AQUA: a novel quality assessment tool for km-scale simulations in the Destination Earth Climate Digital Twin - the core framework

Matteo Nurisso[1], Jost von Hardenberg[2], Silvia Caprioli[2], Supriyo Ghosh[3], Nikolay Koldunov[4], Bruno P. Kinoshita[3], Natalia Nazarova[2], Paolo Ghinassi[5], and Paolo Davini[1]

[1]Istituto di Scienze dell'Atmosfera e del Clima, Consiglio Nazionale delle Ricerche (CNR-ISAC), Corso Fiume 4, 10133 Torino, Italy (m.nurisso@isac.cnr.it)

[2]Department of Environment, Land and Infrastructure Engineering (DIATI) Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

[3]Barcelona Supercomputing Center (BSC), Plaça Eusebi Güell, 1-3, 08034 Barcelona, Spain

[4]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (AWI), Bremerhaven, Germany

[5]Istituto di Scienze dell'Atmosfera e del Clima, Consiglio Nazionale delle Ricerche (CNR-ISAC), Via Piero Gobetti 101, 40129 Bologna, Italy

Destination Earth (DestinE) is a major initiative by the European Commission aiming to create a highly accurate global digital twin of Earth. The Climate Adaptation Digital Twin in DestinE is an ambitious project of several different climate simulations at the km-scale producing a large amount of heavy dataset, difficult to access and analyse with standard data processing pipelines. Each project and each model produces data that may differ in format (NetCDF, GRIB, Zarr), structure and metadata, leading to the necessity of tweaks and complex pipelines in order to prepare data for analysis.

We thus introduce AQUA, an Application for Quality assessment and Uncertainty quAntification. AQUA is composed of a core engine facilitating data access, combined with a series of modular and independent diagnostics to be run continuously to monitor and evaluate climate simulations. In this contribution we present the core engine and its features.

Though many available suites already exist to analyse data from global climate models, AQUA has been specifically developed to deal with large km-scale datasets, with the goal of unifying and simplifying climate data access for all users. AQUA responds to the need for users to have the focus on the development of their data analysis, while datasets are found, retrieved and homogenised by an external tool to which they can connect their pipeline.

Developed in Python, leveraging the power of Dask and Xarray libraries, AQUA prioritises efficiency through lazy data access. Noteworthy is the utilisation of cdo for one-time weight computation, enhancing performances in regridding and averaging operations. A key strength lies in its ability to handle high-resolution, high-frequency data, loading into memory only when necessary. AQUA not only unifies and simplifies climate data access for users but also addresses the crucial need for responsive feedback to climate model developers.