



**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación



HPC challenges of running large ensemble simulations in a Tier-0 machine to model atmospheric desert dust

*F. Macchia, E. Di Tomaso, G. Montané, S. Basart, D. Beltrán,
W. Uruchi, A. Bunuel, M. Castrillo, J. Escribano, O. Jorba,
K. Serradell, C. Pérez García-Pando*

Barcelona Supercomputing Center

PRACEdays21

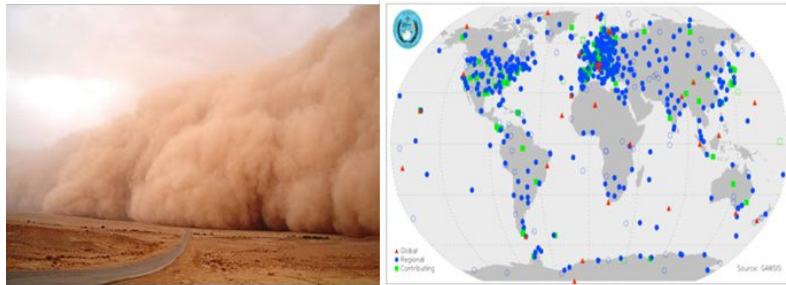
Dust and its impacts



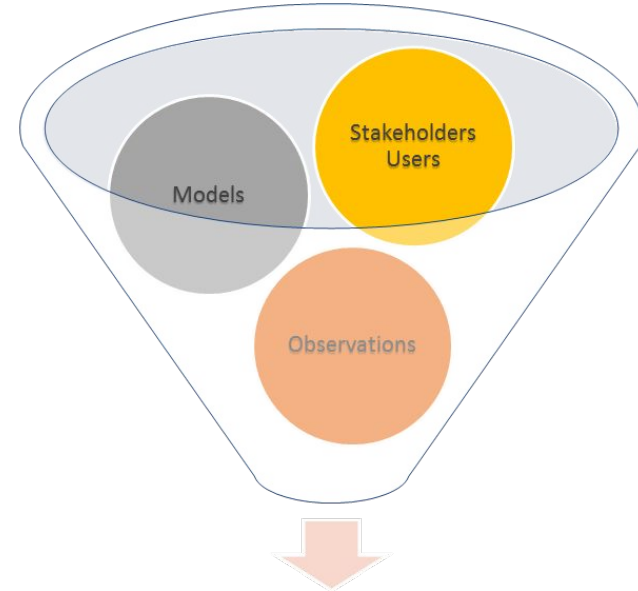
inDust leaflet: <https://cost-indust.eu/>

Dust Storms Assessment for the development of user-oriented Climate Services in Northern Africa, Middle East and Europe

- SDS is a serious hazard for life, health, environment and economy
- Lack of dust observations (past trends and current conditions)

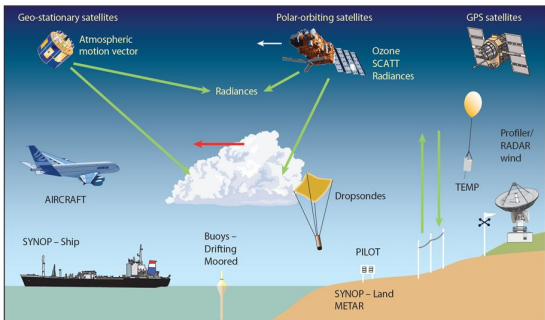


GOAL: Develop dust-related services to specific socio-economic sectors based on an advanced **dust reanalysis** for the NAMEE region



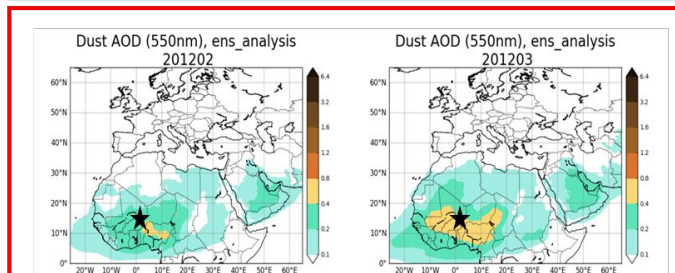
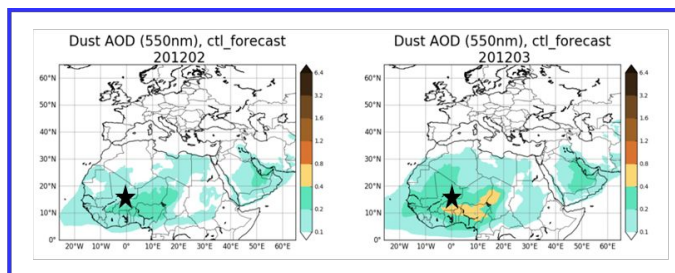
Dust-related Climate Services



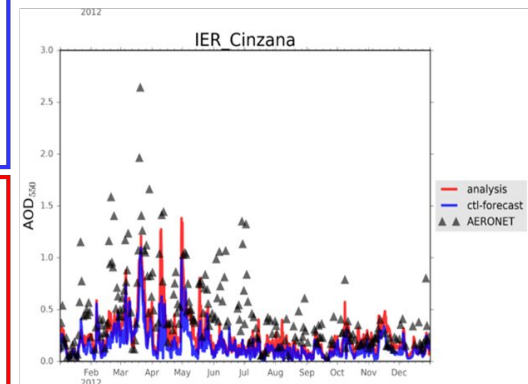


OBSERVATIONS: Our simulations, predictions and services are enhanced by/verified with an intensive use of in-situ and satellite **filtered-dust** observations

REANALISYS: Climatology (2007-2016) of sand and dust storms (SDS) in high spatial-temporal detail (10km, 3-hourly)



Assimilating MODIS DOD Land product (Courtesy P. Ginoux, NOAA)



(Courtesy E. Di Tomaso, BSC)

Development of a dust reanalysis

Dust ensemble forecasts are used at BSC to estimate **flow-dependent forecast uncertainty**, which is used by DA to optimally combine forecast with observations

The ensemble forecast has been designed considering model uncertainties with respect to:

- **surface winds**
- **soil humidity**
- **vertical flux distribution at sources**



The final MONARCH ensemble consider 12 members that are a combination of

- 3 emission schemes
 - Perez, Ginoux and Kok
- 2 global meteorological models (used as initial and boundary conditions)
 - MERRA and ERA
- 2 vertical flux distribution

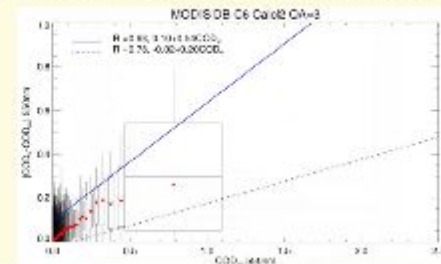
Assimilated observations

MODIS Deep Blue Coarse AOD

- *AE, ω filter, coarse AOD retrieval*
- *highest quality flag (Ginoux et al., 2012; Pu & Ginoux 2016)*
- *uncertainty model based on Sayer et al., 2014*



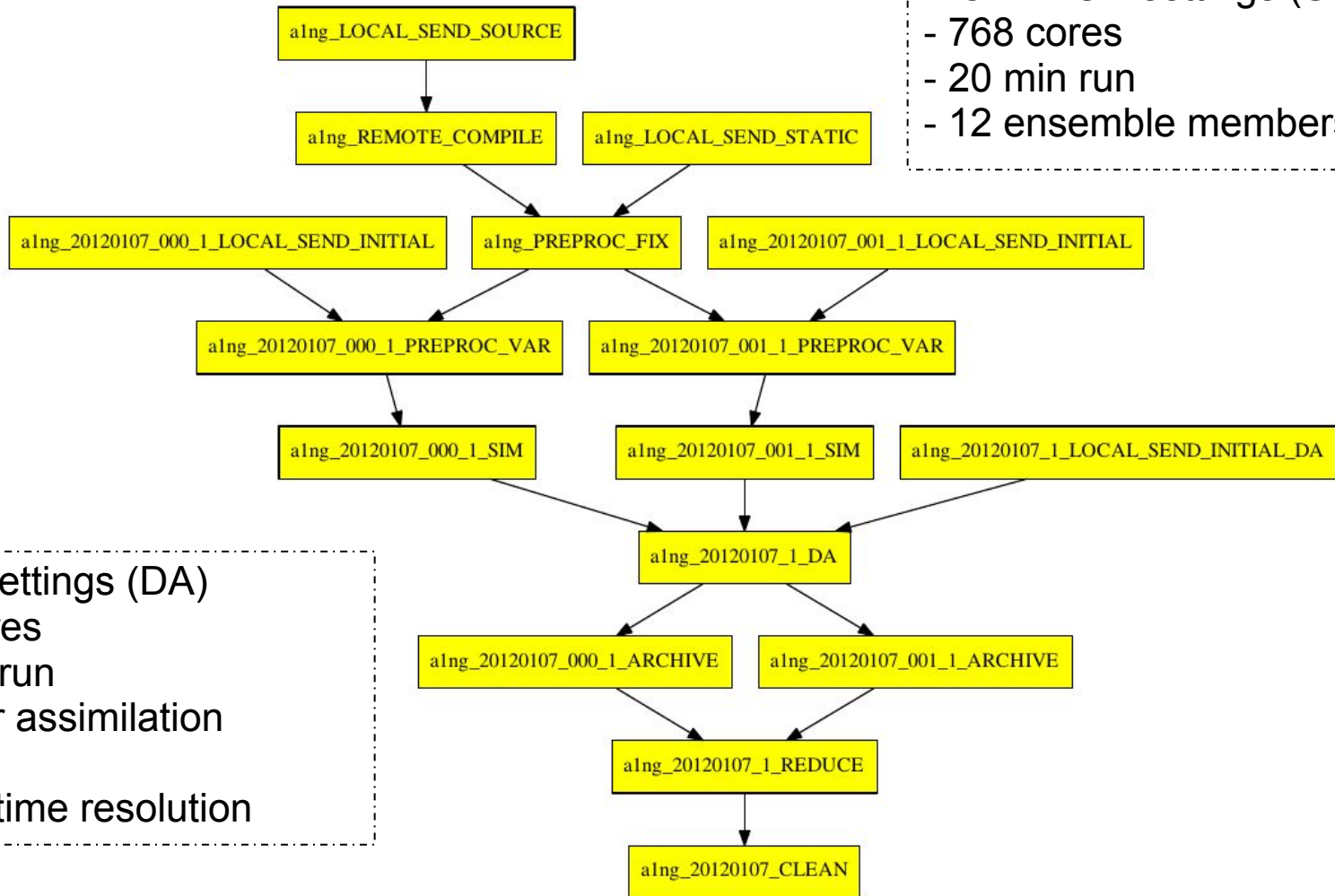
Observation uncertainty characterization



Basic workflow

MONARCH settings (SIM)

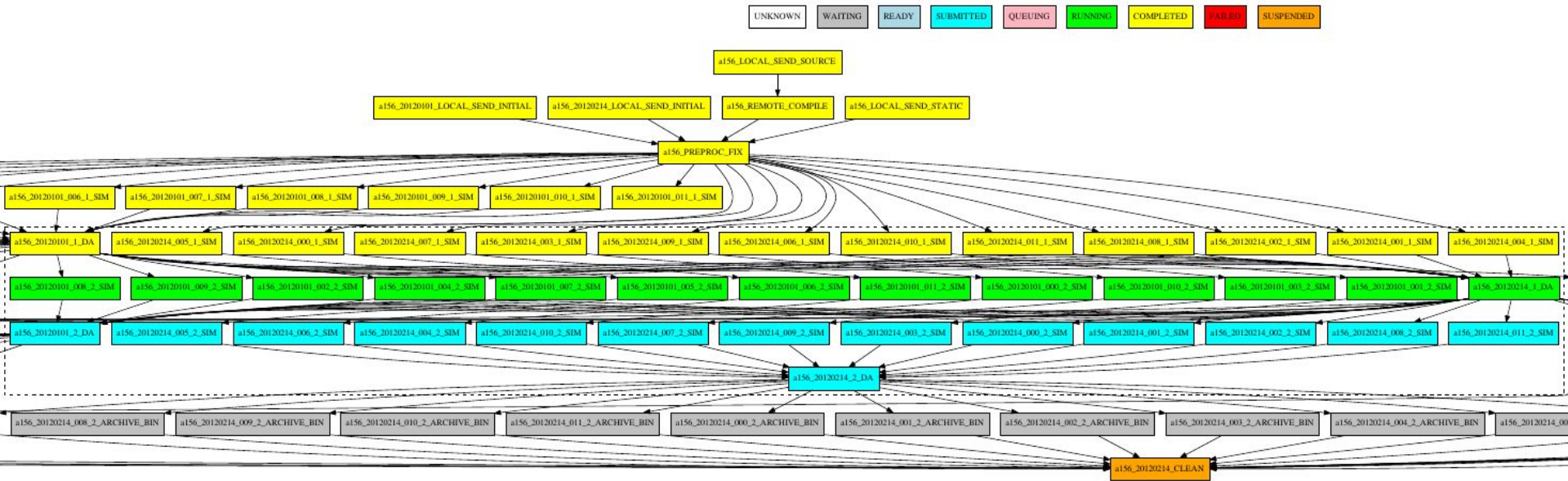
- 768 cores
- 20 min run
- 12 ensemble members



LETKF settings (DA)

- 576 cores
- 20 min run
- 24-hour assimilation window
- 3-hour time resolution

How to run 10-years reanalysis?



Access prace-ri.eu/hpc-access

- Call 14 - eDUST
- Call 17 - eFRAGMENT1
- Call 20 - eFRAGMENT2

What Is Autosubmit

Python-based **workflow manager** that allows to **create, manage & monitor** experiments remotely.

Features:

- **Automatization:** Orchestrating different kind of tasks and environments without intervention.
- **Provenance and reproducibility:** Unique experiment ids, storing all the parameters needed to reproduce them (Autosubmit version, model version, configuration, etc.). Linked with CVS.
- **Failure tolerance:** Automatic retrials and ability to rerun chunks in case of corrupted or missing data, repeating postprocessing and transfers if needed. Recovery capabilities.
- **Versatility:** Variety of workflows supported, from Auto-Models to data-centric or performance analysis, with different potentialities for each case.

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

<https://pypi.python.org/pypi/autosubmit>

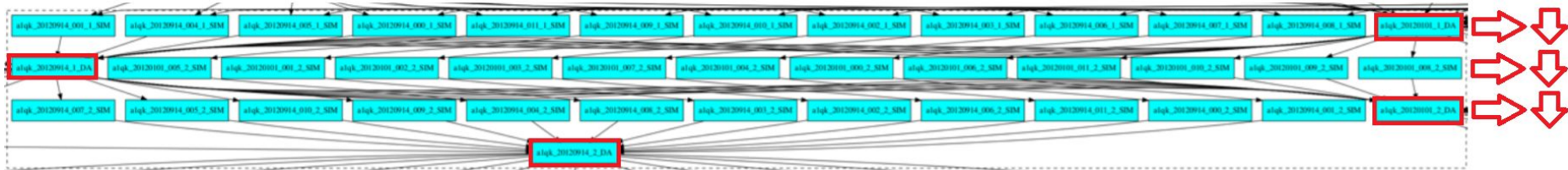
Get involved or contact us:

GitLab:	https://earth.bsc.es/gitlab/es/autosubmit
Mailing List:	autosubmit@bsc.es

Documentation:

Autosubmit:	http://autosubmit.readthedocs.io
FAQ:	https://autosubmit.readthedocs.io/en/latest/faq.html

Hybrid wrapper and crossing-date strategy



A single row (**1 day**) contains:

- **12 SIMs** of the first start date: $12 \times 768 = 9216$ cores
- **1 DA** of the second start date: **576** cores
- **12 post processing jobs**: $12 \times 4 = 48$ cores
- **10 jobs computing statistics**: $10 \times 48 = 480$ cores

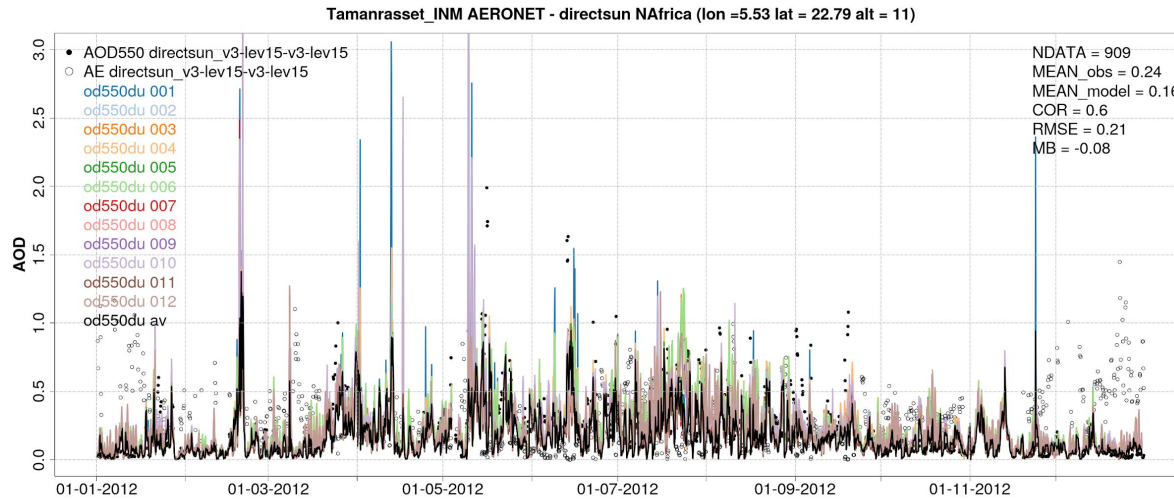
10320 cores
used for
20 hours

Wrapper with **30 days** for each start date

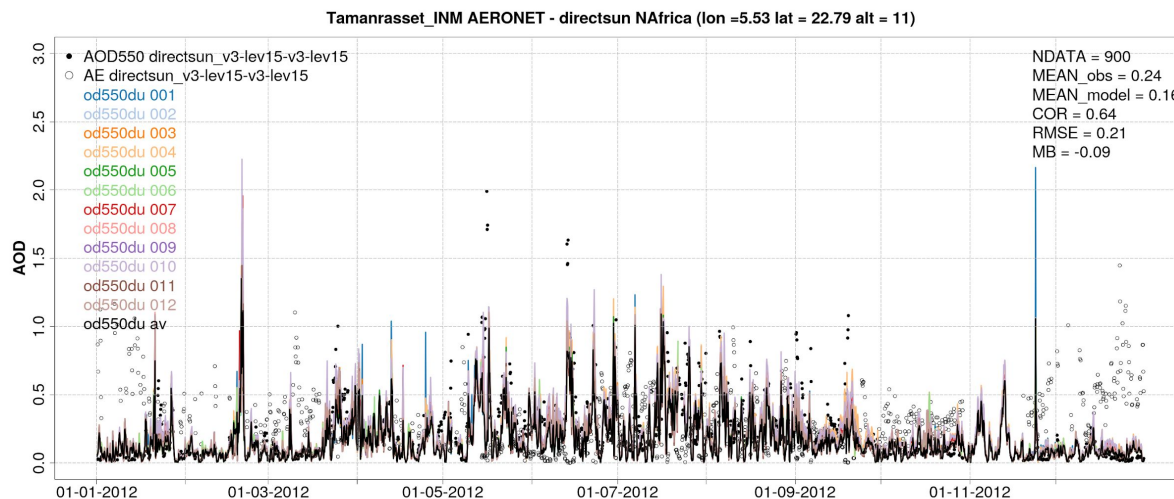
20 minutes to process 1 day

Results

Evaluation results for 2012: AERONET comparison



**FIRST GUEST
run**



**ANALYSIS
with assimilation**

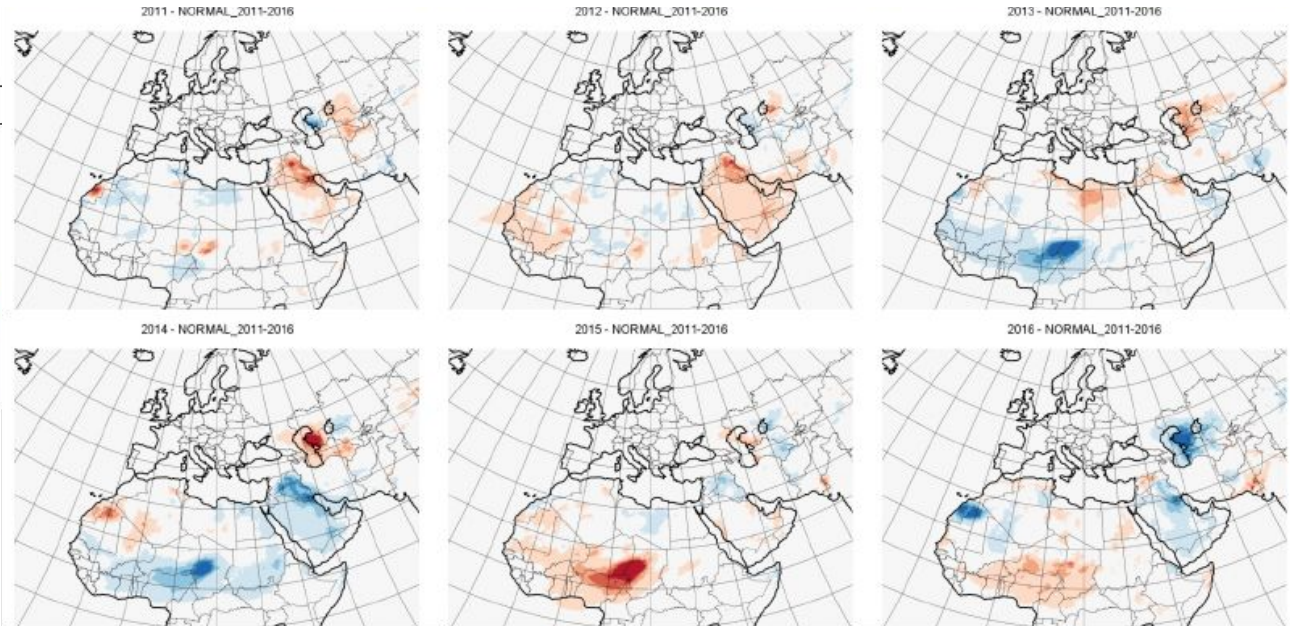
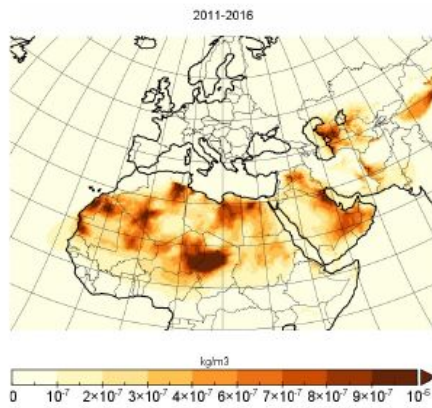
Outcomes and last remarks

Climatology (2011-2016): Surface concentration (in $\mu\text{g}/\text{m}^3$)

sconcdv ANOMALY



NORMAL YEAR, 2011-2016



Our activities are taking as a reference by the **UN SDS initiatives** that searches to **mitigate dust impacts**.



Thank you



The authors acknowledge the DustClim project which is part of ERA4CS, an ERA-NET initiated by JPI Climate, and funded by FORMAS (SE), DLR (DE), BMWFW (AT), IFD (DK), MINECO (ES), ANR (FR) with co-funding by the European Union (435690462); PRACE (eDUST, eFRAGMENT1, eFRAGMENT2); RES (AECT-2020-3-0013) for awarding access to MareNostrum at BSC and for technical support.