https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

The Destination Earth digital twin for climate change adaptation

Francisco J. Doblas-Reyes[1,2], Jenni Kontkanen[3], Irina Sandu[13], Mario Acosta[2], Mohammed Hussam Al Turjmam[4], Ivan Alsina-Ferrer[2], Miguel Andrés-Martínez[4], Leo Arriola[2], Marvin Axness[2], Marc Batlle Martín[2], Peter Bauer[5], Tobias Becker[6], Daniel Beltrán[2], Sebastian Beyer[4], Hendryk Bockelmann[7], Pierre-Antoine Bretonnière[2], Sebastien Cabaniols[8], Silvia Caprioli[9], Miguel Castrillo[2], Aparna Chandrasekar[10], Suvarchal Cheedela[4], Victor Correal[2], Emanuele Danovaro[6], Paolo Davini[11], Jussi Enkovaara[3], Claudia Frauen[7], Barbara Früh[12], Aina Gaya Àvila[2], Paolo Ghinassi[11], Rohit Ghosh[4], Supriyo Ghosh[2], Iker González[2], Katherine Grayson[2], Matthew Griffith[13], Ioan Hadade[6], Christopher Haine[8], Carl Hartick[12], Utz-Uwe Haus[8], Shane Hearne[2], Heikki Järvinen[14], Bernat Jiménez[2,a], Amal John[4], Marlin Juchem[12,b], Thomas Jung[4], Jessica Kegel[4], Matthias Kelbling[10], Kai Keller[2], Bruno Kinoshita[2], Theresa Kiszler[3], Daniel Klocke[5], Lukas Kluft[5], Nikolay Koldunov[4], Tobias Kölling[5], Joonas Kolstela[15], Luis Kornblueh[5], Sergey Kosukhin[5], Aleksander Lacima-Nadolnik[2], Jeisson Javier Leal Rojas[10], Jonni Lehtiranta[15], Tuomas Lunttila[3], Anna Luoma[3], Pekka Manninen[3], Alexey Medvedev[2], Sebastian Milinski[6], Ali Omar Abdelazim Mohammed[8], Sebastian Müller[10], Devaraju Naryanappa[3], Natalia Nazarova[9], Sami Niemelä[15], Bimochan Niraula[4,c], Henrik Nortamo[3], Aleksi Nummelin[15], Matteo Nurisso[11], Pablo Ortega[2], Stella Paronuzzi[2,f], Xabier Pedruzo-Bagazgoitia[6], Charles Pelletier[6], Carlos Peña[2], Suraj Polade[15], Himansu Pradhan[4,c], Rommel Quintanilla[2], Tiago Quintino[13], Thomas Rackow[6], Jouni Räisänen[14], Maqsood Mubarak Rajput[4], René Redler[5], Balthasar Reuter[6], Nuno Rocha Monteiro[2], Francesc Roura-Adserias[2], Silva Ruppert[4], Susan Sayed[12], Reiner Schnur[5], Tanvi Sharma[4,d], Dmitry Sidorenko[4], Outi Sievi-Korte[3], Albert Soret[2], Christian Steger[12], Bjorn Stevens[5], Jan Streffing[4], Jaleena Sunny[4,e], Luiggi Tenorio[2], Stephan Thober[10], Ulf Tigerstedt[3], Oriol Tinto[2], Juha Tonttila[3], Heikki Tuomenvirta[15], Lauri Tuppi[14], Ginka Van Thielen[2], Emanuele Vitali[3], Jost von Hardenberg[9], Ingo Wagner[2], Nils Wedi[13], Jan Wehner[4], Sven Willner[5], Xavier Yepes-Arbós[2], Florian Ziemen[7], and Janos Zimmermann[7]

1 ICREA, Barcelona, Spain
2 Barcelona Supercomputing Center (BSC), Barcelona, Spain
3 CSC – IT Center for Science, Espoo, Finland
4 Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung (AWI), Bremerhaven, Germany

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

5 Max Planck Institute for Meteorology, Hamburg, Germany
6 European Centre for Medium-Range Weather Forecasts, Bonn, Germany
7 German Climate Computing Center (DKRZ), Hamburg, Germany
8 Hewlett Packard France SAS, Puteaux, France
9 Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Turin, Italy
10 Helmholtz Centre for Environmental Research, Leipzig, Germany
11 Consiglio Nazionale delle Ricerche, Istituto di Scienze dell'Atmosfera e del Clima (CNR-ISAC), Torino, Italy
12 Deutscher Wetterdienst, Offenbach, Germany
13 European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK
14 Institute for Atmospheric and Earth System Research, University of Helsinki, Finland
15 Finnish Meteorological Institute, Helsinki, Finland

a now at: Instituto de Geociencias (IGEO), CSIC-UCM, Madrid, Spain
b now at: German Aerospace Center (DLR), Oberpfaffenhofen, Germany
c now at: German Climate Computing Center (DKRZ), Hamburg, Germany
d now at: Siemens AG, Berlin, Germany
e now at: Swiss Seismological Service (SED) – ETH Zurich, Zurich, Switzerland
f now at: Mercator Ocean International (MOi), Toulouse, France

Correspondence to: Francisco J. Doblas-Reyes (francisco.doblas-reyes@bsc.es)

Abstract. The Climate Change Adaptation Digital Twin (Climate DT), developed as part of the European Commission's Destination Earth (DestinE) initiative, is a pioneering effort to build an operational climate information system in support of adaptation. This system produces global climate simulations with local granularity, providing information at scales that matter for decision-making. The Climate DT delivers multi-decadal climate simulations at spatial resolutions of 5–10 km, with hourly outputs, offering globally consistent, frequently updated data. The km-scale simulations address some limitations of current climate models, improving local granularity and reducing longstanding biases, supporting more equitable (understood as accessible and relevant across regions) and credible climate information. The Climate DT is built on cutting-edge infrastructure, expert collaboration, and digital innovation. It supports real-time, on-demand responses to policy questions, with quantified uncertainty. It fosters interactivity by allowing users to influence simulation design, model outputs, and applications through co-design. AI-based tools, including emulators and chatbots, are being developed to enhance flexible scenario exploration and ease climate information access. Sector-specific applications are embedded in the system to generate tailored climate-impact indicators, with examples for energy, water, and forest management. The applications have been co-designed with informed users. An important innovation is the use of high-resolution storylines. These are physically consistent simulations of extreme events under different climate conditions that provide contextual insights to support concrete adaptation decisions. A unified workflow across platforms orchestrates all components ensuring automation, containerisation for portability, and traceability. The unified data management ensures consistency and eases the usability of the data. Data is delivered using standard grids (HEALPix) at high-frequency (hourly) and follows a strict governance policy. Streaming enables real-time data use and unlocks access to the unprecedented data produced by the high-resolution simulations. Monitoring tools provide real-time quality control of both data and models and provide diagnostics during the Climate DT operation. The compute-intensive system is powered by world-class supercomputing capabilities

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

through a strategic partnership with the European High Performance Computing Joint Undertaking (EuroHPC). Despite high computational demands, the Climate DT sets a new benchmark for delivering equitable, credible, and actionable climate information. In this way it complements existing initiatives like CMIP, CORDEX, and national and European climate services, and aligns with global climate science goals for climate adaptation.

1 Introduction

Providing reliable climate information is essential for enabling effective climate change adaptation (Orlove, 2022). The latest report from the Intergovernmental Panel on Climate Change (IPCC) Working Group I (IPCC, 2021) highlights major gaps in our understanding and ability to simulate regional and local climate change and to deliver climate information that supports decision-making (Collins et al., 2024; Shaw et al., 2024). It also emphasises the challenge of ensuring that the information is salient[1], equitable, and credible for diverse audiences (Doblas-Reyes et al., 2021).

With globally annually averaged temperatures, for the first time, exceeding 1.5 ºC global warming level threshold[2,3] set in the Paris Agreement (Betts et al., 2023) and with the profound relevance of climate change impacts on both the economy and society (Kotz et al., 2022), new challenges emerge for the sources of future climate information at all the scales where the impacts of climate change are already observed and adaptation is needed. Solutions to these challenges are required to move from plausibility assessments of changes in local and regional climate (e.g., Collins et al., 2024) to comprehensive climate adaptation plans and targeted measures at both large (e.g., EUCRA, 2024) and small scales (e.g., Schubert et al., 2024). The urgency for a new approach to provide climate information is motivated by, among others, policy frameworks like the European Green Deal[4], which aims to guide Europe toward a carbon-neutral, resilient economy. For instance, a cornerstone of the Green Deal is the creation of accessible and interoperable infrastructures that facilitate climate-related decision-making for adaptation challenges.

For climate information sources to estimate future climate to be salient, equitable, and credible, these sources have to consider a wide range of spatial scales, be timely and innovative, and be co-produced with decision-makers to adequately address societally-relevant questions (e.g., Pitman et al., 2022; Doblas-Reyes et al., 2024). These requirements have unveiled an important climate-change information gap for effective climate adaptation strategies to be designed and implemented successfully:

- Equitability: Climate information for adaptation must be provided at various spatial scales, and particularly at the scales where the impacts of climate change are observed and expected, to align with the specific decision-making needs. Given the broad range of climate-sensitive sectors, climate information is required globally, although always

[1] Meaning the quality of being particularly noticeable or important.
[2] https://climate.metoffice.cloud/current_warming.html
[3] https://climate.copernicus.eu/june-2024-marks-12th-month-global-temperatures-15degc-above-pre-industrial-levels
[4] https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

matching the spatial scale of the decision (Schubert et al., 2024). Ensuring equitable access to climate information by all decision-makers (Hazeleger et al., 2024) therefore requires globally-consistent climate information sources with the highest possible resolution. Because global climate simulations typically do not provide information at resolutions tailored to support adaptation and mitigation decisions at local or regional scale, they are usually complemented with local and regional simulation efforts (e.g., Soares et al., 2024). Ensembles of regional simulations usually take time to cover every region, leave some regions out, and must be continuously updated in a laborious process, illustrating the dependencies and power relations between developed countries and regions that do not have the resources to produce simulations for their own area. The interest in more frequent updates has been formulated by some existing initiatives (EUCRA, 2024).

- Timeliness and innovation: Climate projections are currently updated as part of cycles spanning several years during which the research community undertakes model developments and coordinates climate experiments, to align with the preparation of IPCC assessment reports (e.g., Eyring et al., 2016; Jones et al., 2024). The research community regularly contributes to highly policy-relevant reports[5] as well. The data production cycles take between five and seven years to complete because the community that has been made responsible for these projections is often not structured or funded to address such requirements on a continuous and on-demand basis (Stevens, 2024), with many of those playing a role in the research domain. Moreover, there is a need to more directly address the dialogue towards policy needs by exploiting innovative technologies to make climate information more engaging, accessible, interpretable, understandable, and ultimately actionable (Jones et al., 2024). Given the rapid innovations the climate information sector is experiencing, for example in digital solutions and artificial intelligence (AI) technologies (Bauer et al., 2023), climate projections could consider, in addition to the updates provided by the Coupled Model Intercomparison Project (CMIP[6]) and the Coordinated Regional Climate Downscaling Experiment (CORDEX[7]), strategies and infrastructures to quickly and efficiently integrate these innovations.
- Co-production: Climate-related decisions are based on all available evidence from multiple disciplines. As a result, decision makers at times encounter challenges requiring climate information that cannot be foreseen a priori by physical climate scientists. Hence, the absence of regular channels for climate-sensitive sectors to both request timely data sources and influence the design of climate information hinders effective adaptation efforts. For instance, participatory processes that involve selected decision makers (Baulenas et al., 2023) in discussions about the transformation of climate data into decision-ready climate information can uncover user needs requiring modifications in how climate data is generated. Climate adaptation decision-makers continuously formulate "what-if" questions that require timely answers developed with the participation of a heterogeneous set of actors. The fact that this dialogue does not happen more often tends to be the result of the relationships established between the

[5] https://10insightsclimate.science/
[6] https://wcrp-cmip.org/
[7] https://cordex.org/

research and service communities (Rodrigues and Shepherd, 2022). The need for systematic co-production mechanism (Bojovic et al., 2021) and the difficulties experienced by the physical climate science community to react on time to a variety of decision-making requirements, in spite of the acknowledgement of this need (Jones et al., 2024), can weaken the effective provision of future climate information and limit the diversity of user contexts and values included (Kruk et al., 2017; Fiedler et al., 2021).

Current practice, represented by the international modelling exercises of CMIP and CORDEX international modelling exercises, has demonstrated to be both useful and relevant for the above-mentioned purposes. However, some aspects have not yet been fulfilled (Jakob et al., 2023; Stevens, 2024). The CMIP community has recently undertaken efforts to consider some of them in their next phase (Dunne et al., 2025), although some challenges remain to respond to non-research needs. In this paper we present an approach to generate climate data for the future that is complementary to current practice of climate projection production and delivery. This approach aims to bridge the gap between timely decision-relevant climate information and current climate modelling practices (Hewitt and Stone, 2021). It involves, among other things, operationalising[8] climate-projection production and delivery within a framework that fosters continuous interaction between data producers and "data consumers" following an inclusive pathway that includes mutual recognition. A data consumer is any "application" using the climate model data, be it for scientific evaluation and understanding, climate impact and risk assessment, or for policy-making purposes.

A promising way to design and implement the interactions between data producers and data consumers is to co-develop the required systems using the digital twin concept (Wright and Davidson, 2020). Digital twins are replicas of physical assets, processes, and systems using a highly interconnected workflow, including the possibility for interaction and, when relevant, a link to the physical reality. This paper describes the characteristics and novelties of a digital twin for climate change adaptation (Climate DT henceforth) implemented in the framework of the Destination Earth[9] initiative of the European Union (DestinE henceforth; Sandu, 2024; Hoffmann et al., 2023; Wedi et al., 2025). The paper summarises the achievements of the Climate DT so far and describes its capabilities to address the salience, credibility, and equity principles of climate information for adaptation.

2 The Climate DT concept

A digital twin of the climate system targeting adaptation is expected to make use of observations, integrate several climate models to consider uncertainty sources, include applications for climate-sensitive sectors directly connected to the climate models, and provide interfaces to configure the simulations, output, and data consumers. The climate models embedded in the digital twin could be either process-based, to explicitly represent the essential physical processes, or based on AI with the

[8] Operational is understood as a set of processes delivering timely and application-ready data, following quick updating cycles, continuous quality monitoring, using a DevOps methodology, and generating machine learning ready datasets.
[9] https://destination-earth.eu/

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

capacity to be explainable. For the digital twin to be usable and useful, its components, including the climate models and decision-oriented applications, should be co-designed and offer full traceability. The resulting output must be of sufficient quality (Dee et al., 2024) to support well-informed decisions while the results should be available fast enough for the decisions to be made within acceptable time scales (Wright and Davidson, 2020). To build trust in the digital twin, documentation, verification (full traceability and testing protocols), and validation (comparison with reality whenever possible) procedures must be included. The existence of uncertainty means that validation needs to be treated as a statistical process, which in climate is traditionally addressed using multi-model and multi-member ensembles of solutions (Bauer et al., 2021b; Jones et al., 2024) for a number of emission scenarios. Additionally, a climate digital twin should make use, in a structured, interactive, and iterative manner, of the instruments developed by social sciences for climate adaptation (Tao and Qi, 2019).

This is the concept that inspires the Climate DT (Figure 1). DestinE is a European Union funded initiative launched in 2022, with the aim to build digital replicas of the Earth system by 2030. The initiative is being jointly implemented, under the lead of the European Commission Directorate General CNECT, by three entrusted entities: the European Centre for Medium-Range Weather Forecasts (ECMWF), the European Space Agency (ESA), and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) together with over 100 partner organisations in Europe. The Climate DT is being developed since September 2022 by a strong partnership led by the CSC-IT Center for Science including leading climate, weather and supercomputing centres and academic institutions from six European countries[10], in close collaboration with ECMWF, who is responsible for the implementation of the digital twins. The necessary computational resources are provided through a special access call by the European High-Performance Computing Joint Undertaking (EuroHPC JU[11]), which oversees the implementation of the network of European pre- and exascale computing infrastructures.

The Climate DT consists of a number of components integrated into a workflow using a co-design approach to deliver information. It uses three global kilometre-scale (km-scale henceforth) climate models (ICON, Hohenegger et al., 2022; IFS-NEMO and IFS-FESOM, Rackow et al., 2025) that explicitly represent essential physical processes that critically influence the evolution of the climate system. These global models are used to produce multidecadal simulations (from 1990 to 2040 currently) at 5 to 10 km resolutions as well as simulations that allow exploring how recent extreme events would unfold in different climate conditions (John et al., 2025). This leads to the production of global information with local granularity.

The climate models are embedded in an infrastructure built following practices recommended by both the climate services and digital technology communities to ensure the provision of operational (understood as timely and routine production)

[10] The Climate DT team includes personnel from Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research (AWI), Barcelona Supercomputing Center (BSC), Max Planck Institute for Meteorology (MPI-M), Institute of Atmospheric Sciences and Climate (CNR-ISAC), German Climate Computing Centre (DKRZ), National Meteorological Service of Germany (DWD), Finnish Meteorological Institute (FMI), Hewlett Packard Enterprise (HPE), Polytechnic University of Turin (POLITO), Helmholtz Centre for Environmental Research (UFZ), and University of Helsinki (UH)

[11] https://eurohpc-ju.europa.eu/index_en

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

climate projections. The backbone of the infrastructure is a unified workflow that offers full traceability and permits a unified data handling procedure with common variables, grids, and formats, but responding to continuously collected user requirements. The data handling deals efficiently with unprecedented data volumes due to the choice of high spatial resolution to provide climate information at scales required for adaptation decision-making. Efficient data handling is supported by the introduction of a data streaming strategy into the climate information production process. Both the unified workflow and efficient data handling, which work across multiple high-performance computing (HPC) environments and do not require additional development efforts by both climate modellers and climate applications, are fundamental Climate DT innovations.

The Climate DT sets up an operational framework for the production and tailoring of climate information sources for both recent and future climate according to scenarios and priorities set by data consumers, who develop information to be used in the design adaptation strategies. Current climate information sources for future climate projections cannot be considered operational, although some important evolution is taking place (Dunne et al., 2024). For instance, current efforts have emerged mainly, but not only, from the research community, with its own international governance, and often relying on research funding to address policy-driven requirements[12]. A protocol was developed for the Climate DT operationalisation that considers different workflow definitions and actions for

- development, where new capabilities of the Climate DT components are implemented,
- experimental, where the end-to-end workflow and dataflow are tested and fixes implemented,
- and operational, where a scheduled and robust production is ensured,

phases of the production cycle. These different workflow "suites" are inspired by common practice in operational weather and ocean prediction (e.g., Alvarez Fanjul et al., 2024).

An important characteristic of the Climate DT infrastructure is its flexibility (Wedi et al., 2022). This solution enables performing bespoke simulations to address targeted "what-if" questions. It also allows tailoring information to specific user needs, including a service that allows on-demand production in situations where more detailed and targeted climate information sources are required. A number of applications for sectors sensitive to climate variability and change have been embedded in the production workflow, computing climate-relevant indicators simultaneously with the climate model simulations. The capability of this infrastructure to include new applications in the workflow makes the Climate DT able to support fast testing of climate adaptation options. The Climate DT simulations, and the resulting impact indicators, are monitored in real time and regularly evaluated with a specially designed software framework. The results are regularly published in an internal dashboard. This is an essential requirement for a system that aspires to offer the possibility to determine its added value for users as soon as the data is produced, with the aim of better informing decision-making (Hallegate, 2009).

[12] Note that some national meteorological services have operational initiatives for the production of climate projections.

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

The global climate models used are the core, although not the only important element, of the Climate DT. Given their computational cost, efficient use of HPC resources is another defining Climate DT aspect. The climate simulations performed leverage strategic access to the most powerful European supercomputers with dedicated cloud storage and delivery systems. Additional capabilities using AI are under development to ease the exploitation of the climate information generated, enhancing the Climate DT equitability and timeliness. It is important to note that the Climate DT concept and infrastructure do not restrict themselves to the use of km-scale simulations.

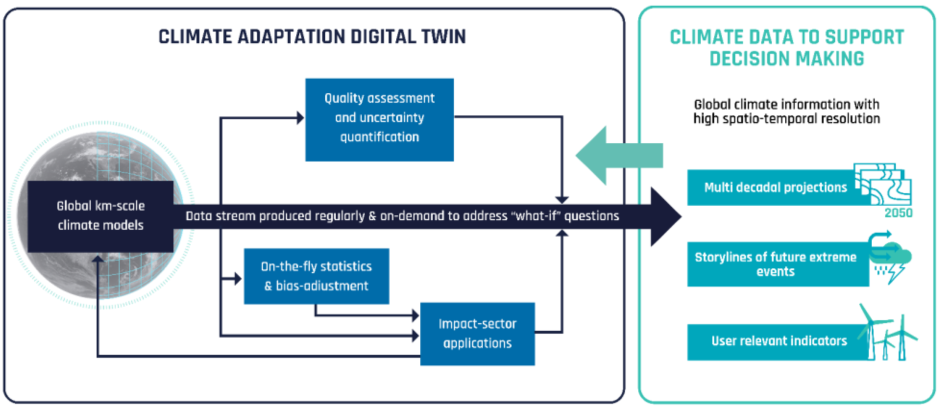

Figure 1: Conceptual diagram illustrating the main Climate DT components. The Climate DT uses data streams from the global high-resolution climate models to satisfy the requirements of a number of data consumers. The data consumers include climate-sensitive applications and a comprehensive monitoring and quality assessment procedure. User feedbacks are regularly incorporated in the Climate DT design (hence the thick arrows pointing in both directions) so that the global information with local granularity produced adequately supports climate adaptation decision-making.

3 End-to-end work and data flow

Digital twins require a multi-layered flexible software infrastructure that allows data consumers to easily access the data and interact with the system. The infrastructure should be designed to deal with models that aim to achieve the highest computing throughput possible, while taking into account end-to-end dataflows with concurrency in the execution of applications. For this approach to work, the Climate DT relies on an orchestrator that executes tasks in a traceable manner (Bauer et al., 2021b) and makes use of a unified interface where developers and operators can configure the whole digital twin production. Domain-oriented workflow managers (e.g., Uruchi et al., 2021; Leo et al., 2024) have proven very useful in the past for similar duties and are used on a regular basis in operational numerical weather prediction and climate simulations. The Climate DT uses a unified workflow approach that seamlessly integrates the three climate models and all data consumers and is controlled seamlessly with a workflow manager. The workflow solution incorporates the necessary elements to configure the Climate DT. The same software solution is employed on the various EuroHPC computing

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

platforms used by the Climate DT. This is an improvement over the current practice of engineering different solutions for each climate model and platform. Data consumers use Singularity[13] containers to simplify their portability with their deployment inside.

The integration of all the Climate DT elements in a single workflow is a fundamental novelty with respect to the current paradigm of climate models where fixed and static flows of components, experimental setups, model output, and a variety of software solutions are managed by layers of experts, often working independently from each other and in isolation from the climate-vulnerable communities. This new approach to perform climate simulations significantly improves workflow maintainability and portability, reduces the cost of maintaining a specific environment on different HPC platforms, and facilitates the introduction of user-relevant Climate DT modifications.

The Climate DT workflow (Figure 2) comprises the full production chain, from setting up and running global climate simulations to serving seamlessly their output to a range of data consumers. Such a groundbreaking holistic approach to deliver climate information includes tasks to monitor and quality control, both scientifically and in terms of data integrity, the climate simulations, and a range of selected applications that estimate user-relevant indicators for climate-sensitive sectors like renewable energy, and forest and water resources management. New data consumers can be included in the workflow as they are considered ready to join; this happens as part of thorough and regular testing, which is a principle of the Climate DT implementation. Given the complexity of the task, the development, deployment, testing, and operation of the end-to-end workflow plays a central role. The development and deployment of this end-to-end workflow on the LUMI and MareNostrum5 (MN5) pre-exascale supercomputers (both part of the EuroHPC network) constituted one of the main priorities during the first phase of DestinE (2022-2024), while in the second phase (2024-2026) the main focus has shifted towards its transitioning towards a prototype operational status and the demonstration of a routine delivery of climate information.

The workflow software is part of the so-called Digital Twin Engine (DTE)[14]. The Climate DT requires a flexible underlying software infrastructure to efficiently orchestrate and run it, as well as to access, handle, and interact with the vast amounts of data they produce. As part of the DTE, the backend of the Climate DT workflow uses the workflow manager Autosubmit (Manubens-Gil et al., 2016). Autosubmit is a lightweight workflow manager designed to meet climate research needs. It integrates the capabilities of both a workflow manager and a workflow orchestrator, all in a self-contained application. Autosubmit allows workflow suites to be centrally configured, communicates and leverages resources seamlessly in a range of computing platforms, and monitors the execution of the tasks. It offers full traceability (Leo et al., 2024), utilising well-established solutions to capture and manage prospective and retrospective provenance, in addition to a notification system that informs users of the progress of climate information production.

[13] https://sylabs.io/docs/
[14] https://stories.ecmwf.int/the-digital-twin-engine/

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

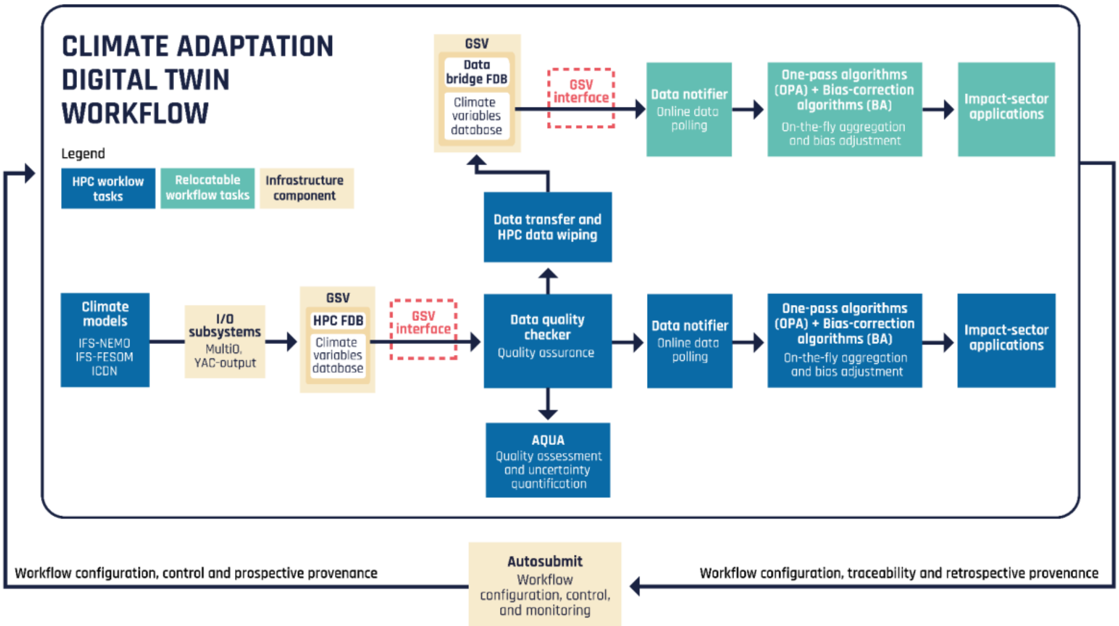

Figure 2: Climate DT workflow and dataflow. The end-to-end workflow tasks (blue colour) are executed on a HPC, supported by the digital twin engine software infrastructure (yellow colour). The data bridge allows the relocatable tasks (which run using containers) to be executed also out of the main workflow (green colour) for additional flexibility and data reuse. The Autosubmit workflow manager is used to configure, control, and orchestrate the end-to-end workflow, as well as tasks that are not part of the critical production path. The workflow manager allows the execution of all tasks to be monitored and traced.

The Climate DT infrastructure has been developed with a generalised use of continuous integration and delivery, agile software development methodologies, documentation of best practices, performance analysis, and portability. Continuous integration provides measurable indicators about reproducibility, replicability, and efficiency, and ensures code quality. Autosubmit handles an automatic testing framework to validate any changes in any Climate DT component and helps with the continuity of operational production.

The unified workflow facilitates a homogeneous and fast data treatment with traceability of its availability through the data notifier tasks, which are used along the production chain. The details are presented in the next section.

4 Seamless and homogeneous data access

To ease the use of climate data by all data consumers, the output of the different climate models is homogenised by introducing a "generic state vector" (GSV). In the GSV, the output of all climate models is unified in terms of parameters[15] and units (similarly to what CMIP does), as well as in spatial and temporal resolution, following a strict data governance that

[15] Ocean vertical levels are specific to each of the three global ocean models included in the Climate DT, while atmospheric data are output using a common set of pressure levels.

regularly accommodates new user requirements. The GSV concept facilitates consistency across models and enables a step change in the interoperability and usability of climate model output for climate-sensitive sectoral applications. This solution allows data consumers to access climate model data from any of the Climate DT models in the same way as the simulations progress.

Climate model data is individually encoded in a message format. This choice was made to align with the World Meteorological Organisation (WMO) member states' national authorities represented through the WMO Integrated Processing and Prediction System (WIPPS[16]) distribution system. The GRIB format has been chosen as it conveniently aligns with these standards and is proven to mitigate the risks of data loss and recovery since each data message is self-contained with its metadata, in contrast to newly evolving formats designed for speed of access in a distributed computational environment. The data under this format is written in the HPC storage using the Field DataBase (FDB[17]) technology, to provide standardised access. The metadata governance solution follows WMO standards[18]. However, it is recognised that other formats and standards combined with suitable conventions (CF convention[19]) are used in the climate community and with a substantial shift away from a focus on long-term storage of global fields to a focus on scalable cloud-based access of vast volumes of area-specific and global data, new access patterns are emerging. Particular features are lazily loaded xarray, compressed representations and zarr as an interface and/or storage format with specific choices for chunking and structuring the data that is gaining acceptance as a new community standard. This is recognised in DestinE so far with interoperable interfaces being developed to read and process global GRIB fields, thus mitigating storage and compatibility risks, but enabling similar access patterns to native zarr stores. Additionally, compliance with existing metadata conventions that synchronise both grib definitions and established CF conventions is ongoing work. Notwithstanding, the models underpinning the Climate DT may choose additional output streams in native zarr stores, which some recent European research projects such as nextGEMS and EERIE have shown to provide performant data access for users across a wide range of applications.

Computational grids used internally by the Climate DT model components are different from one another and from the common output grid used in the GSV. The common grid of choice for the GSV is the Hierarchical Equal Area isoLatitude Pixelation (HEALPix; Górski et al. 2005). While this deviates from common practice in weather and climate, where historically data sharing in common grids has used a regular latitude-longitude grid, managing unprecedented output volumes at high temporal frequency and fine spatial resolution required rethinking the approach. The advantages of the HEALPix grid are its equal-area uniformity, reducing over-precision near the poles and hence minimizing the storage requirement for any given resolution. HEALPix also simplifies hierarchical access patterns, AI training and data-driven workflows. It is also well suited to the aforementioned data access capabilities (e.g., zarr with chunking) allowing to

[16] https://community.wmo.int/en/wipps-web-portal
[17] https://github.com/ecmwf/fdb
[18] http://codes.wmo.int/grib2
[19] https://cfconventions.org/

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

minimise data movement and facilitating subsetting and nesting efforts when requesting local information from global fields. The integration into end-to-end workflows and data delivery services benefits from quick data transformations (e.g., interpolating to coarser grids) to suit consumer needs.

Climate models use their input/output (I/O) software to interpolate and write the output fields in the HEALPix grid. For IFS, FESOM, and NEMO the efficiency of the process is ensured by the use of the parallelised I/O software MultIO (Sarmany et al., 2024). ICON handles the parallel output using external processes coupled with YAC, as in Hanke et al. (2016), although using version 2.6.1. An experimental alternative to further improve efficiency of model output is now being explored with the use of the Maestro middleware (Haine et al., 2021).

The list of climate variables produced by the Climate DT simulations is defined in the data portfolio[20]. The data portfolio is flexible between production cycles to serve any new requirements from the data consumers. The output is offered with a daily frequency for the (two- and three-dimensional) ocean and sea ice variables, and hourly for the (two- and three-dimensional in pressure levels) atmosphere and land. Yet higher frequency output, which is desirable for some applications, is not yet possible, but is a focus of future model developments. The data is available in the HPC file systems to be used by the data consumers included in the workflow as soon as it is produced. The HEALPix resolution closest to the Climate DT climate model components is H1024 nested[21] (where 1024 corresponds to the *Nside* parameter described in the documentation[22]) for simulations of 5 km nominal horizontal resolution and H512 for those of 10 km resolution. As an example, each climate model produces around 100 GB of output per simulated day at H1024 resolution. With the target throughput of 1 simulated year per day (SYPD) at 5 km, each of the Climate DT models thus generates more than 35 TB per wallclock day in routine production mode. These production estimates could be at times higher when climate models perform more than one simulation simultaneously if a large enough share of the HPC is available. The Climate DT is inspired by the weather forecasting experience where similar data challenges are regularly dealt with.

To handle these large data volumes that are continuously produced, both in terms of storage and efficient consumption, the Climate DT employs the streaming concept. In the Climate DT streaming, data is made available to the data consumers embedded in the workflow as soon as a set of automatic checks[23] to detect missing fields, flawed metadata, and other potential errors such as unrealistic physical values, have been passed. If any of the critical checks fails, the workflow manager stops the production and warns the operators, who follow agreed procedures to proceed after diagnosing the problem. During the climate data production the workflow regularly notifies the data consumers about both the simulation

[20] https://destine-data-lake-docs.data.destination-earth.eu/en/latest/dedl-discovery-and-data-access/DestinE-Data-Portfolio/DestinE-Data-Portfolio.html

[21] Descriptions of the specific HEALPix grids can be found in https://easy.gems.dkrz.de/Processing/healpix/index.html#healpix-spatial-resolution, where the grid resolution is expressed by the parameter Nside, which defines the number of divisions along the side of a base-resolution pixel that is needed to reach a desired high-resolution partition; see also https://healpix.sourceforge.io/html/intro_Geometric_Algebraic_Propert.htm#SECTION420.

[22] https://healpix.sourceforge.io/html/intro_Geometric_Algebraic_Propert.htm#SECTION420

[23] https://pypi.org/project/gsv-interface/ and https://github.com/DestinE-Climate-DT/GSV-Interface

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

progress and the availability of new data. The Climate DT continuity is controlled by the data notifier task, which can be called by the consumer to maximise the computational efficiency of each application software. In this way, the streaming is ensured by the workflow but configured by the consumer, facilitating the inclusion of new data consumers in the production workflow.

To the best of our knowledge, this is the first time that such a detailed simulated multi-model climate state is offered to data consumers as the climate simulations progress with a homogeneous data model (to which they can contribute) and automated quality control. Additionally, in an alternative experimental setup, relying on the Maestro middleware for data streaming, data will be in the future directly accessible for consumers over the high-speed network while the model is running, bypassing and hence also alleviating the disk intermediary.

As the Climate DT produces regularly multi-decadal simulations from several models, accumulating high-frequency, high-resolution data soon becomes prohibitive for any HPC disk system. This has been solved by addressing both data storage and access with a tiering system. Data (so far all of it in phases 1 and 2 of DestinE) is transferred soon after it is produced to a managed storage cloud system attached to the HPC. These "data bridges" are implemented by EUMETSAT, as part of the DestinE "data lake", and managed jointly together with ECMWF (Fig. 2). The residence time of the data in the HPC, which is the time data consumers with HPC access have to exploit the full GSV, is known as the streaming window. The streaming window has been so far of up to one year, but as the simulation production ramps up it is expected to be reduced to several weeks, striking a balance between the limited data storage in the HPC and the access patterns in both the HPC and the cloud.

The streaming approach offers a unique opportunity, without precedents in climate adaptation, for data consumers to efficiently exploit the full model state at native resolution with highly-efficient data management. Once the data has been transferred to the data bridges, it is gradually erased from the HPC filesystem to make space for new climate simulation output. The transfer to the data bridge is another task managed by the Climate DT workflow.

The approach used by the Climate DT connects the HPC platforms (which have access restrictions) to infrastructures with cloud-like data access and services. This spares data consumers not embedded in the production workflow the complexities of the HPC data access and operational procedures, offering user-oriented access to federated, very large data sources through DestinE's data lake services implemented by EUMETSAT and the DestinE platform implemented by ESA, always applying DestinE's data access policy. The first set of the Climate DT datasets (DestinE, 2025) has been used in this manuscript.

A challenge of data streaming is that a straight implementation only allows an instantaneous view of the data. To address this limitation a variety of one-pass algorithms (OPAs; Grayson et al., 2025; Figure 3) have been developed to estimate statistics, such as time averages, variances, threshold exceedances, percentiles or histograms, and offer temporal buffering in the streamed variables. The OPAs can handle the output from any of the climate models seamlessly thanks to the homogeneity of the GSV. The OPAs can be called by the applications embedded in the workflow to access seamlessly the data available either in the HPC or the data bridge, and deliver user-relevant processed variables and indicators. The functions have been optimised to deal with the high-resolution, high-frequency fields, making requests to the GSV interface prompted by data

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

notifiers. They provide full traceability of the data processing. The OPAs are another Climate DT novelty to deliver user-oriented sources of climate information in different formats, including NetCDF on disk and Python's Xarray in memory. When they are integrated as a task in the production workflow, they make use of a local buffer to restart the data consumer operations in case of a climate model failure that forces the climate simulation to restart from the last checkpoint file.

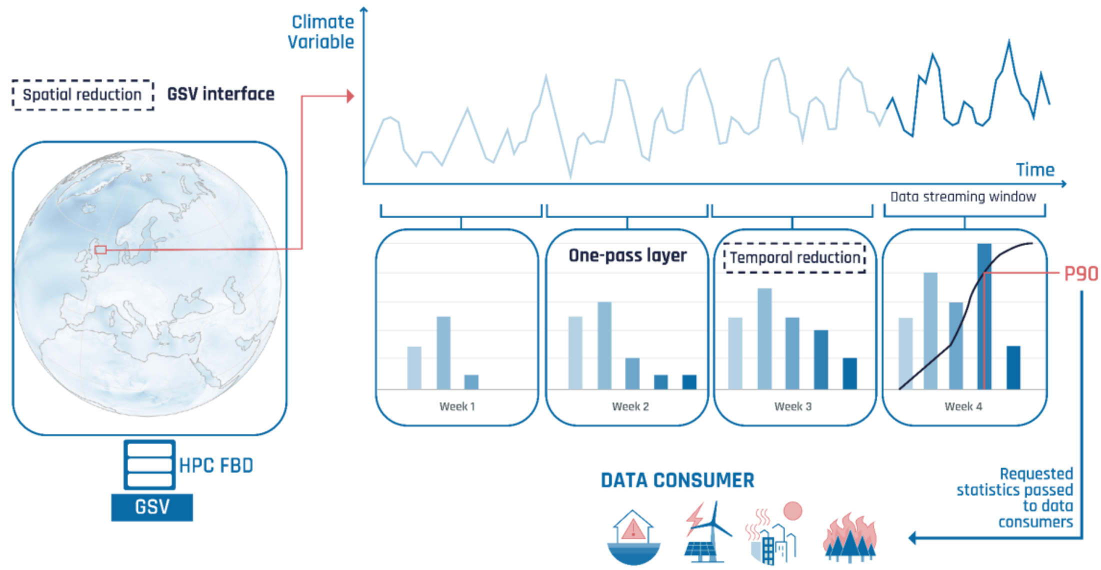

Figure 3: Schematic showing how the climate model output stored as the GSV in the FDB is retrieved via the GSV interface, including data interpolation if requested, and passed into the one-pass layer. The one-pass algorithms (OPAs) provide temporal reduction by computing a statistic from native data as requested by a data consumer, in this case percentiles from weekly histograms of a specific variable. The statistic is continuously updated, controlled by data notifiers tasks, as new data from the climate simulation becomes available, always within the streaming window allowed by the data management process. The statistic is regularly exposed to the data consumers and feedback is sought within a co-production approach.

With its data streaming capability, providing the full access to the climate model state as the models perform their simulations and data processing enabled by the combination of GSV, data notifier and OPA capabilities, the Climate DT functions as a virtual instance of the historical and future climate. It can be considered a metaphor in the climate modelling realm of an observing instrument, which samples reality with the highest frequency possible relevant to the process under study, while in parallel the data is thinned from the high-frequency sample to make the data stream manageable. In this metaphor, if a data consumer cannot make use of the data available in the streaming window (i.e., before the HPC storage is flushed to make space for more model output), they will have to either 1) use model output from the data bridge as described previously if the necessary variables are available or 2) wait some time for a new climate simulation, either with a new model version or from a new member of an ensemble.

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

5 Production of the first set of global climate projections and storyline simulations with local granularity

The Climate DT exploits and co-develops a new generation of global storm-resolving and eddy-rich models, often referred to as km-scale models, and HPC advances. The three global climate models used, ICON, IFS-NEMO, and IFS-FESOM, have been adapted to perform km-scale simulations through a cooperative development approach supported by the European-funded research projects nextGEMS[24] and EERIE[25], as well as national initiatives such as WarmWorld[26] in Germany and Gloria[27] in Spain, involving many climate, weather, and supercomputing centres, and academic partners throughout Europe. In the Climate DT, these models are used to perform climate simulations at nominal resolutions ranging from 5 and 10 km for the different components of the Earth system, exploiting the supercomputing facilities of the EuroHPC JU. The current throughput is approximately 0.5 simulated years per wall clock day (SYPD) at 5 km resolution, and 2–3 SYPD at 10 km, using around 200 and 100 computing nodes, respectively. Efforts are underway to improve the HPC adaptation of the models to enhance their energy efficiency, while aiming to reach a throughput of about 1 SYPD through various optimizations such as reduced precision or efficient GPU porting.

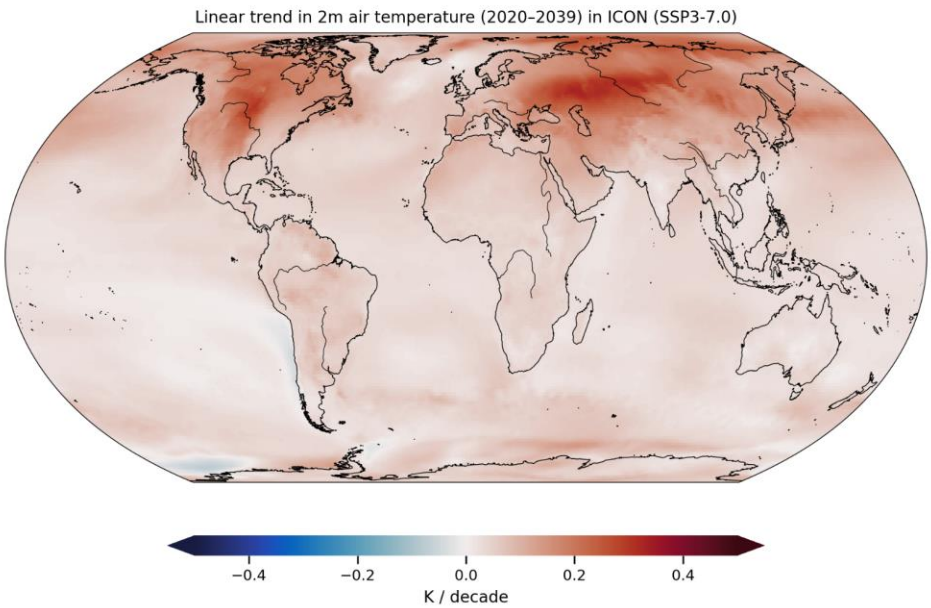

Figure 4: Projected trend (in K/decade) for the annual-mean surface air temperature for the period 2020–2039 in a SSP3-7.0 scenario simulation performed with the ICON model.

The Climate DT infrastructure is not limited to deterministic km-scale simulations but also allows the production of climate simulations following standard CMIP protocols, including ensembles, following production cycles of less than a year. It can

[24] https://nextgems-h2020.eu/
[25] https://eerie-project.eu/
[26] https://www.warmworld.de/
[27] https://www.bsc.es/research-and-development/projects/gloria-global-digital-twin-regional-and-local-climate-adaptation

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

also be used to perform bespoke simulations to assess the impacts of different scenarios and policy decisions and to address "what-if" questions.

The first set of simulations performed by the Climate DT have followed a streamlined version of the HighResMIP protocol (Haarsma et al., 2016) on both the LUMI and MN5 EuroHPC pre-exascale supercomputers. The protocol consists of 1) 30-year long control simulations with constant 1990 forcing to evaluate model drift, 2) historical simulations starting in 1990 and ending in 2020, and 3) projections for the period 2021–2040 following the SSP3-7.0 scenario defined in the sixth phase of CMIP (CMIP6; Eyring et al., 2016). The projection simulations with IFS-NEMO were carried out at a horizontal resolution of 4.5 km for the atmosphere and land and 1/12º (around 8 km at the Equator) for the ocean and sea-ice, with IFS-FESOM using the same atmospheric and land resolution of 4.5 km and about 5 km over most of the globe for the ocean and sea ice, while the ICON simulations were performed at 5 km resolution for all Earth system components. The historical and control runs were carried out at about 10 km for the atmosphere and land, and about 10 km for both ocean and sea ice. In all simulations, the ocean/sea-ice models were spun up for five years using stand-alone ocean runs forced with the Copernicus Climate Change Service ERA5 reanalysis (Hersbach et al., 2020) for the corresponding initial date (1990 for the historical and control and 2020 for the projection) that were started from EN4.2 ocean estimates interpolated to the corresponding ocean grid (Good et al., 2013). The ocean spin up was followed by a two-year coupled ocean-atmosphere spin up with constant forcing in the IFS-NEMO and IFS-FESOM cases.

This experimental configuration allows for some comparisons with the simulations produced by CMIP and CORDEX, always taking into account the important differences in the experimental setups, while allowing the development and testing of the Climate DT prototype. Figure 4 shows a simple example of the future near-surface annual-mean air temperature trend calculated over 2020–2039. In a fully operational context, additional simulations will be performed on a regular basis as part of the production cycle, either with the same versions to increase the ensemble size or with new ones as part of a new operational cycle to take advantage of recent improvements in the Climate DT workflow elements. New operational cycles typically occur every nine months to a year. The simulations performed after the initial set described above use homogeneous resolutions for the control, historical, and projection and include simulations at both about 5 and about 10 km horizontal resolutions.

In addition, the Climate DT features a class of high-resolution storyline simulations that allow to "rewind and replay" recent extreme weather events, such as heat waves or floods, and explore how they would unfold under different climate conditions, from pre-industrial to warmer futures (John et al., 2025). These physically consistent "what-if" experiments link directly to observed events through spectral nudging, anchoring the simulations to the observations while exploring alternative climate trajectories (e.g., Athanase et al., 2024; Sánchez-Benítez et al., 2022).

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

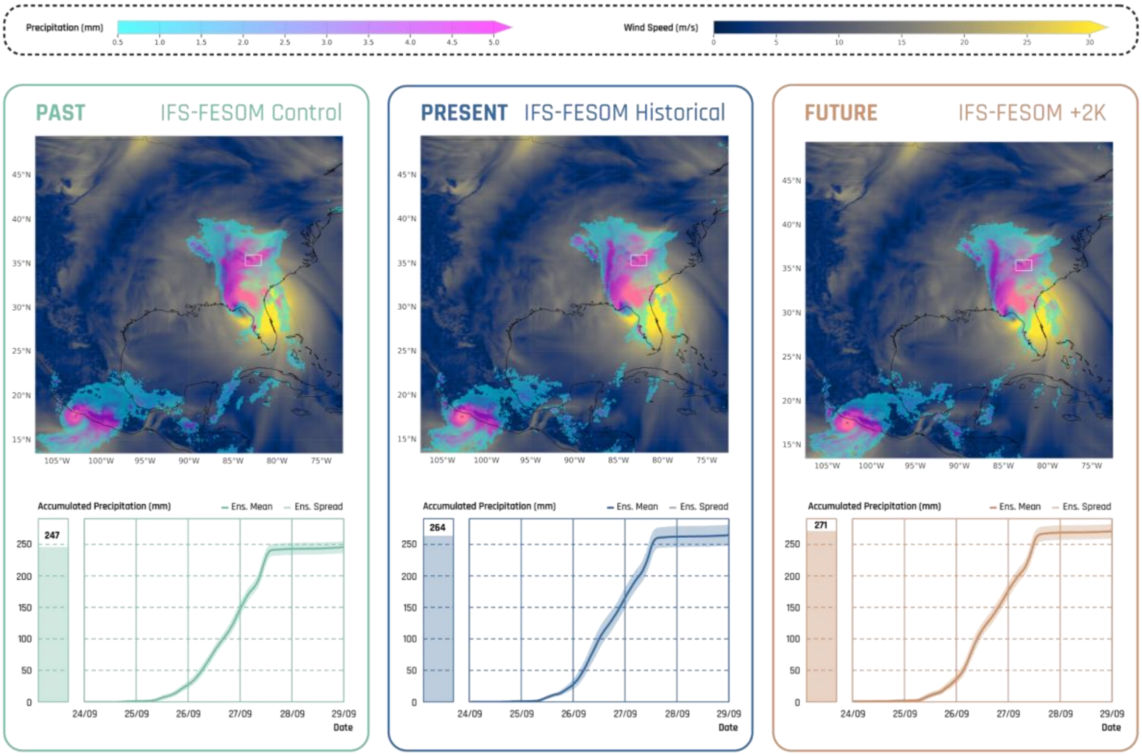

Figure 5: Snapshots of 850-hPa wind speed and precipitation associated with hurricane Helene just after landfall on the 27th of September 2024 at 04:00 UTC simulated using IFS-FESOM in which large-scale tropospheric winds nudged toward ERA5. The three panels represent different climate scenarios: PAST (IFS-FESOM Control, ~1950s climate), PRESENT (IFS-FESOM Historical, current climate), and FUTURE (IFS-FESOM +2K global warming level). The lower panels show accumulated precipitation (mm) in the highlighted region with a white box in each map from the 24th to the 29th of September, with totals increasing from 247 mm (Past) to 264 mm (Present) and 271 mm (Future). Simulation details are described in John et al. (2025).

The storyline simulations are performed with IFS-FESOM (Sánchez-Benítez et al., 2022; John et al., 2025) at resolutions of about 10 km for the atmosphere and land and 5 km for the ocean and sea ice. They represent an important application of the Climate DT's capabilities because they provide concrete, location-specific insights into how climate change is reshaping extremes, making risks more tangible and adaptation planning more actionable. These simulations reconstruct the evolution of the climate system, including extreme events such as heatwaves, floods, storms, and drought, from 2017 to the present, with continuous updates close to real time, under three distinct climate conditions: a past climate resembling the 1950s, the present-day climate, and a future scenario with 2 °C warming above pre-industrial levels (assumed to be reached around the 2050s). To ensure realism, the large-scale atmospheric circulation of IFS-FESOM is nudged to ERA5 reanalysis data above the boundary layer, while other components evolve freely under the prescribed climate forcing. This approach allows the same physical event to be simulated across different climate states, offering insights into how thermodynamic changes modulate its intensity, duration, and impact. For example, Figure 5 shows how the simulations reproduce essential features of hurricane Helene, including mesoscale structures, and how climate change may have influenced its characteristics,

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

supporting both event attribution and forward-looking scenario analysis. Because the simulations are global and consistently high-resolution, they enable the examination of multiple, concurrent extremes worldwide, providing locally relevant insights into compound climate risks.

The Climate DT simulations require the extreme computing power and data handling capacities provided by the EuroHPC JU machines. A substantial investment has been made in adapting the models to hybrid CPU-GPU architectures. Currently, a 20-year simulation at 5-km resolution typically requires around 0.6 million GPU-hours in the case of ICON or 23 million CPU core hours in the IFS-NEMO case on the LUMI HPC. With these figures, the computational resources available allow performing a multi-model ensemble in each production cycle. Adaptation efforts have been accompanied by a systematic performance analysis of all the steps necessary to complete a climate simulation and bottlenecks are continuously addressed. In addition, automated performance information collection (Acosta et al., 2024) has been implemented in the workflow to monitor the progress of all models and detect anomalous behaviours. The analysis identified issues that limited performance and suggested code and runtime modifications. It inspired work leading to code refactoring, transfer of code sections to accelerators, pipeline restructuring, input/output choices, and runtime optimisation for the different computing platforms.

5.1 Real-time scientific quality assessment

The simulations of an operational system require a monitoring service that evaluates the scientific performance and allows detecting unexpected climate model behaviour. The scientific evaluation of the climate simulations is carried out by two specific applications developed for the Climate DT.

The Application for QUality Assessment (AQUA; Nurisso et al., 2025) is an open-source Python package built upon a core engine facilitating data access based on catalog access and designed to process high-resolution global data in an efficient and scalable way (Fig. 6). AQUA has similar properties to other climate diagnostic packages (e.g., ESMValTool; Lauer et al., 2024). It uses a series of modular and independent diagnostics. It has been designed to analyse the simulations from all models in a homogeneous way and benefits from the data checks performed upstream in the workflow to detect corrupted data. It is another data consumer that can be used either embedded in the workflow (to monitor the simulation progress) or as a stand-alone tool to address specific questions and comparisons with either other simulations or observational references. The climate model output is compared, as the climate simulations progress, with either reanalyses or benchmark simulations by AQUA to evaluate and monitor aspects of their quality. This innovative capability helps to quickly identify potential problems by benchmarking the climate simulations against existing experiments and allows continuous monitoring of the progress and quality of the simulations.

The assessment focuses on several state-of-the-art diagnostics and metrics targeting the mean state model biases over the historical period for selected regions (e.g., performance indices from Reichler and Kim, 2008) and variables or basic metrics (global mean air temperature, top-of-the-atmosphere energy balance, atmospheric teleconnections, metrics of ocean circulation, etc.) that are essential to understand the evolution of the climate system and its internal modes of variability. This information contributes to the quality assurance of the climate simulations and builds trust into its use for any adaptation

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

decision. AQUA also includes a reduced set of frontier diagnostics that exploit the native spatial resolution and high frequency of the data. They aim at providing statistics on physical processes that could not be systematically investigated in km-scale climate simulations due to the huge data management challenge they pose. These diagnostics include statistical properties of the tropical rainfall or tropical cyclone characteristics. The AQUA diagnostics are designed to account for multi-model ensembles to provide a measure of uncertainty for both historical simulations and climate projections.

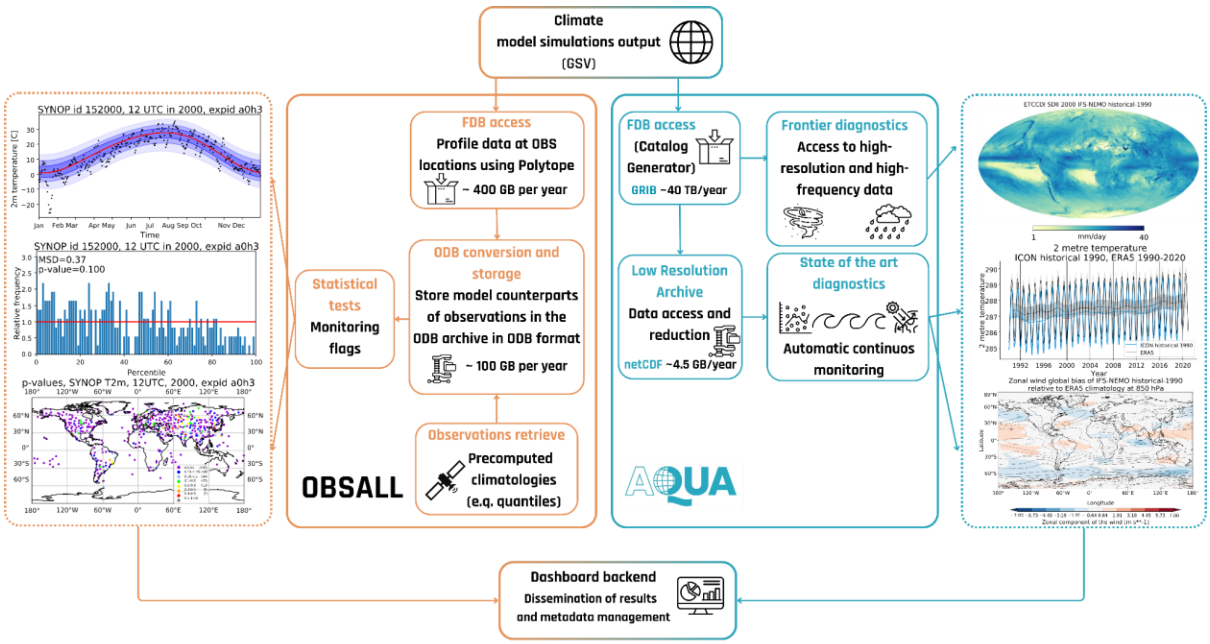

Figure 6: Scientific model evaluation strategy of the Climate DT. The AQUA processing (blue boxes) enables two-way model evaluation through automatic-generated catalogue entries. High-resolution data supports high-granularity frontier diagnostics, such as the precipitation-based simple daily intensity index (SDII) from the ETCCDI collection, estimated for the year 2000 from an IFS-NEMO 10-km historical simulation (top-right panel). Simultaneously, data is aggregated into a low-resolution archive (monthly averages on a one-degree regular grid) for continuous monitoring via state-of-the-art diagnostics. Another example diagnostic shows globally averaged temperature and one-year moving average of the temperature from the ICON historical simulation and ERA5 (middle-right panel). Yet another example shows the mean bias of 850 hPa zonal wind in the IFS-NEMO historical simulation with respect to ERA5 (lower-right panel). The OBSALL processing (orange boxes) involves time-critical GSV access, model profile retrieval, application of observation operators, and comparison with observed climatologies. The left panels show results for 12 UTC two-metre air temperature from an IFS-NEMO 10-km historical simulation. Temperatures for the year 2000 at Arad, Romania, are shown against observed quantiles in the top-left panel. The histogram (middle-left panel) illustrates the probability of differences over the whole year arising by chance, with a p-value of 0.1. The p-values for other stations are shown in the bottom left panel.

The Climate DT also includes a bespoke observation-based quality assessment system, named OBSALL (Fig. 6). It generates an image of the full-resolution climate simulation in the observation space. The image is a trace of the simulation as if recorded by an observing system. In stand-alone mode, OBSALL allows the quality monitoring to focus on, e.g., the formation of surface temperature inversions in the polar night as viewed by the synoptic surface network. In on-line mode, OBSALL makes a statistical assessment of whether the simulation stays within the envelope formed by the quantiles of the observed daily climatology and if not, it raises flags. Technically, OBSALL can be interpreted as an OPA that can access the full-resolution, high-frequency GSV data. It extracts model profiles at observation locations with the Polytope tool (Leuridan et al., 2023) and applies standard observation operators (ECMWF, 2015) to compute the model counterparts for observations. Both are archived into an observation database (ODB; ECMWF, 2010) using the ECMWF Python package pyodc (ECMWF, 2025). High-quality and stable observational platforms and networks are selected so that informative observation-based climate statistics can be precomputed and stored in the ODB and used in the quality assessment. The ODB input archive for Climate DT simulations contains observations from synoptic surface stations (Dunn et al., 2014), upper-air soundings (Madonna et al., 2022), and AMSU-A radiances (NOAA, 2025) as a remote sensing data demonstrator. The ODB archive and the observation projection in general are designed to be easily extendable. Results of both AQUA and OBSALL are regularly displayed on a Climate DT internal dashboard.

6 Prototype climate-change impact applications

A characteristic of the Climate DT is its ability to generate tailored and timely information for climate-sensitive sectors. The end-to-end production of user-relevant climate information is illustrated with a selected number of sectoral climate applications that have been embedded in the workflow as data consumers and can compute relevant indicators as the climate simulations are performed. At the same time, these applications play an important role in shaping the Climate DT operation. By being an integral part of the Climate DT, they provide continuous feedback about the digital twin development like the data portfolio, the experimental setup, the software to maximise the application throughput, the data strategy, the OPA capabilities, etc., all through a co-production process. Every embedded application has been implemented with selected users, who participated in the process through regular interactions. The co-production allows to illustrate the relevance for the climate adaptation community. Innovative climate information can be generated by making use of the full climate model state described above as the simulations are performed, while more traditional asynchronous access to the resulting data included in the DestinE portfolio is still possible for those applications that are not embedded in the workflow by accessing data in the data bridge. These applications can also be included in the production workflow in later production cycles.

The climate-adaptation applications integrated in the Climate DT so far are:

- Wind-energy management and energy demand: This application illustrates the value of the Climate DT for informing the investment and management strategies of the wind energy sector. Wind resources are per se highly heterogeneous and intermittent. Wind-energy generation is particularly sensitive to sudden variations in resource availability, which are typically not captured in current climate simulations. To address this need, this application provides indicators, such as high- and low-wind event frequency (Rapella et al., 2023), capacity factors, and annual energy production, among others, for a range of turbine types (Lledó et al., 2019), at global scale and high resolution from high-frequency wind data (Fig. 7). These indicators have been selected in collaboration with a company that provides estimates of the wind resource to energy producers. In addition, for specific areas with offshore installations, the application provides indicators relevant to the construction and maintenance that require high-frequency data, such as sea-ice occurrence, ice-related stress to structures, and expected navigability conditions. Beyond these indicators, robust statistics for extreme clustering events will require the accumulation of several decades of data (Priestley et al., 2018).

- Freshwater availability and flood occurrence: This application, named HydroLand, provides global estimates of river runoff at the same spatial scale as the climate simulations. This is an improvement over current hydrology information systems for climate time scales. The HydroLand application is based on the mesoscale Hydrologic Model (mHM; Samaniego et al., 2010) and the multiscale Routing Model (mRM; Thober et al., 2019). mHM uses Climate DT data at the highest spatial and temporal resolution available and mRM provides a wide range of hydrological variables and indicators depicting global freshwater availability to support insights about the requirements for climate adaptation from global to local level in areas of particular interest (Fig. 7). The application can also take advantage of the simultaneous bias-adjustment of the climate simulations, where bias adjustment parameters are estimated using the model output of existing members using the OPA.
- Characteristics of hydrometeorological extreme events: The HydroMet application summarises user-selected statistics of hydrological extremes by identifying extreme rainfall events over Germany. This leads to information about the spatial and temporal extent, precipitation amounts, return time, and the frequency of extreme events. The application is based on KOSTRA-DWD-2020 (Junghänel et al., 2022), a solution for the evaluation of precipitation levels, duration, and the annual return interval, and on the Catalogue of Radar-based heavy Rainfall Events (CatRaRE; Lengfeld et al., 2021) software package that detects spatially and temporally independent heavy precipitation events leading to flash floods. High global spatial and temporal data frequency over what is currently available as climate sources is required to estimate these indicators at the level required by the range of users that collaborate with the developers.
- Fire weather and wildfire spread: This application focuses on wildfire management strategies. The high-resolution Fire Weather Index (FWI) identifies where meteorological drivers of fire weather conditions are conducive to the occurrence and persistence of fires (Abatzoglou et al., 2019). Additionally, the fire spread model WISE simulates critical fire parameters (e.g., burnt area and fire intensity) at high spatial resolution, utilizing input data co-developed with users, including land-use and ignition information (Touma et al., 2021). By leveraging the high-frequency, high-resolution data along with user-defined local land-use scenarios, WISE is used to assess wildfire management options such as fuel and firebreaks management. This supports the user-driven planning and implementation of adaptation strategies to increasing fire risk (Hetzer et al., 2024).

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

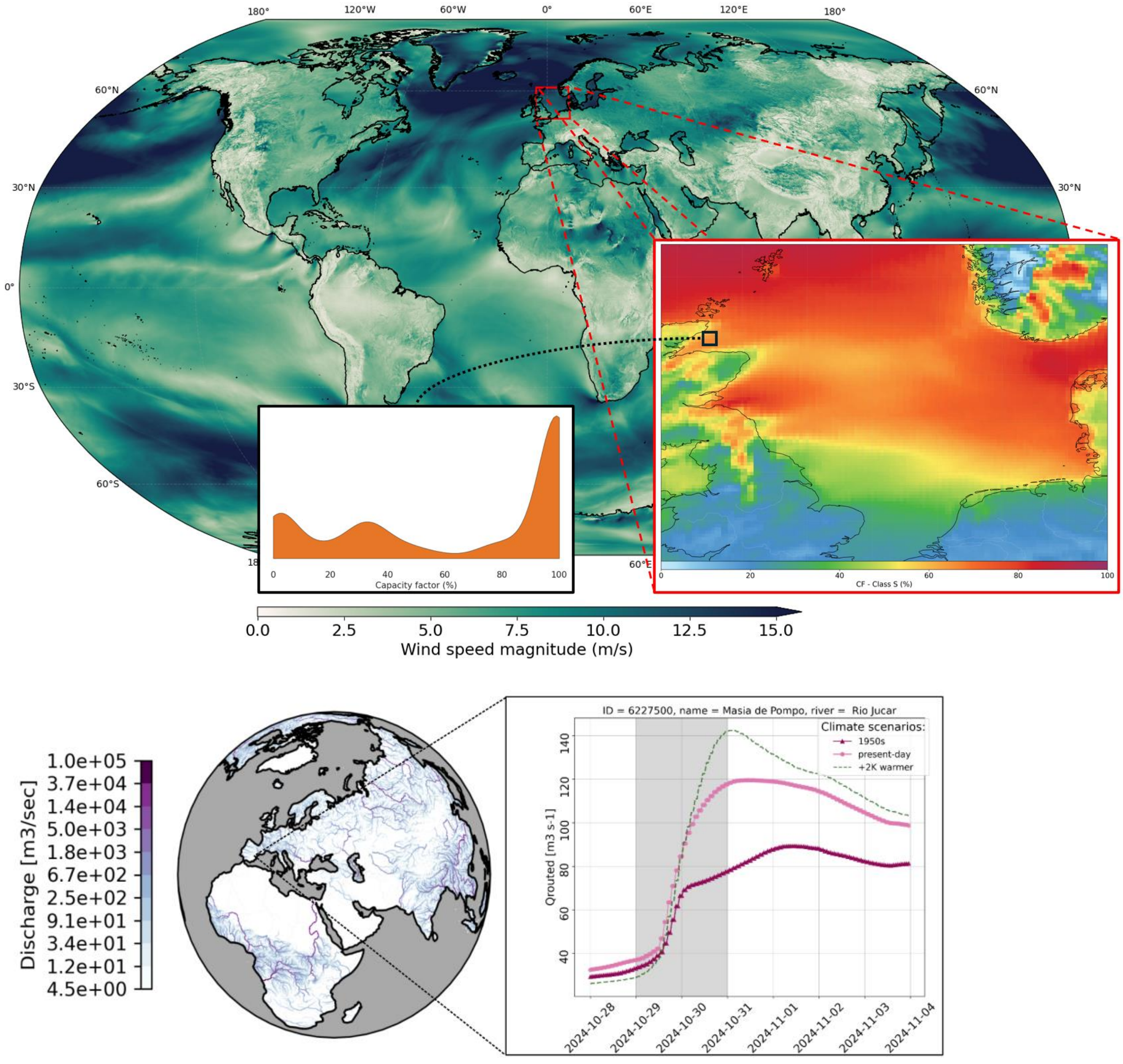

Figure 7: (Top panel) Global map with a snapshot of wind speed at 100 metres from the 10-km IFS-NEMO historical simulation. The regional zoom, highlighted with a red rectangle, shows the capacity factor averaged over a week for a class S Vestas V164 wind turbine over the North Sea computed from 1-hourly wind components (zonal and meridional 100-metre wind). The black square marks the location of the Moray East wind farm, off the coast of Scotland, which operates this specific type of turbine. The curve represents the distribution of the hourly capacity factor for Moray East during the week. (Bottom panel) The HydroLand application provides global estimates of terrestrial hydrological processes like river discharge over small areas (~80 km2), soil moisture deficits, and extreme flooding events. The time series represent the river discharge values over the Jucar River in the Valencia region (Spain) during the period of the storm event at the end of October 2024 using climate data from the 1950s, present-day, and future climate (+2K global warming level) IFS-FESOM storyline simulations.

Commonalities among the applications have been identified to develop a protocol that reuses as much of the technical developments as possible to offer a timely access to the climate data. The GSV, through its unified access and formats, offers a flexible and responsive environment to address the emerging requirements, while reducing computational and data handling overheads. It also hides some of the complexity of the climate data from the application developers.

The experience gained with these applications underpins additional efforts to implement applications that address other user requirements. The continuous Climate DT production, generating new ensemble members in succession, can leverage newly acquired knowledge about the digital twin performance from previous simulations and be used to modify any element of the production workflow. For instance, in each new production cycle the list of output variables can be modified to allow the creation of new user indicators or climate model diagnostics missing in previous simulations.

7 Looking forward in the Climate DT

While previous multi-decadal climate projections have been performed every five to seven years, the Climate DT aims at producing operational, quality-assured simulations with frequent updates (at least yearly) and real-time access to the data generated, either in raw format or processed according to the requirements as collected by climate services. The frequent updates allow the integration of both the latest advances in science and technology and the new requirements from users to support decision-making supported by the best climate science. The Climate DT also aims at producing climate simulations on demand, to answer what-if questions from relevant actors like climate services providers, public administrations, and the European Commission.

The previous sections describe the prototype Climate DT developed in the first phase of DestinE (2022-2024). Currently, in the second phase (2024–2026), the focus is on consolidating, operationalising, and further evolving the Climate DT system, aiming to gradually enhance its capabilities and response to user requirements for efficient climate adaptation. This section summarises some of the developments undertaken or planned.

7.1 Operational framework

The Climate DT is laying out the basis for the operationalisation of user-oriented, multi-decadal, multi-model climate projections. This is an initiative complementary to the existing CMIP and CORDEX international efforts and to the activities of national climate services and the Copernicus Climate Change Service in the provision of future climate information sources. Previous sections have described the Climate DT concept and its goal to ensure the robustness and traceability of the whole production chain. The regular production of climate simulations at this scale can only be maintained by ensuring that the computational performance of all elements is monitored, analysed, and optimised as the underlying HPC platforms evolve.

The Climate DT simulations are not time-critical, compared to weather forecasts, but the operational context requires that they are regularly performed to address continuously emerging data consumer requirements. The simulations are long, relatively slow, and produce enormous amounts of data that need to be exposed, served and handled. A break of a few days in the production due to the unavailability of the HPC system is not a problem, but discontinuities should be appropriately

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

disseminated to the data consumers, especially to those not embedded in the workflow. At the same time, the production should be flexible enough to continue on alternative systems whenever needed. This requires data consumers embedded in the workflow to integrate and schedule their own restarting mechanisms, an aspect on which the development team is already supporting the selected applications.

The Climate DT is updated with improved climate models (whenever significant developments are available and have been thoroughly tested), more adequate workflow management, more relevant data checks, faster and more user-driven transfers to the data bridge, richer monitoring and evaluation, and a broader range of adaptation indicators. For instance, climate model releases follow developments that are both scientific, where the diagnostic capabilities play a central role, and technical, linked to aspects like data governance or code scalability. Most of these developments are included in development suites. The experimental suites apply a thorough testing to the proposed developments to be included in the operational suite with stringent acceptance criteria, ensuring that transitions from development to experimental and later operational phases are robust. This makes the operational production increasingly stable, traceable, and fit-for-purpose (Dee et al., 2024) for both data consumers and users accessing products from the data bridges.

A fundamental objective of the Climate DT in the second phase of DestinE is to consolidate the prototype system developed in the first phase and transition it toward operational use. This includes planning and executing a wide range of climate simulations with different objectives. These include increasing the ensemble size for uncertainty quantification (in historical, projection, and storyline simulations), generating projections for new scenarios or updated model versions, and conducting bespoke simulations tailored to specific policy or user needs, such as exploring "what-if" questions or extreme event storylines. The scheduling of such simulations is guided by a dedicated committee and is subject to the availability of computing resources.

An efficient operationalisation requires that computing resources are readily available, not just for the operational suite but also for scheduling all the development and experimental suites that a flexible, well-tested, regular production requires. Given the current computational cost of the Climate DT operation, ensuring the resource availability is a continuous challenge despite the strategic EuroHPC allocation that makes possible the current production.

There are many other aspects under discussion relevant to the operationalisation of decision-oriented climate projections. A close collaboration with the CMIP and CORDEX communities, as well as with those developing emission scenarios, climate-impact models, and climate services is necessary to ensure that all the components of the system remain state-of-the-art and current practices are taken into account. This collaboration is required for an efficient transfer between research development and user-oriented operations.

7.2 Increasingly flexible data management

The Climate DT data is a combination of a moderate volume set located in the data bridge for traditional climate applications and research, and an agile, rich, short-lived set in the HPCs for simulation evaluation, embedded applications, and online machine-learning training. All of it to be accessed and consumed using the same tools. The challenge is even more ambitious because it needs to strike a balance between storage (both in the HPCs and the data bridge), the ability to recompute specific parts of the climate trajectories (using stored checkpoint files) with enhanced data output, and the commitment to serve global, high-resolution data to a wide range of users. This balance needs input from as many stakeholders as possible. For instance, the selection of the data transferred to the data bridge, such as relevant vertical levels to retain, the possibility of reducing the numerical precision or resolution for some variables or the adequate data frequency for each variable, is currently discussed with the data consumers with experience in the Climate DT production. However, this interaction should be widened.

The data stored and made available in the data bridge are not necessarily homogeneous across all the members of a given production cycle, and there might also be differences across ensemble members. This happens because the archive is enriched over time with simulations that respond to emerging data consumer needs that lead to modifications of the data portfolio. For this reason, a solution that fully describes the data exposed from the bridge needs to be developed to ensure an efficient use of the datasets.

7.3 Penetration of artificial intelligence solutions

Another important aspect for the future development of the Climate DT is the integration of the current fast-evolving AI and machine learning breakthroughs. The Climate DT has started to exploit these recent advances, acknowledging that they require the best input data and traceable sources to build trust (Bracco et al., 2025; Eyring et al., 2024). Ongoing efforts in this regard focus on:

- enhancing the user experience through the development of chatbots that make the Climate DT data more accessible to decision-makers by including it into context-relevant reports (Koldunov and Jung, 2023) using fine-tuned large language models with the help of a climate-adaptation specific corpus;
- exploiting the wealth of the unprecedented high-resolution, high-frequency global data produced by the Climate DT models for AI model training to build emulators[28] of the physical climate models; the emulators will be used for uncertainty quantification of the simulated trajectories, for the fast reconstruction (Kadow et al., 2020) of the simulated trajectories as reruns for selected time slices (Rackow et al., 2024) that can provide high-frequency data on demand as described in the "AI on top" concept (Bauer et al., 2023), and for the exploration of the climate evolution for a range of emission scenarios not sampled by the physical models.

Results from these efforts will be reported in separate papers in the coming months.

7.4 Interactivity to address user needs

One area of increased focus in future phases is the interactivity of the Climate DT system. The operationalisation leads to significantly shortened production cycles compared to existing climate projection sources, enabling the possibility of

[28] https://destine.ecmwf.int/news/spains-predictia-to-build-a-climate-emulator-for-destine/

enhancing the interactivity in the access to climate projection sources. As a result, some users will at some point be able to explore different scenarios and address their what-if climate-related questions with the Climate DT. The options could include requesting new projections, slices of existing projections, or increased ensembles of storyline simulations. Impact sector applications can consider what-if analyses specific to their domains in separate workflows supported by the Climate DT infrastructure. An example of this option consists in assessing how changes in land cover modify flood impacts using HydroLand. In addition, AI solutions introduce new possibilities for interactivity, such as fast, on-demand data generation through emulators supported by chatbot-based interfaces, allowing users to formulate requirements using natural language.

8 Summary and conclusions

The Climate DT is a pioneering effort to build an operational climate information system based on a digital infrastructure to support climate change adaptation decisions. It provides globally consistent data with very high temporal (hourly) and spatial resolution (between 5 and 10 km), delivering updated climate simulations frequently, and considering user needs in terms of the output variables, access patterns, data processing, and simulation design. It benefits from a convergence of an infrastructure, dedicated resources, digital innovation, and domain experts with the adequate knowledge, perspectives, and experience, in a transdisciplinary endeavour. The Climate DT illustrates how to tackle some of the existing challenges (Stevens et al., 2024) in providing climate-vulnerable communities with timely, decision-oriented global climate information. It complements existing sources of climate information for adaptation focusing on the need for salience, equitability, and credibility.

The features introduced by the Climate DT can be summarised as follows:

- Operational multi-decadal projections: The operational nature of the Climate DT ensures the continuous production and delivery of climate simulations that support the availability of up-to-date information for climate up to 2050. This is made possible by leveraging EuroHPC capabilities, using operational end-to-end workflows, performing real-time monitoring to ensure the accuracy, fitness-for-purpose, and credibility of the climate data, quantifying uncertainty to offer reliable information, and involving professional workforce and practices that lead to regular tests and software optimisation for a range of computing platforms. The experimental setup is inspired by CMIP (more specifically by HighResMIP), illustrating the alignment between CMIP and the Climate DT. This set of capabilities enables the production of on-demand climate projections to address new policy-relevant questions, filling an important gap concerning the timeliness of the climate information requested by a range of stakeholders.
- Global km-scale climate modelling: The Climate DT benefits from a global km-scale modelling capability with three climate models. The current generation of global climate models often use spatial resolutions that are insufficiently fine to accurately represent important regional aspects and extreme events (Slingo et al., 2022; Collins et al., 2024). The improved representation of critical processes, associated with a substantial increase in spatial resolution with respect to the state-of-the-art models, has been shown to reduce some long-standing biases

(Moreno-Chamarro et al., 2021; Moreno-Chamarro et al., 2022; Rackow et al., 2022). The simulation of these processes and their feedbacks on the global climate circulation require eddy-rich, storm-resolving climate models with a grid-spacing of around or less than 10 km (Moreton et al., 2020; Beech et al., 2024; Roberts et al., 2025; Segura et al., 2025). This category of models can be also seen to provide global downscaling capability, where the simulations provide information with local granularity globally, supporting the need for equitability in climate information. This novelty has only been possible through enormous code refactoring efforts (including code porting to use GPUs) to adapt the models to the new HPC generation. A handful of these models have been built by several global actors, among which those participating in the Climate DT (Bauer et al., 2021a; Rackow et al., 2022; Segura et al., 2022; Streffing et al., 2022; Rackow et al, 2025; Wedi et al., 2022: Taylor et al., 2023; Donahue et al., 2024).

- High-resolution climate storylines: The Climate DT includes a type of simulation named storylines, where the atmospheric model follows the observed atmospheric circulation but with conditions corresponding to a specific global warming level (John et al., 2025). Three simulations for 1950, current, and two-degree warming levels are performed in parallel. These simulations start in 2017 and are produced continuously close to real time. They lead to tangible and relatable estimates of the effect of climate change on emerging extreme events, as they occur, supporting adaptation decisions where the future information is linked to personal experiences, enhancing the salience of the climate-change information. This approach offers the separation of the extreme event drivers, because the storyline simulations split the dynamical and the thermodynamic components of observed events. It can support near-real time attribution efforts that will soon be offered by the Copernicus Climate Change Service.
- Unified workflow: The infrastructure is designed to deal with climate models that aim to achieve the highest computing throughput possible while handling end-to-end dataflows where the streaming allows concurrency and efficiency in the data consumer execution. The end-to-end nature of the workflow suites and the unified approach to facilitate the use of any HPC is a unique Climate DT feature. The climate digital twin relies on Autosubmit (Manubens-Gil et al., 2016), an orchestrator that executes tasks with full traceability (Leo et al., 2024), includes a notification system of the production progress, and makes use of a unified interface to configure the digital twin production. The workflow software seamlessly integrates all the climate models and data consumers, with the same solution applied in all the computing platforms available, avoiding the current practice of engineering different solutions for each model, application, and platform. The modular end-to-end workflow, along with a generalised use of containers for the portability of embedded data consumers, significantly improves workflow maintainability, the extension to other climate models and data consumers, and the operational practice at large.
- Unified data handling: The Climate DT has introduced several data-related innovations to enhance the efficient production, exploitation, and handling of climate-relevant data for the future such as the GSV concept. The GSV provides a standardised representation of the climate data, facilitating homogeneity, consistency, and interoperability. All climate models deliver their output using the HEALPix grid, which offers numerous advantages for the analysis of high-resolution data, optimising the handling of spherical variables. A common list of

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

variables is delivered by all models at high frequency and close-to-native resolution with a strict data governance policy. Data is made available using the streaming concept, where data consumers can access the quality-controlled climate model output as the models run. At this point data and workflow become closely linked together. The unified end-to-end workflow facilitates a homogeneous data treatment with full traceability of its availability through the generalised access to the data notifier tasks. To cope with the large volume of data generated, the OPAs offer efficient data processing as the climate models perform the simulations, reducing computational overhead and improving responsiveness to the data consumer requirements. FAIR data principles (Wilkinson et al., 2016) are at the core of the data governance and management to facilitate data dissemination and analysis.

- AI integration: AI technologies are an emerging activity of the Climate DT. It aims to both enhance its interactivity capabilities and improve the user experience in a similar way as it is suggested in the Earth Virtualization Engines initiative (Stevens et al., 2024). AI-powered chatbots are being developed to ease the interactive engagement and exploitation of the Climate DT data (Koldunov and Jung, 2024). The chatbots ease access to climate insights based on the Climate DT climate data and indicators, addressing the requirements of equity and salience. The DTE enables the creation of AI-ready datasets to also support the development of data-driven climate emulators.
- Real-time monitoring and model evaluation: The Climate DT workflow includes two different tools to provide an online model assessment. AQUA is a software framework that delivers a continuous quality assessment of the most relevant climate aspects, by leveraging state-of-the-art technologies to ingest and process high-resolution climate data and to analyse them with multiple model evaluation diagnostics addressing, for example, model biases, performance indices and measures of variability such as teleconnection indices. OBSALL provides a full-resolution trace of the climate simulation in observation space. This is currently done for data from the SYNOP, TEMP, and AMSU-A networks, allowing thus on-line monitoring of the simulation quality as well as posterior diagnosis with respect to observations. The monitoring results from both AQUA and OBSALL are summarised via a dashboard used by operators and climate model developers to assess the quality of the simulations while they are still running.
- Action-oriented climate information: The Climate DT has embedded climate-impact applications in the workflow. They act as data consumers that produce climate-related indicators relevant for climate adaptation. Current examples generate indicators for renewable energy, and water and forest management. They have been co-designed with selected users with the aim of generating salient climate information. The applications use the high-frequency, high-spatial resolution of the Climate DT data taking advantage of the streaming capabilities to provide indicators with unprecedented spatial resolution. The tiered data protocol that makes available either part or all of the simulation results through the DestinE platform allows other applications to use the Climate DT data offline. In every case, the information sources are accompanied by quality (scientific and technical) estimates, an essential element to build trust to generate credibility (Dee et al., 2024), while the uncertainty component requires additional work to move beyond the multi-model estimates currently available.

The Climate DT ambitions respond to many of the objectives of the World Climate Research Programme and WMO's Scientific Advisory Panel[29] strategic vision. The Climate DT is an effort complementary to the sources offered by initiatives like CMIP and CORDEX, among others. These initiatives, which are evolving to work with faster cycles (e.g., Dunne et al., 2025), have different business models and governance, but provide invaluable opportunities to exchange experiences and solutions. The exchanges can follow the WMO concept of research to operations and operations to research transfers. Synergies of the Climate DT with those initiatives require mutual recognition, identification of strengths and synergies, and continuous efforts to develop salient, equitable, and credible operational solutions for climate-sensitive sectors.

The Climate DT, while computationally more expensive than existing climate simulation approaches, is a trailblazer for the use of computational resources so far largely unused by part of the climate community, such as GPU-based platforms. It provides a necessary step for the delivery of equitable and timely climate information for climate adaptation anywhere on the globe. The flexibility in the simulation definition and data availability are highly innovative aspects of the Climate DT operational ambition. The climate information is co-designed and co-produced with its users, while the interactivity uncovers new capabilities for users of this new climate information source.

Code and data availability

The Climate DT dataset is accessible via https://doi.org/10.21957/d3f982672e. There is a unique semantic data access for the data that allows users to clearly delineate the contributing models, experiment, dates, and variables, among other parameters. The availability policy of datasets in the data lake is defined in https://destine-data-lake-docs.data.destination-earth.eu/en/latest/dedl-discovery-and-data-access/DestinE-Data-Policy-for-DestinE-Digital-Twin-Outputs/DestinE-Data-Policy-for-DestinE-Digital-Twin-Outputs.html.

The ICON source code is available under https://gitlab.dkrz.de/icon/icon-model with doi:10.35089/wdcc/iconrelease2025.04, including all modifications used for this study. It can be used without any restrictions. The version of FESOM2.5 used in the Climate DT simulations is available at https://doi.org/10.5281/zenodo.10225420. The FESOM2.5 model is also available from GitHub https://github.com/FESOM/fesom2. The IFS source code is available subject to a licence agreement with ECMWF. ECMWF member-state weather services and approved partners have granted access. The IFS code without modules for data assimilation is also available. IFS used version cycle 48r1, available under an openIFS licence (http://www.ecmwf.int/en/research/projects/openifs, OpenIFS licence, 2024) for educational and academic purposes, with corresponding modifications available at https://doi.org/10.5281/zenodo.10223577. The NEMO4.0 source code version used in the Climate-DT is available at https://doi.org/10.5281/zenodo.5566313. The interfaces to NEMO can be obtained from ECMWF on request under the licence described above. Autosubmit is available at https://github.com/BSC-ES/autosubmit and the version used in the manuscript is available at https://doi.org/10.5281/zenodo.1559052. The Climate DT workflow is available in https://github.com/DestinE-Climate-DT/Workflow, and the version used is archived available at

[29] https://community.wmo.int/en/activity-areas/scientific-advisory-panel

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

https://doi.org/10.5281/zenodo.15607598. The AQUA code is available at https://doi.org/10.5281/zenodo.14906075. The OBSALL and OPA versions used are available at https://doi.org/10.5281/zenodo.15628903 and https://doi.org/10.5281/zenodo.14591827, respectively. The code for the mHM model version used in the HydroLand application is available at https://doi.org/10.5281/zenodo.8279545, while the code for the energy application is at https://doi.org/10.5281/zenodo.15609171. The code developed for the wildfire FWI application is available from https://doi.org/10.5281/zenodo.15610494.

Scripts and data to reproduce the figures of this manuscript can be found at https://zenodo.org/records/15680547.

Author contributions

AM, AS, BF, BS, CH, CS, DK, FD, FZ, HJ, IS, JH, JKon, MAc, MC, NK, NW, PB, PBr, PD, PM, PO, PP, RR, SCa, SMi, ST and UH contributed to conceptualization; AC, ALu, AS, BF, CH, CS, DK, FD, FR, HJ, HT, IS, JH, JKon, MAc, MC, NK, NW, PBr, PM, PO, PP, RR, SCa, SMi, SN, SP and UH contributed to project administration; AA, AC, AJ, AL, AMe, AN, BJ, BR, CF, CH, CHa, CPe, DB, DN, DS, ED, EV, FD, FR, HJ, HT, IA, IH, IG, IS, IW, JE, JJ, JKo, JKon, JRa, KG, LK, LKo, LT, MAc, MC, MG, MK, MM, MN, MR, NK, NN, PBr, PD, PG, PO, PR, PS, RR, SCa, SG, SMi, SMu, SN, SP, ST, SW, TC, UB, UH, UTi, VW and VG contributed to the software; AC, AJ, AL, AMe, AN, BJ, BR, CF, CPe, DS, EV, FR, HP, IA, IH, IW, JE, JKo, KG, LK, LKo, LT, MAc, MC, MG, MN, NK, PBr, PR, SCa, SMi, ST, SW and VG contributed to investigation; AA, BJ, BN, FR, HJ, IG, JH, JRa, JSu, KG, KK, MN, MR, NK, NN, PD, PG, PO, SCa, SG, SMi and SS contributed to validation; AJ, AL, FD, HJ, JW, JKon, JRa, KG, LT, MC, MM, MN, OS, PD, PO and ST contributed to visualization; AA, AM, BN, CH, DN, ED, FD, HJ, IA, JH, JT, LK, LKo, LT, MAc, MC, MG, MN, NW, PBr, SMi and SW contributed to data curation; AN, FD, HJ, HT, JKon, KG, SN and SP contributed to writing – original draft; AL, AN, BF, CS, FD, FR, HJ, HT, IG, JKon, JRa, KG, LT, MAc, MC, NK, NW, PB, PBr, PD, PO, SMi, SN and ST contributed to writing – review & editing. All authors read and approved the final manuscript.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Acknowledgements

The authors would like to thank all their colleagues of the different participating institutes listed at https://destine.ecmwf.int/provider and from ECMWF for their collaboration to make this ambitious endeavour a reality. We also acknowledge the EuroHPC Joint Undertaking (JU) for awarding this project access to the EuroHPC supercomputer LUMI and MareNostrum5 through a EuroHPC JU Special Access call. The work presented in this paper has been produced in the context of the European Union's Destination Earth Initiative and relates to tasks entrusted by the European Union to the European Centre for Medium-Range Weather Forecasts implementing part of this Initiative with funding by the

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

European Union. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. Some of the authors have received support from the EERIE project (Grant Agreement No 101081383) funded by the European Union and some through WarmWorld (Better – 01LK2202A and Faster – 01LK2203A) funded by the German Federal Ministry of Education and Research. Some of the authors have received support of the GLORIA project (TED2021-129543B-I00) funded by the Spanish MICIN/AEI. This work was also supported by the S1 project: “Diagnosis and Metrics in Climate Models” of the Collaborative Research Centre TRR 181 “Energy Transfer in Atmosphere and Ocean”, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, project no. 274762653).

References

Abatzoglou, J. T., Williams, A. P., and Barbero, R.: Global Emergence of Anthropogenic Climate Change in Fire Weather Indices, Geophysical Research Letters, 46, 326–336, https://doi.org/10.1029/2018GL080959, 2019.

Acosta, M. C., Palomas, S., Paronuzzi Ticco, S. V., Utrera, G., Biercamp, J., Bretonniere, P.-A., Budich, R., Castrillo, M., Caubel, A., Doblas-Reyes, F., Epicoco, I., Fladrich, U., Joussaume, S., Kumar Gupta, A., Lawrence, B., Le Sager, P., Lister, G., Moine, M.-P., Rioual, J.-C., Valcke, S., Zadeh, N., and Balaji, V.: The computational and energy cost of simulation and storage for climate science: lessons from CMIP6, Geosci. Model Dev., 17, 3081–3098, https://doi.org/10.5194/gmd-17-3081-2024, 2024.

Alvarez Fanjul, E., Ciliberti, S., Pearlman, J., Wilmer-Becker, K., Bahurel, P., Ardhuin, F., Arnaud, A., Azizzadenesheli, K., Aznar, R., Bell, M., Bertino, L., Behera, S., Brassington, G., Calewaert, J. B., Capet, A., Chassignet, E., Ciavatta, S., Cirano, M., Clementi, E., Cornacchia, L., Cossarini, G., Coro, G., Corney, S., Davidson, F., Drevillon, M., Drillet, Y., Dussurget, R., El Serafy, G., Fearon, G., Fennel, K., Ford, D., Le Galloudec, O., Huang, X., Lellouche, J. M., Heimbach, P., Hernandez, F., Hogan, P., Hoteit, I., Joseph, S., Josey, S., Le Traon, P.-Y., Libralato, S., Mancini, M., Martin, M., Matte, P., McConnell, T., Melet, A., Miyazawa, Y., Moore, A. M., Novellino, A., O’Donncha, F., Porter, A., Qiao, F., Regan, H., Robert-Jones, J., Sanikommu, S., Schiller, A., Siddorn, J., Sotillo, M. G., Staneva, J., Thomas-Courcoux, C., Thupaki, P., Tonani, M., Garcia Valdecasas, J. M., Veitch, J., Von Schuckmann, K., Wan, L., Wilkin, J., Zhong, A., and Zufic, R.: Promoting best practices in ocean forecasting through an Operational Readiness Level, Front. Mar. Sci., 11, 1443284, https://doi.org/10.3389/fmars.2024.1443284, 2024.

Argyroudis, S. A., Mitoulis, S. A., Chatzi, E., Baker, J. W., Brilakis, I., Gkoumas, K., Vousdoukas, M., Hynes, W., Carluccio, S., Keou, O., Frangopol, D. M., and Linkov, I.: Digital technologies can enhance climate resilience of critical infrastructure, Climate Risk Management, 35, 100387, https://doi.org/10.1016/j.crm.2021.100387, 2022.

Athanase, M., Sánchez-Benítez, A., Goessling, H. F., Pithan, F., and Jung, T.: Projected amplification of summer marine heatwaves in a warming Northeast Pacific Ocean, Commun Earth Environ, 5, 53, https://doi.org/10.1038/s43247-024-01212-1, 2024.

Bauer, P., Stevens, B., and Hazeleger, W.: A digital twin of Earth for the green transition, Nat. Clim. Chang., 11, 80–83, https://doi.org/10.1038/s41558-021-00986-y, 2021a.

Bauer, P., Dueben, P. D., Hoefler, T., Quintino, T., Schulthess, T. C., and Wedi, N. P.: The digital revolution of Earth-system science, Nat Comput Sci, 1, 104–113, https://doi.org/10.1038/s43588-021-00023-0, 2021b.

Bauer, P., Dueben, P., Chantry, M., Doblas-Reyes, F., Hoefler, T., McGovern, A., and Stevens, B.: Deep learning and a changing economy in weather and climate prediction, Nat Rev Earth Environ, 4, 507–509, https://doi.org/10.1038/s43017-023-00468-z, 2023.

Baulenas, E., Bojovic, D., Urquiza, D., Terrado, M., Pickard, S., González, N., and Clair, A. L. St.: User Selection and Engagement for Climate Services Coproduction, Weather, Climate, and Society, 15, 381–392, https://doi.org/10.1175/WCAS-D-22-0112.1, 2023.

Beech, N., Rackow, T., Semmler, T., and Jung, T.: Exploring the ocean mesoscale at reduced computational cost with FESOM 2.5: efficient modeling strategies applied to the Southern Ocean, Geosci. Model Dev., 17, 529–543, https://doi.org/10.5194/gmd-17-529-2024, 2024.

Betts, R. A., Belcher, S. E., Hermanson, L., Klein Tank, A., Lowe, J. A., Jones, C. D., Morice, C. P., Rayner, N. A., Scaife, A. A., and Stott, P. A.: Approaching 1.5 °C: how will we know we've reached this crucial warming mark?, Nature, 624, 33–35, https://doi.org/10.1038/d41586-023-03775-z, 2023.

Bojovic, D., St. Clair, A. L., Christel, I., Terrado, M., Stanzel, P., Gonzalez, P., and Palin, E. J.: Engagement, involvement and empowerment: Three realms of a coproduction framework for climate services, Global Environmental Change, 68, 102271, https://doi.org/10.1016/j.gloenvcha.2021.102271, 2021.

Bracco, A., Brajard, J., Dijkstra, H. A., Hassanzadeh, P., Lessig, C., and Monteleoni, C.: Machine learning for the physics of climate, Nature Reviews Physics, 7, 6–20, https://doi.org/10.1038/s42254-024-00776-3, 2025.

Collins, M., Beverley, J. D., Bracegirdle, T. J., Catto, J., McCrystall, M., Dittus, A., Freychet, N., Grist, J., Hegerl, G. C., Holland, P. R., Holmes, C., Josey, S. A., Joshi, M., Hawkins, E., Lo, E., Lord, N., Mitchell, D., Monerie, P.-A., Priestley, M. D. K., Scaife, A., Screen, J., Senior, N., Sexton, D., Shuckburgh, E., Siegert, S., Simpson, C., Stephenson, D. B., Sutton, R., Thompson, V., Wilcox, L. J., and Woollings, T.: Emerging signals of climate change from the equator to the poles: new insights into a warming world, Front. Sci., 2, 1340323, https://doi.org/10.3389/fsci.2024.1340323, 2024.

Dee, D., Obregon, A., and Buontempo, C.: Are Our Climate Data Fit for Your Purpose?, Bulletin of the American Meteorological Society, 105, E1723–E1733, https://doi.org/10.1175/BAMS-D-23-0295.1, 2024.

DestinE: Destination Earth Climate DT dataset (Version 1), Destination Earth Data Lake (ECMWF/EUMETSAT), https://doi.org/10.21957/d3f982672e, 2025.

Doblas-Reyes, F. J. and others: Linking global to regional climate change, in: Climate Change 2021: The Physical Science Basis, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., and others, Cambridge University Press, Cambridge, UK, 1363–1512, https://doi.org/10.1017/9781009157896.012, 2021.

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

Doblas-Reyes, F. J., Lera St Clair, A., Baldissera Pacchetti, M., Checchia, P., Cortekar, J., Klostermann, J. E. M., Krauß, W., Muñoz, Á. G., Mysiak, J., Paz, J., Terrado, M., Villwock, A., Volarev, M., and Zorita, S.: Standardisation of equitable climate services by supporting a community of practice, Climate Services, 36, 100520, https://doi.org/10.1016/j.cliser.2024.100520, 2024.

Donahue, A. S., Caldwell, P. M., Bertagna, L., Beydoun, H., Bogenschutz, P. A., Bradley, A. M., Clevenger, T. C., Foucar, J., Golaz, C., Guba, O., Hannah, W., Hillman, B. R., Johnson, J. N., Keen, N., Lin, W., Singh, B., Sreepathi, S., Taylor, M. A., Tian, J., Terai, C. R., Ullrich, P. A., Yuan, X., and Zhang, Y.: To exascale and beyond—The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM), a performance portable global atmosphere model for cloud-resolving scales, J Adv Model Earth Syst, 16, e2024MS004314, https://doi.org/10.1029/2024MS004314, 2024.

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., Fuentes-Franco, R., Gröger, M., v. Hardenberg, J., Hieronymus, J., Karami, M. P., Keskinen, J.-P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., van Noije, T., Nolan, P., O'Donnell, D., Ollinaho, P., van den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F., Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., and Zhang, Q.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6, Geoscientific Model Development, 15, 2973–3020, https://doi.org/10.5194/gmd-15-2973-2022, 2022.

Dunn, R. J. H., Willett, K. M., Morice, C. P., and Parker, D. E.: Pairwise homogeneity assessment of HadISD, Clim. Past, 10, 1501–1522, https://doi.org/10.5194/cp-10-1501-2014, 2014.

Dunne, J. P., Hewitt, H. T., Arblaster, J., Bonou, F., Boucher, O., Cavazos, T., Durack, P. J., Hassler, B., Juckes, M., Miyakawa, T., Mizielinski, M., Naik, V., Nicholls, Z., O'Rourke, E., Pincus, R., Sanderson, B. M., Simpson, I. R., and Taylor, K. E.: An evolving Coupled Model Intercomparison Project phase 7 (CMIP7) and Fast Track in support of future climate assessment, https://doi.org/10.5194/egusphere-2024-3874, 2025.

ECMWF: ODB Overview, European Centre for Medium-Range Weather Forecasts (ECMWF), 2010.

ECMWF: IFS documentation – CY41r1, Part I: Observations, European Centre for Medium-Range Weather Forecasts (ECMWF), 2015.

ECMWF: Pyodc documentation, European Centre for Medium-Range Weather Forecasts (ECMWF), 2025.

EUCRA: European Climate Risk Assessment Report, EEA Report 01/2024, 2024.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

Eyring, V., Gentine, P., Camps-Valls, G., Lawrence, D. M., and Reichstein, M.: AI-empowered next-generation multiscale climate modelling for mitigation and adaptation, Nat. Geosci., 17, 963–971, https://doi.org/10.1038/s41561-024-01527-w, 2024.

Fiedler, T., Pitman, A. J., Mackenzie, K., Wood, N., Jakob, C., and Perkins-Kirkpatrick, S. E.: Business risk and the emergence of climate analytics, Nat. Clim. Chang., 11, 87–94, https://doi.org/10.1038/s41558-020-00984-6, 2021.

Good, S. A., Martin, M. J., and Rayner, N. A.: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, JGR Oceans, 118, 6704–6716, https://doi.org/10.1002/2013JC009067, 2013.

Gorski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., and Bartelmann, M.: HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere, ApJ, 622, 759–771, https://doi.org/10.1086/427976, 2005.

Grayson, K., Thober, S., Lacima-Nadolnik, A., Sharifi, E., Lledó, L., and Doblas-Reyes, F. J.: Statistical summaries for streamed data from climate simulations: One-pass algorithms (v0.6.2), https://doi.org/10.5194/egusphere-2025-28, 2025.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., Von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J.-J., Mao, J., Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., and Von Storch, J.-S.: High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6, Geosci. Model Dev., 9, 4185–4208, https://doi.org/10.5194/gmd-9-4185-2016, 2016.

Haine, C., Haus, U.-U., Martinasso, M., Pleiter, D., Tessier, F., Sarmany, D., Smart, S., Quintino, T., and Tate, A.: A Middleware Supporting Data Movement in Complex and Software-Defined Storage and Memory Architectures, in: High Performance Computing, Cham, 346–357, 2021.

Hallegatte, S.: Strategies to adapt to an uncertain climate change, Global Environmental Change, 19, 240–247, https://doi.org/10.1016/j.gloenvcha.2008.12.003, 2009.

Hanke, M., Redler, R., Holfeld, T., and Yastremsky, M.: YAC 1.2.0: new aspects for coupling software in Earth system modelling, Geosci. Model Dev., 9, 2755–2769, https://doi.org/10.5194/gmd-9-2755-2016, 2016.

Hazeleger, W., Aerts, J. P. M., Bauer, P., Bierkens, M. F. P., Camps-Valls, G., Dekker, M. M., Doblas-Reyes, F. J., Eyring, V., Finkenauer, C., Grundner, A., Hachinger, S., Hall, D. M., Hartmann, T., Iglesias-Suarez, F., Janssens, M., Jones, E. R., Kölling, T., Lees, M., Lhermitte, S., Van Nieuwpoort, R. V., Pahker, A.-K., Pellicer-Valero, O. J., Pijpers, F. P., Siibak, A., Spitzer, J., Stevens, B., Vasconcelos, V. V., and Vossepoel, F. C.: Digital twins of the Earth with and for humans, Commun Earth Environ, 5, 463, https://doi.org/10.1038/s43247-024-01626-x, 2024.

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., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hetzer, J., Forrest, M., Ribalaygua, J., Prado-López, C., and Hickler, T.: The fire weather in Europe: large-scale trends towards higher danger, Environmental Research Letters, 19, 084017, https://doi.org/10.1088/1748-9326/ad5b09, 2024.

Hewitt, C. D. and Stone, R.: Climate services for managing societal risks and opportunities, Climate Services, 23, 100240, https://doi.org/10.1016/j.cliser.2021.100240, 2021.

Hoffmann, J., Bauer, P., Sandu, I., Wedi, N., Geenen, T., and Thiemert, D.: Destination Earth – A digital twin in support of climate services, Climate Services, 30, 100394, https://doi.org/10.1016/j.cliser.2023.100394, 2023.

Hohenegger, C., Korn, P., Linardakis, L., Redler, R., Schnur, R., Adamidis, P., Bao, J., Bastin, S., Behravesh, M., Bergemann, M., Biercamp, J., Bockelmann, H., Brokopf, R., Brüggemann, N., Casaroli, L., Chegini, F., Datseris, G., Esch, M., George, G., Giorgetta, M., Gutjahr, O., Haak, H., Hanke, M., Ilyina, T., Jahns, T., Jungclaus, J., Kern, M., Klocke, D., Kluft, L., Kölling, T., Kornblueh, L., Kosukhin, S., Kroll, C., Lee, J., Mauritsen, T., Mehlmann, C., Mieslinger, T., Naumann, A. K., Paccini, L., Peinado, A., Praturi, D. S., Putrasahan, D., Rast, S., Riddick, T., Roeber, N., Schmidt, H., Schulzweida, U., Schütte, F., Segura, H., Shevchenko, R., Singh, V., Specht, M., Stephan, C. C., von Storch, J.-S., Vogel, R., Wengel, C., Winkler, M., Ziemen, F., Marotzke, J., and Stevens, B.: ICON-Sapphire: simulating the components of the Earth system and their interactions at kilometer and subkilometer scales, Geosci. Model Dev., 16, 779–811, https://doi.org/10.5194/gmd-16-779-2023, 2023.

Intergovernmental Panel On Climate Change (Ipcc): 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, 2023.

Jakob, C., Gettelman, A., and Pitman, A.: The need to operationalize climate modelling, Nat. Clim. Chang., 13, 1158–1160, https://doi.org/10.1038/s41558-023-01849-4, 2023.

John, A., Beyer, S., Athanase, M., Benítez, A. S., Goessling, H., Hossain, A., Nurisso, M., Aguridan, R., Andrés-Martínez, M., Gaya-Àvila, A., Cheedela, S. K., Geier, P., Ghosh, R., Hadade, I., Koldunov, N. V., Pedruzo-Bagazgoitia, X., Rackow, T., Sandu, I., Sidorenko, D., Streffing, J., Vitali, E., and Jung, T.: Global Storyline Simulations at the Kilometre-scale, https://doi.org/10.22541/essoar.173160166.64258929/v1, 2025.

Jones, C. G., Adloff, F., Booth, B. B. B., Cox, P. M., Eyring, V., Friedlingstein, P., Frieler, K., Hewitt, H. T., Jeffery, H. A., Joussaume, S., Koenigk, T., Lawrence, B. N., O'Rourke, E., Roberts, M. J., Sanderson, B. M., Séférian, R., Somot, S., Vidale, P. L., van Vuuren, D., Acosta, M., Bentsen, M., Bernardello, R., Betts, R., Blockley, E., Boé, J., Bracegirdle, T., Braconnot, P., Brovkin, V., Buontempo, C., Doblas-Reyes, F., Donat, M., Epicoco, I., Falloon, P., Fiore, S., Frölicher, T., Fučkar, N. S., Gidden, M. J., Goessling, H. F., Graversen, R. G., Gualdi, S., Gutiérrez, J. M., Ilyina, T., Jacob, D., Jones, C. D., Juckes, M., Kendon, E., Kjellström, E., Knutti, R., Lowe, J., Mizielinski, M., Nassisi, P., Obersteiner, M., Regnier, P., Roehrig, R., Salas y Mélia, D., Schleussner, C.-F., Schulz, M., Scoccimarro, E., Terray, L., Thiemann, H., Wood, R. A., Yang, S., and Zaehle, S.: Bringing it all together: science priorities for improved understanding of Earth system change and to support international climate policy, Earth System Dynamics, 15, 1319–1351, https://doi.org/10.5194/esd-15-1319-2024, 2024.

Junghänel, T., Bär, F., Deutschländer, T., Haberlandt, U., Otte, I., Shehu, B., Stöckel, H., Stricker, K., Thiele, L.-B., and Willems, W.: Methodische Untersuchungen zur Novellierung der Starkregenstatistik für Deutschland (MUNSTAR). Synthesebericht, German Weather Service (DWD), Germany, 2022.

Kadow, C., Hall, D. M., and Ulbrich, U.: Artificial intelligence reconstructs missing climate information, Nature Geoscience, 13, 408–413, https://doi.org/10.1038/s41561-020-0582-5, 2020.

Koldunov, N. and Jung, T.: Local climate services for all, courtesy of large language models, Commun Earth Environ, 5, 13, https://doi.org/10.1038/s43247-023-01199-1, 2024.

Kotz, M., Levermann, A., and Wenz, L.: The effect of rainfall changes on economic production, Nature, 601, 223–227, https://doi.org/10.1038/s41586-021-04283-8, 2022.

Kruk, M. C., Parker, B., Marra, J. J., Werner, K., Heim, R., Vose, R., and Malsale, P.: Engaging with Users of Climate Information and the Coproduction of Knowledge, Weather, Climate, and Society, 9, 839–849, https://doi.org/10.1175/WCAS-D-16-0127.1, 2017.

Lauer, A., Bock, L., Hassler, B., Jöckel, P., Ruhe, L., and Schlund, M.: Monitoring and benchmarking Earth system model simulations with ESMValTool v2.12.0, Geosci. Model Dev., 18, 1169–1188, https://doi.org/10.5194/gmd-18-1169-2025, 2025.

Lean, P., Hólm, E. V., Bonavita, M., Bormann, N., McNally, A. P., and Järvinen, H.: Continuous data assimilation for global numerical weather prediction, Quart. J. Royal Meteor. Soc., 147, 273–288, https://doi.org/10.1002/qj.3917, 2021.

Lengfeld, K., Walawender, E., Winterrath, T., and Becker, A.: CatRaRE: A Catalogue of radar-based heavy rainfall events in Germany derived from 20 years of data, Meteor. Zeitschrift, 30, 469–487, https://doi.org/10.1127/metz/2021/1088, 2021.

Leo, S., Crusoe, M. R., Rodríguez-Navas, L., Sirvent, R., Kanitz, A., De Geest, P., Wittner, R., Pireddu, L., Garijo, D., Fernández, J. M., Colonnelli, I., Gallo, M., Ohta, T., Suetake, H., Capella-Gutierrez, S., De Wit, R., Kinoshita, B. P., and Soiland-Reyes, S.: Recording provenance of workflow runs with RO-Crate, PLoS ONE, 19, e0309210, https://doi.org/10.1371/journal.pone.0309210, 2024.

Leuridan, M., Hawkes, J., Smart, S., Danovaro, E., and Quintino, T.: Polytope: An Algorithm for Efficient Feature Extraction on Hypercubes, https://arxiv.org/abs/2306.11553, 2023.

Lledó, L., Torralba, V., Soret, A., Ramon, J., and Doblas-Reyes, F. J.: Seasonal forecasts of wind power generation, Renewable Energy, 143, 91–100, https://doi.org/10.1016/j.renene.2019.04.135, 2019.

Madonna, F., Tramutola, E., Sy, S., Serva, F., Proto, M., Rosoldi, M., Gagliardi, S., Amato, F., Marra, F., Fassò, A., Gardiner, T., and Thorne, P. W.: The New Radiosounding HARMonization (RHARM) Data Set of Homogenized Radiosounding Temperature, Humidity, and Wind Profiles With Uncertainties, JGR Atmospheres, 127, e2021JD035220, https://doi.org/10.1029/2021JD035220, 2022.

Manubens-Gil, D., Vegas-Regidor, J., Prodhomme, C., Mula-Valls, O., and Doblas-Reyes, F. J.: Seamless management of ensemble climate prediction experiments on HPC platforms, in: 2016 International Conference on High Performance Computing & Simulation (HPCS), 2016 International Conference on High Performance Computing & Simulation (HPCS), Innsbruck, Austria, 895–900, https://doi.org/10.1109/HPCSim.2016.7568429, 2016.

Moreno-Chamarro, E., Caron, L.-P., Ortega, P., Loosveldt Tomas, S., and Roberts, M. J.: Can we trust CMIP5/6 future projections of European winter precipitation?, Environ. Res. Lett., 16, 054063, https://doi.org/10.1088/1748-9326/abf28a, 2021.

Moreno-Chamarro, E., Caron, L.-P., Loosveldt Tomas, S., Vegas-Regidor, J., Gutjahr, O., Moine, M.-P., Putrasahan, D., Roberts, C. D., Roberts, M. J., Senan, R., Terray, L., Tourigny, E., and Vidale, P. L.: Impact of increased resolution on long-standing biases in HighResMIP-PRIMAVERA climate models, Geosci. Model Dev., 15, 269–289, https://doi.org/10.5194/gmd-15-269-2022, 2022.

Moreton, S. M., Ferreira, D., Roberts, M. J., and Hewitt, H. T.: Evaluating surface eddy properties in coupled climate simulations with 'eddy-present' and 'eddy-rich' ocean resolution, Ocean Modelling, 147, 101567, https://doi.org/10.1016/j.ocemod.2020.101567, 2020.

NOAA: Advanced Microwave Sounding Unit-A, National Oceanic and Atmospheric Administration, https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C01681, 2025

Nurisso, M., Caprioli, S., Davini, P., von Hardenberg, J., Nazarova, N., Ghosh, S., Ghinassi, P., Cadau, M., Tovazzi, E., Koldunov, N., Massonnet, F., Rajput, M. M., Sayed, S., Sharma, T., Sunny, J., Kluft, L., Kinoshita, B., and Ortega, P.: AQUA, https://doi.org/10.5281/zenodo.15044749, 2025.

Orlove, B.: The concept of adaptation, Annu. Rev. Environ. Resour., 47, 535–581, https://doi.org/10.1146/annurev-environ-112320-095719, 2022.

Palmer, T. N.: A personal perspective on modelling the climate system, Proc. R. Soc. A., 472, 20150772, https://doi.org/10.1098/rspa.2015.0772, 2016.

Pitman, A. J., Fiedler, T., Ranger, N., Jakob, C., Ridder, N., Perkins-Kirkpatrick, S., Wood, N., and Abramowitz, G.: Acute climate risks in the financial system: examining the utility of climate model projections, Environ. Res.: Climate, 1, 025002, https://doi.org/10.1088/2752-5295/ac856f, 2022.

Priestley, M. D. K., Dacre, H. F., Shaffrey, L. C., Hodges, K. I., and Pinto, J. G.: The role of serial European windstorm clustering for extreme seasonal losses as determined from multi-centennial simulations of high-resolution global climate model data, Nat. Hazards Earth Syst. Sci., 18, 2991–3006, https://doi.org/10.5194/nhess-18-2991-2018, 2018.

Rackow, T., Danilov, S., Goessling, H. F., Hellmer, H. H., Sein, D. V., Semmler, T., Sidorenko, D., and Jung, T.: Delayed Antarctic sea-ice decline in high-resolution climate change simulations, Nat Commun, 13, 637, https://doi.org/10.1038/s41467-022-28259-y, 2022.

Rackow, T., Koldunov, N., Lessig, C., Sandu, I., Alexe, M., Chantry, M., Clare, M., Dramsch, J., Pappenberger, F., Pedruzo-Bagazgoitia, X., Tietsche, S., and Jung, T.: Robustness of AI-based weather forecasts in a changing climate, https://arxiv.org/abs/2409.18529, 2024.

Rackow, T., Pedruzo-Bagazgoitia, X., Becker, T., Milinski, S., Sandu, I., Aguridan, R., Bechtold, P., Beyer, S., Bidlot, J., Boussetta, S., Deconinck, W., Diamantakis, M., Dueben, P., Dutra, E., Forbes, R., Ghosh, R., Goessling, H. F., Hadade, I., Hegewald, J., Jung, T., Keeley, S., Kluft, L., Koldunov, N., Koldunov, A., Kölling, T., Kousal, J., Kühnlein, C., Maciel, P., Mogensen, K., Quintino, T., Polichtchouk, I., Reuter, B., Sármány, D., Scholz, P., Sidorenko, D., Streffing, J., Sützl, B., Takasuka, D., Tietsche, S., Valentini, M., Vannière, B., Wedi, N., Zampieri, L., and Ziemen, F.: Multi-year simulations at kilometre scale with the Integrated Forecasting System coupled to FESOM2.5 and NEMOv3.4, Geoscientific Model Development, 18, 33–69, https://doi.org/10.5194/gmd-18-33-2025, 2025.

Rapella, L., Faranda, D., Gaetani, M., Drobinski, P., and Ginesta, M.: Climate change on extreme winds already affects off-shore wind power availability in Europe, Environmental Research Letters, 18, 034040, https://doi.org/10.1088/1748-9326/acbdb2, 2023.

Reichler, T. and Kim, J.: How Well Do Coupled Models Simulate Today's Climate?, Bulletin of the American Meteorological Society, 89, 303–312, https://doi.org/10.1175/BAMS-89-3-303, 2008.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195–204, https://doi.org/10.1038/s41586-019-0912-1, 2019.

Roberts, M. J., Reed, K. A., Bao, Q., Barsugli, J. J., Camargo, S. J., Caron, L.-P., Chang, P., Chen, C.-T., Christensen, H. M., Danabasoglu, G., Frenger, I., Fučkar, N. S., ul Hasson, S., Hewitt, H. T., Huang, H., Kim, D., Kodama, C., Lai, M., Leung, L.-Y. R., Mizuta, R., Nobre, P., Ortega, P., Paquin, D., Roberts, C. D., Scoccimarro, E., Seddon, J., Treguier, A. M., Tu, C.-Y., Ullrich, P. A., Vidale, P. L., Wehner, M. F., Zarzycki, C. M., Zhang, B., Zhang, W., and Zhao, M.: High-Resolution Model Intercomparison Project phase 2 (HighResMIP2) towards CMIP7, Geoscientific Model Development, 18, 1307–1332, https://doi.org/10.5194/gmd-18-1307-2025, 2025.

Rodrigues, R. R. and Shepherd, T. G.: Small is beautiful: climate-change science as if people mattered, PNAS Nexus, 1, pgac009, https://doi.org/10.1093/pnasnexus/pgac009, 2022.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resources Research, 46, 2008WR007327, https://doi.org/10.1029/2008WR007327, 2010.

Sánchez-Benítez, A., Goessling, H., Pithan, F., Semmler, T., and Jung, T.: The July 2019 European Heat Wave in a Warmer Climate: Storyline Scenarios with a Coupled Model Using Spectral Nudging, Journal of Climate, 35, 2373–2390, https://doi.org/10.1175/JCLI-D-21-0573.1, 2022.

Sandu, I.: Destination Earth's digital twins and Digital Twin Engine – state of play, ECMWF Newsletter, 14–23, https://doi.org/10.21957/is1fc736jx, 2024.

Sarmany, D., Valentini, M., Maciel, P., Geier, P., Smart, S., Aguridan, R., Hawkes, J., and Quintino, T.: MultIO: A Framework for Message-Driven Data Routing For Weather and Climate Simulations, in: Proceedings of the Platform for Advanced Scientific Computing Conference, PASC '24: Platform for Advanced Scientific Computing Conference, Zurich Switzerland, 1–12, https://doi.org/10.1145/3659914.3659938, 2024.

Schubert, J. E., Mach, K. J., and Sanders, B. F.: National-Scale Flood Hazard Data Unfit for Urban Risk Management, Earth's Future, 12, e2024EF004549, https://doi.org/10.1029/2024EF004549, 2024.

Segura, H., Hohenegger, C., Wengel, C., and Stevens, B.: Learning by Doing: Seasonal and Diurnal Features of Tropical Precipitation in a Global-Coupled Storm-Resolving Model, Geophysical Research Letters, 49, e2022GL101796, https://doi.org/10.1029/2022GL101796, 2022.

Segura, H., Pedruzo-Bagazgoitia, X., Weiss, P., Müller, S. K., Rackow, T., Lee, J., Dolores-Tesillos, E., Benedict, I., Aengenheyster, M., Aguridan, R., Arduini, G., Baker, A. J., Bao, J., Bastin, S., Baulenas, E., Becker, T., Beyer, S., Bockelmann, H., Brüggemann, N., Brunner, L., Cheedela, S. K., Das, S., Denissen, J., Dragaud, I., Dziekan, P., Ekblom, M., Engels, J. F., Esch, M., Forbes, R., Frauen, C., Freischem, L., García-Maroto, D., Geier, P., Gierz, P., González-Cervera, Á., Grayson, K., Griffith, M., Gutjahr, O., Haak, H., Hadade, I., Haslehner, K., ul Hasson, S., Hegewald, J., Kluft, L., Koldunov, A., Koldunov, N., Kölling, T., Koseki, S., Kosukhin, S., Kousal, J., Kuma, P., Kumar, A. U., Li, R., Maury, N., Meindl, M., Milinski, S., Mogensen, K., Niraula, B., Nowak, J., Praturi, D. S., Proske, U., Putrasahan, D., Redler, R., Santuy, D., Sármány, D., Schnur, R., Scholz, P., Sidorenko, D., Spät, D., Sützl, B., Takasuka, D., Tompkins, A., Uribe, A., Valentini, M., Veerman, M., Voigt, A., Warnau, S., Wachsmann, F., Wacławczyk, M., Wedi, N., Wieners, K.-H., Wille, J., Winkler, M., Wu, Y., Ziemen, F., Zimmermann, J., Bender, F. A.-M., Bojovic, D., Bony, S., Bordoni, S., Brehmer, P., Dengler, M., Dutra, E., Faye, S., Fischer, E., van Heerwaarden, C., Hohenegger, C., Järvinen, H., Jochum, M., et al.: nextGEMS: entering the era of kilometer-scale Earth system modeling, EGUsphere, 2025, 1–39, https://doi.org/10.5194/egusphere-2025-509, 2025.

Shaw, T. A., Arias, P. A., Collins, M., Coumou, D., Diedhiou, A., Garfinkel, C. I., Jain, S., Roxy, M. K., Kretschmer, M., Leung, L. R., Narsey, S., Martius, O., Seager, R., Shepherd, T. G., Sörensson, A. A., Stephenson, T., Taylor, M., and Wang, L.: Regional climate change: consensus, discrepancies, and ways forward, Front. Clim., 6, 1391634, https://doi.org/10.3389/fclim.2024.1391634, 2024.

Slingo, J., Bates, P., Bauer, P., Belcher, S., Palmer, T., Stephens, G., Stevens, B., Stocker, T., and Teutsch, G.: Ambitious partnership needed for reliable climate prediction, Nat. Clim. Chang., 12, 499–503, https://doi.org/10.1038/s41558-022-01384-8, 2022.

Soares, P. M. M., Careto, J. A. M., Cardoso, R. M., Goergen, K., Katragkou, E., Sobolowski, S., Coppola, E., Ban, N., Belušić, D., Berthou, S., Caillaud, C., Dobler, A., Hodnebrog, Ø., Kartsios, S., Lenderink, G., Lorenz, T., Milovac, J., Feldmann, H., Pichelli, E., Truhetz, H., Demory, M. E., De Vries, H., Warrach-Sagi, K., Keuler, K., Raffa, M., Tölle, M., Sieck, K., and Bastin, S.: The added value of km-scale simulations to describe temperature over complex orography: the

CORDEX FPS-Convection multi-model ensemble runs over the Alps, Clim Dyn, 62, 4491–4514, https://doi.org/10.1007/s00382-022-06593-7, 2024.

Stevens, B.: A perspective on the future of CMIP, AGU Advances, 5, e2023AV001086, https://doi.org/10.1029/2023AV001086, 2024.

Stevens, B., Adami, S., Ali, T., Anzt, H., Aslan, Z., Attinger, S., Bäck, J., Baehr, J., Bauer, P., Bernier, N., Bishop, B., Bockelmann, H., Bony, S., Brasseur, G., Bresch, D. N., Breyer, S., Brunet, G., Buttigieg, P. L., Cao, J., Castet, C., Cheng, Y., Dey Choudhury, A., Coen, D., Crewell, S., Dabholkar, A., Dai, Q., Doblas-Reyes, F., Durran, D., El Gaidi, A., Ewen, C., Exarchou, E., Eyring, V., Falkinhoff, F., Farrell, D., Forster, P. M., Frassoni, A., Frauen, C., Fuhrer, O., Gani, S., Gerber, E., Goldfarb, D., Grieger, J., Gruber, N., Hazeleger, W., Herken, R., Hewitt, C., Hoefler, T., Hsu, H.-H., Jacob, D., Jahn, A., Jakob, C., Jung, T., Kadow, C., Kang, I.-S., Kang, S., Kashinath, K., Kleinen-von Königslöw, K., Klocke, D., Kloenne, U., Klöwer, M., Kodama, C., Kollet, S., Kölling, T., Kontkanen, J., Kopp, S., Koran, M., Kulmala, M., Lappalainen, H., Latifi, F., Lawrence, B., Lee, J. Y., Lejeun, Q., Lessig, C., Li, C., Lippert, T., Luterbacher, J., Manninen, P., Marotzke, J., Matsouoka, S., Merchant, C., Messmer, P., Michel, G., Michielsen, K., Miyakawa, T., Müller, J., Munir, R., Narayanasetti, S., Ndiaye, O., Nobre, C., Oberg, A., Oki, R., Özkan-Haller, T., Palmer, T., Posey, S., Prein, A., Primus, O., Pritchard, M., Pullen, J., Putrasahan, D., et al.: Earth Virtualization Engines (EVE), Earth Syst. Sci. Data, 16, 2113–2122, https://doi.org/10.5194/essd-16-2113-2024, 2024.

Streffing, J., Sidorenko, D., Semmler, T., Zampieri, L., Scholz, P., Andrés-Martínez, M., Koldunov, N., Rackow, T., Kjellsson, J., Goessling, H., Athanase, M., Wang, Q., Hegewald, J., Sein, D. V., Mu, L., Fladrich, U., Barbi, D., Gierz, P., Danilov, S., Juricke, S., Lohmann, G., and Jung, T.: AWI-CM3 coupled climate model: description and evaluation experiments for a prototype post-CMIP6 model, Geosci. Model Dev., 15, 6399–6427, https://doi.org/10.5194/gmd-15-6399-2022, 2022.

Tao, F. and Qi, Q.: Make more digital twins, Nature, 573, 490–491, https://doi.org/10.1038/d41586-019-02849-1, 2019.

Taylor, M., Caldwell, P. M., Bertagna, L., Clevenger, C., Donahue, A., Foucar, J., Guba, O., Hillman, B., Keen, N., Krishna, J., Norman, M., Sreepathi, S., Terai, C., White, J. B., Salinger, A. G., McCoy, R. B., Ruby Leung, L., Bader, D. C., Wu, D.: The simple cloud-resolving E3SM atmosphere model running on the Frontier exascale system, Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 7, 1-11, https://doi.org/10.1145/3581784.3627044, 2023.

Thober, S., Cuntz, M., Kelbling, M., Kumar, R., Mai, J., and Samaniego, L.: The multiscale routing model mRM v1.0: simple river routing at resolutions from 1 to 50 km, Geosci. Model Dev., 12, 2501–2521, https://doi.org/10.5194/gmd-12-2501-2019, 2019.

Touma, D., Stevenson, S., Lehner, F., and Coats, S.: Human-driven greenhouse gas and aerosol emissions cause distinct regional impacts on extreme fire weather, Nature Communications, 12, 212, https://doi.org/10.1038/s41467-020-20570-w, 2021.

https://doi.org/10.5194/egusphere-2025-2198
Preprint. Discussion started: 13 August 2025

Uruchi, W., Castrillo, M., and Beltrán, D.: Autosubmit GUI: A Javascript-based Graphical User Interface to Monitor Experiments Workflow Execution, JOSS, 6, 3049, https://doi.org/10.21105/joss.03049, 2021.

Virman, M., Bister, M., Sinclair, V. A., Järvinen, H., and Räisänen, J.: A New Mechanism for the Dependence of Tropical Convection on Free-Tropospheric Humidity, Geophysical Research Letters, 45, 2516–2523, https://doi.org/10.1002/2018GL077032, 2018.

Virman, M., Bister, M., Räisänen, J., Sinclair, V. A., and Järvinen, H.: Radiosonde comparison of ERA5 and ERA-Interim reanalysis datasets over tropical oceans, Tellus A: Dynamic Meteorology and Oceanography, 73, 1929752, https://doi.org/10.1080/16000870.2021.1929752, 2021.

Wedi, N., Bauer, P., Sandu, I., Hoffmann, J., Sheridan, S., Cereceda, R., Quintino, T., Thiemert, D., and Geenen, T.: Destination Earth: High-performance computing for weather and climate, Comput. Sci. Eng., 24, 29–37, https://doi.org/10.1109/MCSE.2023.3260519, 2022.

Wedi, N., Sandu, I., Bauer, P., Acosta, M., Andersen, R. C., Andrae, U., Auger, L., Balsamo, G., Baousis, V., Bennett, V., Bennett, A., Buontempo, C., Bretonnière, P.-A., Capell, R., Castrillo, M., Chantry, M., Chevallier, M., Correa, R., Davini, P., Denby, L., Doblas-Reyes, F.J., Dueben, P., Fischer, C., Frauen, C., Frogner, I.-L., Früh, B., Gascón, E., Gérard, E., Gorwits, O., Geenen, T., Grayson, K., Guenova-Rubio, N., Hadade, I., von Hardenberg, J., Haus, U.-U., Hawkes, J., Hirtl, M., Hoffmann, J., Horvath, K., Järvinen, H., Jung, T., Kann, A., Klocke, D., Koldunov, N., Kontkanen, J., Sievi-Korte, O., Kristiansen, J., Kuwertz, E., Mäkelä, J., Maljutenko, I., Manninen, P., McKnight, U. S., Milinski, S., Mueller, A., McNally, A., Modigliani, U., Narayanappa, D., Nielsen, K. P., Nipen, T., Nortamo, H., Peuch, V.-H., Polade, S., Quintino, T., Schicker, I., Reuter, B., Smart, S., Sleigh, M., Suttie, M., Termonia, P., Thober, S., Randriamampianina, R., Theeuwes, N., Thiemert, D., Vannière, B., Vannitsem, S., Wittmann, C., Yang, X., Pontaud, M., Stevens, B., Pappenberger, and F.: Implementing digital twin technology of the earth system in Destination Earth, Journal of the European Meteorological Society, 3, 100015, https://doi.org/10.1016/j.jemets.2025.100015, 2025.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 'T Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., Van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., Van Der Lei, J., Van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B.: The FAIR Guiding Principles for scientific data management and stewardship, Sci Data, 3, 160018, https://doi.org/10.1038/sdata.2016.18, 2016.

Wright, L. and Davidson, S.: How to tell the difference between a model and a digital twin, Adv. Model. and Simul. in Eng. Sci., 7, 13, https://doi.org/10.1186/s40323-020-00147-4, 2020.