24-228180: Development and Implementation of AI4Clouds for Enhanced Extreme DT Cloud Forecasts

Fernando Iglesias-Suarez

Predictia Intelligent Data Solutions (**Predictia**)
Consejo Superior de Investigaciones Científicas (**CSIC**)

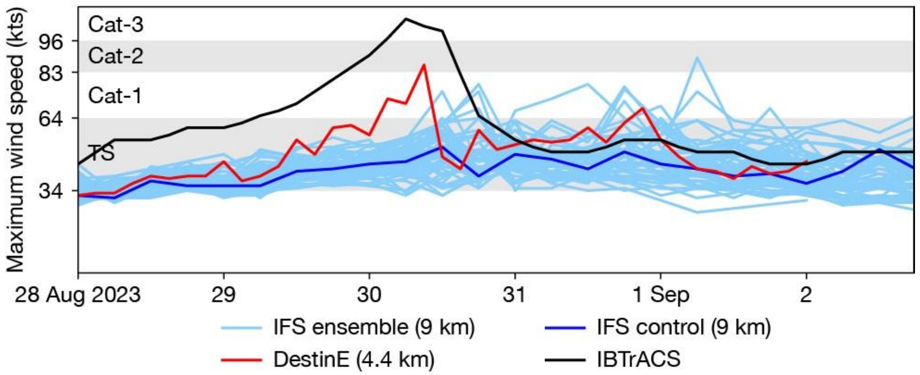
17 June 2025

AI4Clouds: Motivation

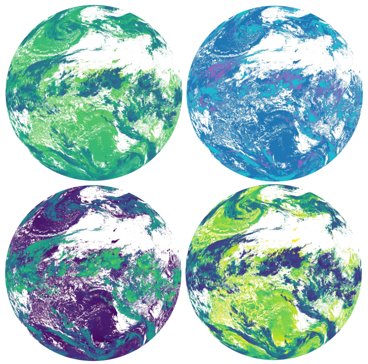
Destination Earth (**DestinE**) aims to create a digital twin of the Earth to enhance Europe's ability to respond and adapt to extreme weather (Extreme DT) and climate change (Climate DT).

The Weather-Induced **Extreme DT** will support decision-making for a rapid response to meteorological, hydrological and air quality extremes, on a timescale of a few days ahead.

Extreme DT relies on high-resolution simulations (4.4 km in the atmosphere) of the Integrated Forecasting System, coupled to NEMO ocean model, (**IFS-NEMO**).



Tropical cyclone Idalia forecasts. Figure adapted from ECMWF newsletter 178.



EUMETSAT's **satellite** (MSG/MTG) **cloud products**.

Can **artificial intelligence** be used to **fuse Extreme DT simulations and satellite observations** to deliver **enhanced short-term cloud forecasts** — with a concrete application in the solar energy sector?

AI4Clouds: At a Glance

Scope: Develop and implement an **artificial intelligence (AI) solution to enhance short-term (up to 12-hour) cloud forecasts for DestinE Extreme DT simulations** by merging satellite observations (SEVIRI/FCI) and Extreme DT, fully integrated into the DestinE Data Lake.

Goals:

- Provide accurate cloud forecasts (cloud cover, cloud optical depth & particle phase)
- Incorporate uncertainty representation via ensemble forecasting
- Deliver open-source models, code, and documentation
- Enable deployment and integration into DestinE operational ecosystem
- Demonstrate its potential for the energy sector (over the Iberian Peninsula) via a partnership with an end-user

Content

- Approach: end-to-end AI forecast system
- AI Framework: AnemoI for AI weather forecasting models
- Validation & MLOps/Infrastructure

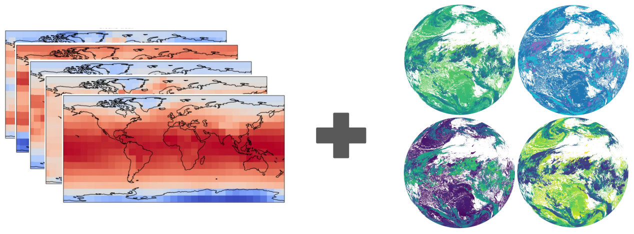
anemoi

HARMONISED DATA ACCESS

ISLET

AI4Clouds - Approach: Extreme DT-Satellites Mapping

Although **Extremes DT simulations can exhibit systematic biases compared to observations**, they are initialized from analysis fields that assimilate real-world data, and their high-resolution and short forecast horizons help limit error growth.



Outputs
(Extreme DT)

Clouds
(MSG/MTG)

How can we used them and improve clouds forecasts?

Bias correction typically adjusts model outputs by removing systematic errors:

- ★ Low complexity with well established approaches
- ★ High transparency understanding the biases that have been corrected
- ➔ Limited scope via statistical correction — missing structural deficiencies in simulated cloud processes, addressing regime-specific errors, or spatio-temporal mismatches
- ➔ No dynamic forecast learning only post-processing it — missing the potential to extract richer spatial-temporal patterns
- ➔ Extreme DT cloud fields currently not available (!)

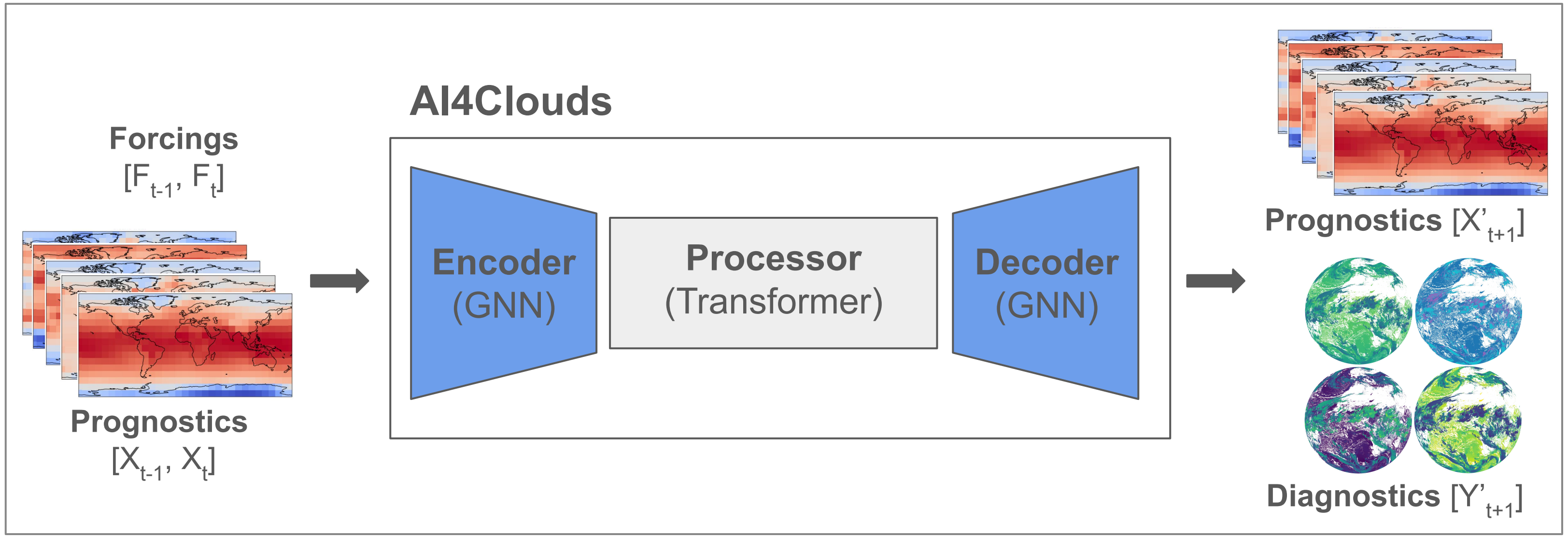
End-to-end AI forecast system, AI4Clouds, will be developed to implicitly capture and correct these biases through the mapping from Extreme DT thermodynamics (and, if available, clouds) to satellite-derived cloud fields (SEVIRI/FCI).

AI4Clouds - AI Framework: Anemol (I)

We will use Anemol framework to develop **AI4Clouds**, an AI-based cloud emulator that merges Extreme DT (predictors) and EUMETSAT satellite cloud observations (target) for short-term cloud forecasts (up to 12-hour).

Anemol is a framework developed by ECMWF and different meteorological agencies in Europe (e.g., NMI or KNMI) largely adopted within DestinE for developing AI weather and climate forecasting models (Lang et al., 2024a; Nipen et al., 2024). It comprises different components for preparing training datasets, conducting ML model training and inference.

Modelling (Anemol-core)



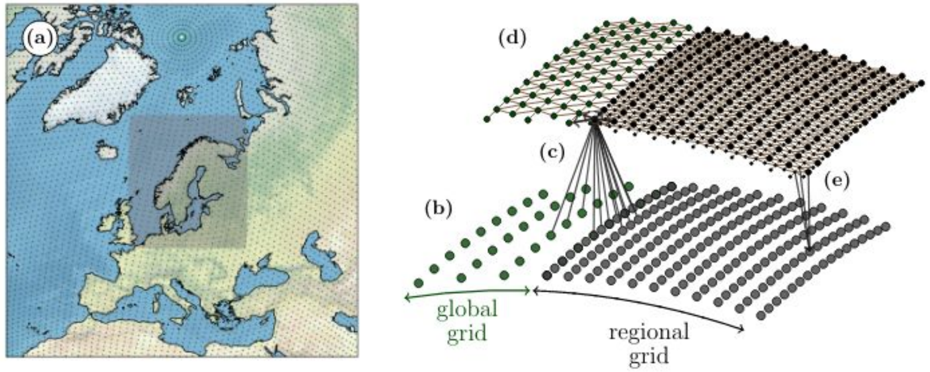
AI4Clouds specifications:

- Stretched-grid or limited area modeling (LAM)
- Forecast uncertainty: *Deterministic model* via ensemble forecasting or *Stochastic model* (Lang et al., 2024b) that learns a distribution of the outputs while preserving uncertainty

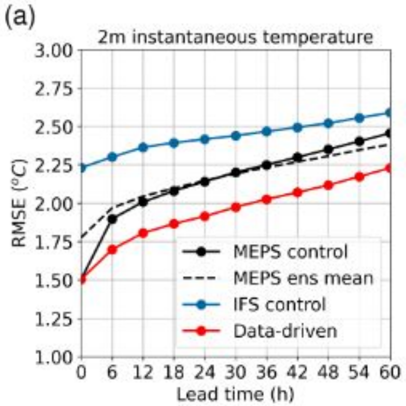
AI4Clouds - AI Framework: Setup (II)

Based on personal communication and by attending DestinE and other events, such as AnemoI Community meetings, there is already empirical evidence about the strengths of **stretched-grid and LAM approaches**:

- Competitive and **similar performance**, including computational costs;
- but **stretched-grid seems to be easier to optimise** with better large-scale circulation patterns emulation



Stretched-grid. Adapted from Fig. 1 in Nipen et al. (2024).



2m temperature. Fig. 5a in Nipen et al. (2024).

There is active research to account for forecast uncertainties via a **stochastic approach** (Lang et al., 2024b), but current evidence suggests that a **deterministic approach** may be more sensible until the first is more mature.

AI4Clouds: Validation & MLOps/Infrastructure

As evaluation framework, we adopt **AQUA** (Application for Quality assessment and Uncertainty quAntification), a Python-based tool designed to facilitate **access, processing, and analysis** of extensive output generated by DestinE DT simulations. Currently, AQUA includes a number of diagnostics for climate time-scales (e.g., means and distributions), and we will expand it with **additional AI metrics and weather-scale diagnostics** (e.g., rmse, acc and power spectra).

As infrastructure and services, we use DestinE Platform (DESP) and the various services, e.g.:

- **ISLET** (cloud computing and storage)
- **HDA** (Harmonised Data Access) to access DestinE and Federated data
- **EUMETSAT's Gitlab**, ensuring high code quality, security checks and unit testing

As MLOps tools, we use **MLFlow** to log training metrics and save model weights (model repository with versioning).

Work is in progress (!) to retrieve and align Extreme DT and satellite data, train AI4Clouds prototypes, and extend AQUA for evaluation — all in close collaboration with an end-user to ensure practical relevance for the energy sector.

Funded by the European Union

Thank you!

Questions are welcome.