Funded by the European Union **Destination Earth** implemented by ECMWF esa EUMETSAT

Contact
Carmen Piñero-Megías
carmen.pinero@bsc.es

The Climate DT workflow: orchestrating multi-model climate simulations across heterogeneous HPC platforms

Carmen Piñero-Megías[1], Damien McClain[1], Egor Tiavlovsky[1], Francesc Roura-Adserias[1], Ivan Alsina-Ferrer[1], Leo Arriola[1], Marc Romeu[1], Miguel Andrés-Martínez[2], Muhammad Shafeeque[2], Francesca Macchia[1]

[1] Barcelona Supercomputing Center (BSC) [2] Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (AWI)

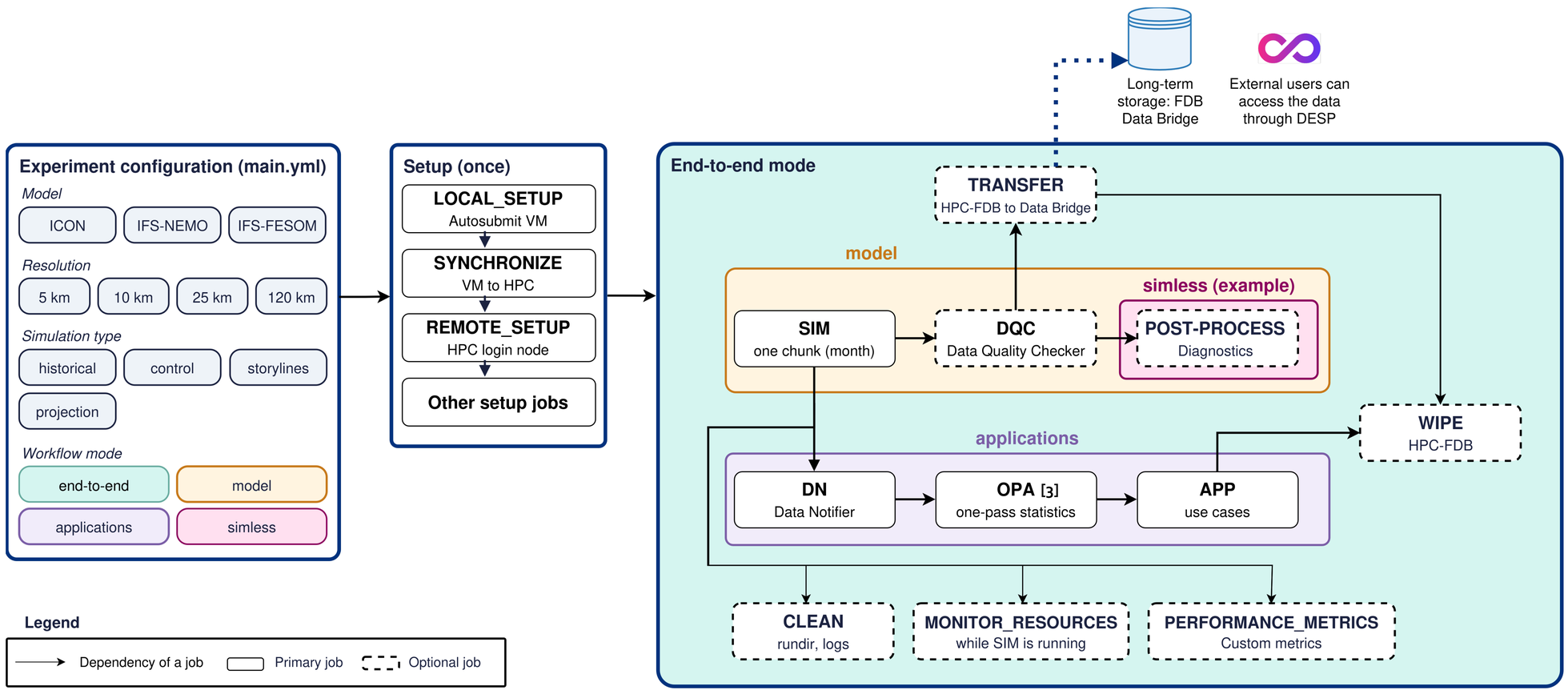
The challenge of Climate DT

- The **Destination Earth Climate Change Adaptation Digital Twin** (Climate DT) is a climate information system for producing **multi-decadal, multi-model global climate projections** and **translating climate data into climate impact information** to support adaptation efforts. [1]
- It runs three models (ICON, IFS-NEMO and IFS-FESOM) on **different EuroHPC machines**, such as LUMI and MareNostrum5, at **high temporal and spatial resolutions** (hourly, 5 km).
- **Model output is standardised** on a common grid for all models with the Generic State Vector (GSV), and then **streamed** into the impact-sector applications **while the model runs**. These applications cover use cases in energy, water, and forest management.

The workflow

We developed a **modular**, **flexible** and **portable** workflow that executes the climate simulations, data processing, applications, and data archiving on the remote machines. It contains the necessary **logic, job templates and configuration** to run the different components, which are then **orchestrated and monitored with Autosubmit.** [2] The workflow can run in different modes:

- **End-to-end:** operational production runs.
- **Model:** research and development experiments. Different simulation types can be specified: historical, control, storylines and projections.
- **Applications:** use-case applications run offline, using pre-existing data.
- **Simless:** used for different purposes: cleaning, rerunning post-processing or turning Climate DT output into AI-ready datasets for emulators.



Workflow configuration and dependencies (simplified) [4]

Our workflow manager: AUTOSUBMIT

Autosubmit is an experiment manager, orchestrator and monitoring tool in one lightweight **open-source** package, **built with climate research needs in mind** [5]. It runs in a virtual machine, independent of the HPC platforms.

Some key Autosubmit functionalities that have helped build the Climate DT workflow are:

- **Configuring different job granularity:** members, chunks and splits.
- **Task aggregation [6][7]:** allows combining several workflow tasks into one allocation, reducing queueing times significantly and optimizing resources.
- **Recovery from connectivity or I/O failures**.
- The Autosubmit API provides **statistics and performance metrics** for telemetry, and the **Autosubmit GUI** facilitates monitoring for users.

Portability, reproducibility and traceability

- General task templates are kept separate from platform-specific libraries, so **introducing new models and platforms needs only minor adjustments**.
- Applications are packaged as **Singularity containers, deployed automatically** on both EuroHPC machines and portable to future ones [8].
- Configurations and templates are **tracked in Git** and the workflow manages their **automatic deployment** on the different machines.
- **Configuration and logs** for each task are **tracked**, facilitating debugging.
- Autosubmit can archive **full experiments as RO-Crate**, packaging each run with its metadata.
- Autosubmit facilitates testing and sharing previous experiments, with the possibility of making **test cases** and **copying previous experiments**, which helps us keep testing close to production.

Future tasks

- Maintain operations, add features and update models to next cycle versions.
- **Enable external usage** of the applications workflow.
- **Port to other machines** such as JUPITER.
- Generate **AI-ready datasets online,** alongside the simulations.
- Duplicate the Autosubmit VM to **isolate operations from development**.
- **Improve test coverage:** testing covers unit tests, a testing suite run through Autosubmit, and an improved pipeline through **Jacamar CI**.
- Improve general performance by **adjusting task aggregation and job granularity**.

References

[1] Doblas-Reyes, F. J., et al. (2026). Geosci. Model Dev., 19, 2821.
[2] Gaya-Àvila, A., et al. (2024). EGU General Assembly 2024, EGU24-2533.
[3] Grayson, K., et al. (2025). Geosci. Model Dev., 18, 5873.
[4] Climate DT Workflow documentation, v6.2.7. climate-dt-workflow.readthedocs.io
[5] Manubens-Gil, D., et al. (2016). HPCS 2016. doi:10.1109/HPCSim.2016.7568429
[6] Marciani, M. G., et al. (2025). Geosci. Model Dev., 18, 9709.
[7] Goitia, P., et al. (2026). Geosci. Model Dev., 19, 8501.
[8] Roura-Adserias, F., et al. (2025). EGU25-4466. doi:10.5194/egusphere-egu25-4466

We 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.

Earth Sciences Department
Barcelona Supercomputing Center
Centro Nacional de Supercomputación