricultural stress as the previous decade, extending into France, central Italy, and further into eastern Europe. The bias-corrected projections exhibit an even more severe and widespread distribution of stress, significantly affecting nearly all southern European regions, including intensified impacts in Spain, France, Italy and eastern European countries.

For the final years (2045–2050), non-bias corrected results depict even more severe climatic concerns across the Mediterranean basin, southern, and eastern Europe, with further expansion into France and central Europe. The bias-corrected scenario projects an even greater severity and geographical coverage, with almost all agricultural areas in southern, central, and eastern Europe under significant climatic stress, including previously less affected northern regions.

Overall, the bias-corrected scenarios consistently illustrate broader, and more intense agricultural vulnerability than the non-bias corrected projections. The bias correction clearly underscores that climate risks to agricultural productivity, particularly in southern and eastern Europe, could be significantly underestimated without such adjustments, highlighting the importance of incorporating bias adjustment for impact assessments.

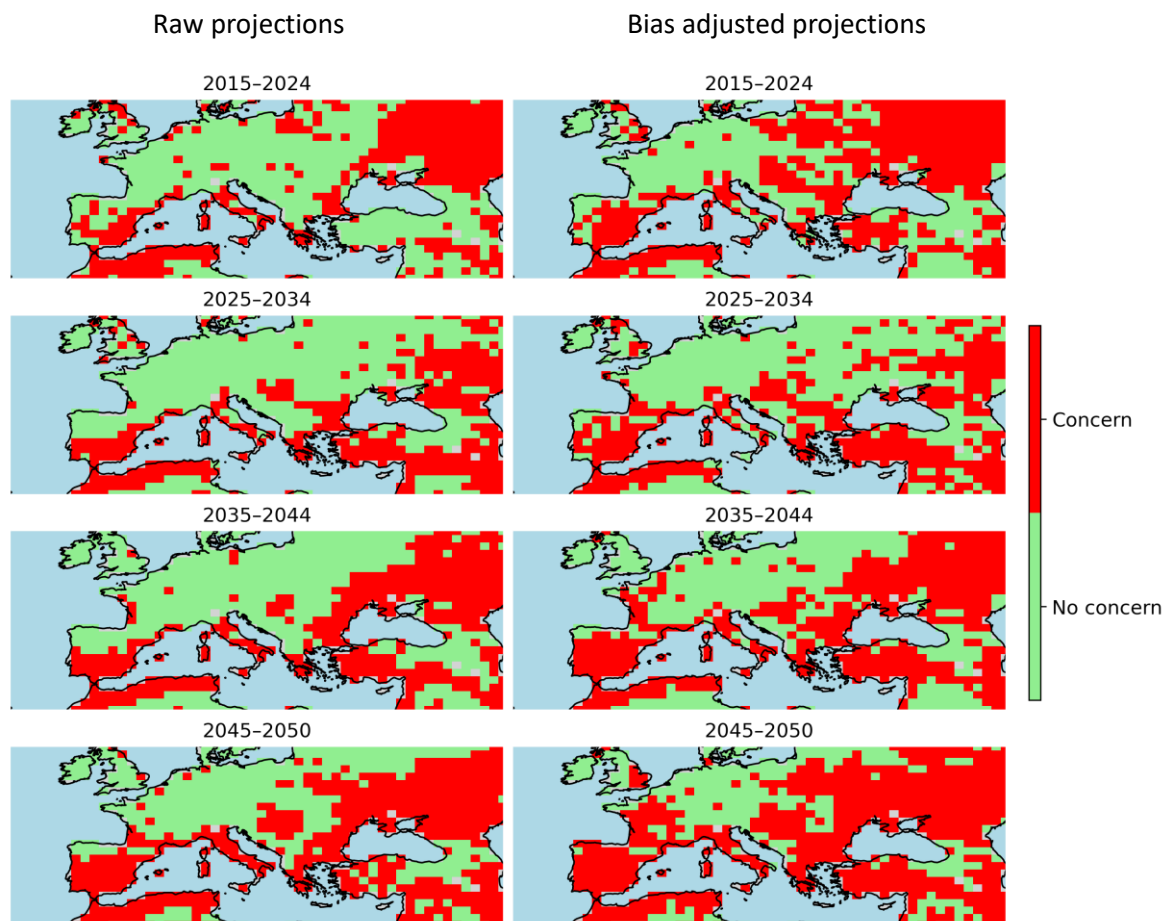


Figure 25 - 10-yearly average AoC maps for GFDL-ESM4 under SSP5-8.5 based on the surrogate model; Left: raw projections; Right: bias adjusted projections.

6. Summary and conclusions

This deliverable demonstrates how AI/ML significantly enhance CSs by addressing core challenges ranging from hydrological patterns reconstruction to the development of surrogate models towards trustworthy, fast and economically viable CSs products. Sector-specific applications such as multi-basin LSTM streamflow post-processing in the water sector, ML-based hydropower projections in the energy sector, and RNN surrogate models for crop yield in the food sector—highlight how AI improves accuracy, scalability, and operational relevance across diverse contexts. These innovations support climate services along the full temporal spectrum: real-time monitoring, seasonal forecasting, and long-term scenario-based projections. Beyond individual sectors, the integration of AI across the water-energy-food nexus offers a powerful pathway to address compound risks and cascading impacts, addressing the competing nature of those sectors, aligning with the overarching goals of CSs to transform complex climate data into actionable, sector-specific insights, and thus providing science informed policy information.

6.1. AI-enhanced climate services for the water sector

In this work, we develop an AI-enhanced post-processing framework using a multi-basin LSTM model to improve streamflow predictions from the process-based E-HYPE model. By learning from residual errors, the AI model refines streamflow estimates across both gauged and ungauged catchments in the pan-European region. The results demonstrate that AI post-processing significantly reduces biases, particularly for high flows (Q90), where the raw E-HYPE model tends to underestimate runoff. The AI-enhanced model provides a more detailed and spatially heterogeneous representation of streamflow, particularly in hydrologically complex regions such as mountainous and coastal areas. The robustness of these improvements is confirmed through temporal and spatial validation, indicating that AI-enhancement increases predictive accuracy. Overall, this approach can be applied to other ungauged basins and provides AI-enhanced streamflow simulations in these ungauged basins, making hydrological assessments more reliable and spatially refined.

6.2. AI-enhanced climate services for the energy sector

The analysis underscores that hydropower is not only central to Europe's clean energy transition but also highly sensitive to climate-induced inflow variability. To ensure long-term system reliability, especially in regions exposed to pronounced extremes, investments in adaptive infrastructure, flexible storage, and regional cooperation will be essential. The divergence across RCP scenarios highlights the need for robust planning approaches that integrate uncertainty, ensuring that power systems can respond effectively to a changing climate while advancing toward net-zero targets.

- Regions such as Austria, Finland, and Germany generally experience moderate shifts in inflow and ROR generation patterns.
- Southern European countries (e.g., greater Iberian Peninsula including Portugal, Spain and France) and some Eastern European countries (e.g., Romania, Croatia) face higher variability and more pronounced extremes, necessitating further adaptation of flexibility options in power generation sector.
- The notable divergence in projections between different RCP scenarios underscores the importance of adaptive capacity, robust infrastructure, and diversified flexibility measures to maintain reliable hydropower operations in the face of climatic uncertainties.

6.3. AI-enhanced climate services for the food sector

We developed and presented a well-functioning AI surrogate crop growth model which performs in a low cost and resource effective way, emulating yield growth, from flowering to harvest time, as simulated by the ECroPS model. The output consists of grain maize simulations for ERA5, six CMIP6 ESMs and SEAS5.1 seasonal forecast system from ECMWF. We illustrate potential applications of the AI model in historical (reanalysis) and retrospective forecasts (SEAS5.1). Regarding future projections, we present and evaluate the sub-service for Areas of Concern, showing the potential of the model to estimate areas that exhibit high probability for elevated risk in comparison to past yields. The AI surrogate model can be deployed for various datasets within the European domain it was trained with and can easily be adapted by stakeholders.

Bibliography

[Chapter 2.1]

Berg, P., Almén, F., & Bozhinova, D. (2021). HydroGFD3.0 (Hydrological Global Forcing Data): A 25 km global precipitation and temperature data set updated in near-real time. *Earth System Science Data*, 13(4), 1531–1545. <https://doi.org/10.5194/essd-13-1531-2021>

Crochemore, L., Perrin, C., Andréassian, V., Ehret, U., Seibert, S. P., Grimaldi, S., Gupta, H., & Paturel, J.-E. (2015). Comparing expert judgement and numerical criteria for hydrograph evaluation. *Hydrological Sciences Journal*, 60(3), 402–423. <https://doi.org/10.1080/02626667.2014.903331>

Du, Y., Pechlivanidis, G. I. (2024) Hybrid approaches enhance hydrological model usability for local streamflow prediction, 23 October 2024, PREPRINT (Version 1) available at Research Square [<https://doi.org/10.21203/rs.3.rs-5142634/v1>]

Du, Y., Olsson, J., Isberg, K., Strömqvist, J., Hundecha, Y., Silva, B. C. da, Rafee, S. A. A., Fragoso Jr, C. R., Beldring, S., Hansen, A., Uvo, C. B., & Sørensen, J. (2024). Application-based evaluation of multi-basin hydrological models. *Journal of Hydrology*, 641, 131727. <https://doi.org/10.1016/j.jhydrol.2024.131727>

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Comput.*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>

Hundecha, Y., Arheimer, B., Donnelly, C., & Pechlivanidis, I. (2016). A regional parameter estimation scheme for a pan-European multi-basin model. *Journal of Hydrology: Regional Studies*, 6, 90–111. <https://doi.org/10.1016/j.ejrh.2016.04.002>

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, 22(11), 6005–6022. <https://doi.org/10.5194/hess-22-6005-2018>

Lees, T., Reece, S., Kratzert, F., Klotz, D., Gauch, M., De Bruijn, J., Kumar Sahu, R., Greve, P., Slater, L., & Dadson, S. J. (2022). Hydrological concept formation inside long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, 26(12), 3079–3101. <https://doi.org/10.5194/hess-26-3079-2022>

Li, D., Marshall, L., Liang, Z., Sharma, A., & Zhou, Y. (2021). Characterizing distributed hydrological model residual errors using a probabilistic long short-term memory network. *Journal of Hydrology*, 603, 126888. <https://doi.org/10.1016/j.jhydrol.2021.126888>

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, 10(3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)

Pechlivanidis, I. G., Crochemore, L., Rosberg, J., & Bosshard, T. (2020). What Are the Key Drivers Controlling the Quality of Seasonal Streamflow Forecasts? *Water Resources Research*, 56(6), e2019WR026987. <https://doi.org/10.1029/2019WR026987>

Willmott, C., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, 79–82. <https://doi.org/10.3354/cr030079>

[Chapter 2.3]

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., ... Zheng, X. (2016). TensorFlow: A system for large-scale machine learning. *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation*, 265–283.

Ceglar, A., van der Wijngaart, R., de Wit, A., Lecerf, R., Boogaard, H., Seguini, L., van den Berg, M., Toreti, A., Zampieri, M., Fumagalli, D. & Baruth, B. (2019). Improving WOFOST model to simulate winter wheat phenology in Europe: Evaluation and effects on yield. *Agricultural Systems*, 168, 168–180, <https://doi.org/10.1016/j.agry.2018.05.002>

Chollet, F. (2015). Keras. Seattle, WA, USA.

Chung, J., Ahn, S., & Bengio, Y. (2017). *Hierarchical Multiscale Recurrent Neural Networks* (arXiv:1609.01704). arXiv. <https://doi.org/10.48550/arXiv.1609.01704>

de Wit, A., Boogaard, H., Fumagalli, D., Janssen, S., Knapen, R., van Kraalingen, D., Supit, I., van der Wijngaart, R., & van Diepen, K. (2019). 25 years of the WOFOST cropping systems model. *Agricultural Systems*, 168, 154–167. <https://doi.org/10.1016/j.agry.2018.06.018>

Eurostat, Agricultural production—Crops, Retrieved April 15, 2025, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Agricultural_production_-_crops

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., ... Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>

Huber, P. J. (1964). Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, 35(1), 73–101. <https://doi.org/10.1214/aoms/1177703732>

Kingma, D. P., & Ba, J. (2017). *Adam: A Method for Stochastic Optimization* (arXiv:1412.6980). arXiv. <https://doi.org/10.48550/arXiv.1412.6980>

Toreti, A., Bassu, S., Asseng, S., Zampieri, M., Ceglar, A. & Royo, C. (2022). Climate service driven adaptation may alleviate the impacts of climate change in agriculture. *Communications Biology* volume 5, 1235. <https://doi.org/10.1038/s42003-022-04189-9>

[Chapter 4]

European Commission. (2019). *The European Green Deal* (COM(2019) 640 final). Brussels. Retrieved from <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52019DC0640>

Andresen, G. B., Gøtske, E. K., Neumann, F., & Victoria, M. (2024). Designing a sector-coupled European energy system robust to 60 years of historical weather data. *Nature Communications*, 12.

JRC. (2020). Power system flexibility in a variable climate. Joint Research Centre.

L. van der Most, K. v.-L. (2024). Temporally compounding energy droughts in European electricity systems with hydropower. *Nature Energy*(9), 1474–1484.

[Chapter 5]

Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate*, 28(17), 6938–6959. <https://doi.org/10.1175/JCLI-D-14-00754.1>

Climate Action Tracker: 2024 warming projection update. (n.d.). Climate Analytics. Retrieved March 17, 2025, from <https://climateanalytics.org/publications/cat-global-update-as-the-climate-crisis-worsens-the-warming-outlook-stagnates>

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L.-P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., ... Zhang, Q. (2022). The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6. *Geoscientific Model Development*, 15(7), 2973–3020. <https://doi.org/10.5194/gmd-15-2973-2022>

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., ... Zhao, M. (2020). The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics. *Journal of Advances in Modeling Earth Systems*, 12(11), e2019MS002015. <https://doi.org/10.1029/2019MS002015>

Intergovernmental Panel On Climate Change (Ippc). (2023). *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781009157896>

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremmer, D., Weisheimer, A., Balsamo, G., Keeley, S. P. E., Mogensen, K., Zuo, H., & Monge-Sanz, B. M. (2019). SEAS5: The new ECMWF seasonal forecast system. *Geoscientific Model Development*, 12(3), 1087–1117. <https://doi.org/10.5194/gmd-12-1087-2019>

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblueh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., ... Marotzke, J. (2018). A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). *Journal of Advances in Modeling Earth Systems*, 10(7), 1383–1413. <https://doi.org/10.1029/2017MS001217>

Ridley, J., Menary, M., Kuhlbrodt, T., Andrews, M., & Andrews, T. (2019). *MOHC HadGEM3-GC31-MM model output prepared for CMIP6 CMIP* [Dataset]. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.420>

Ridoutt, B. G., Pfister, S., Manzardo, A., Bare, J., Boulay, A.-M., Cherubini, F., Fantke, P., Frischknecht, R., Hauschild, M., Henderson, A., Jolliet, O., Levasseur, A., Margni, M., McKone, T., Michelsen, O., Milà i Canals, L., Page, G., Pant, R., Raugei, M., ... Veronesi, F. (2016). Area of concern: A new paradigm in life cycle assessment for the development of footprint metrics. *The International Journal of Life Cycle Assessment*, 21(2), 276–280. <https://doi.org/10.1007/s11367-015-1011-7>

Seguini, L., Bussay, A., & Baruth, B. (2019). From extreme weather to impacts: The role of the areas of concern maps in the JRC MARS bulletin. *Agricultural Systems*, 168, 213–223. <https://doi.org/10.1016/j.agry.2018.07.003>

Seland, Ø., Bentsen, M., Olivié, D., Toniazzi, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., ... Schulz, M. (2020). Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geoscientific Model Development*, 13(12), 6165–6200. <https://doi.org/10.5194/gmd-13-6165-2020>

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., Colin, J., Guérémy, J.-F., Michou, M., Moine, M.-P., Nabat, P., Roehrig, R., Salas y Méliá, D., Séférián, R., Valcke, S., Beau, I., Belamari, S., Berthet, S., Cassou, C., ... Waldman, R. (2019). Evaluation of CMIP6 DECK Experiments With CNRM-CM6-1. *Journal of Advances in Modeling Earth Systems*, 11(7), 2177–2213. <https://doi.org/10.1029/2019MS001683>