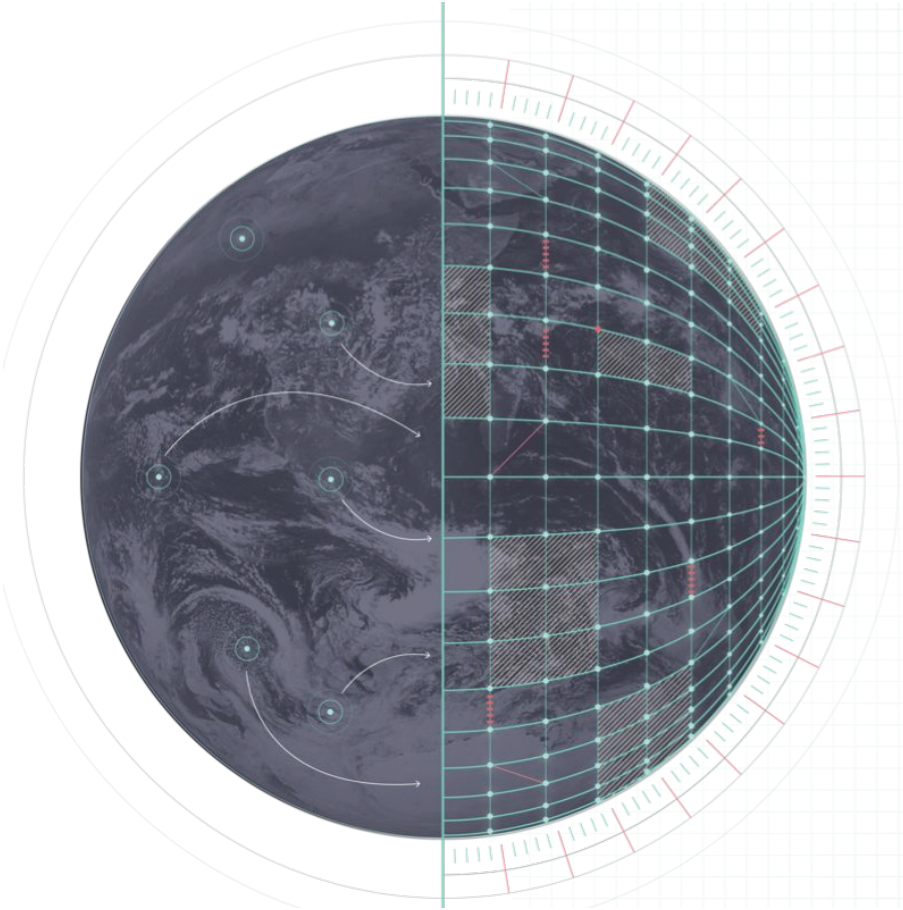


DESTINATION EARTH

Development of the on-demand extreme Digital Twin workflow

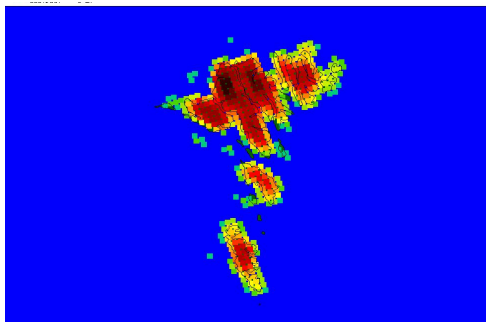
Xiaohua Yang, DMI
Ulf Andrae, SMHI
& On-demand DTE Workflow Team

PASC2023/MS4B, 28 June 2023, Davos

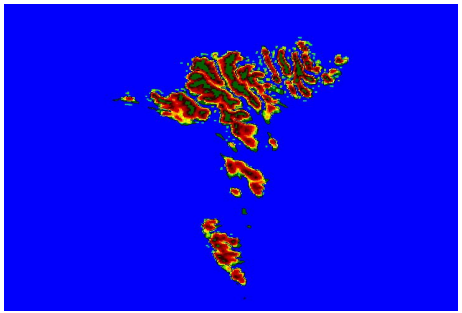
Funded by the European Union
Destination Earth
implemented by
ECMWF esa EUMETSAT

Recent years have seen promising developments in the mesoscale NWP which is advancing from km-scale to hectometric scale

Faroe Islands

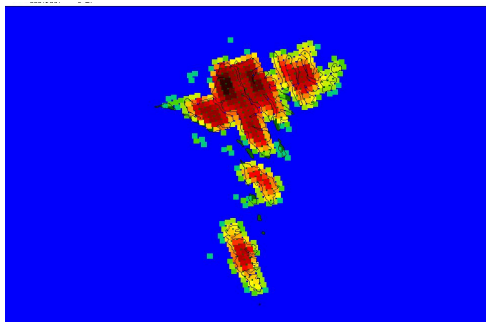


500 m grid

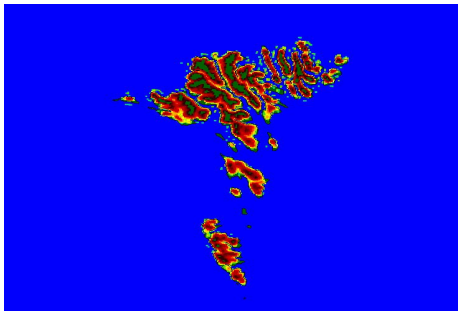


150 m grid

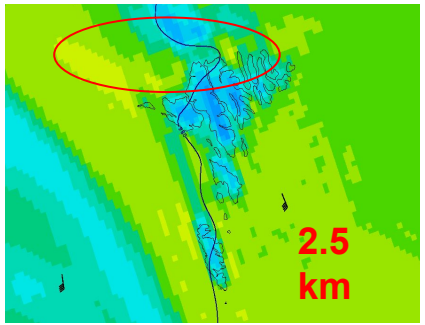
Recent years have seen promising developments in the mesoscale NWP which is advancing from km-scale to hectometric scale

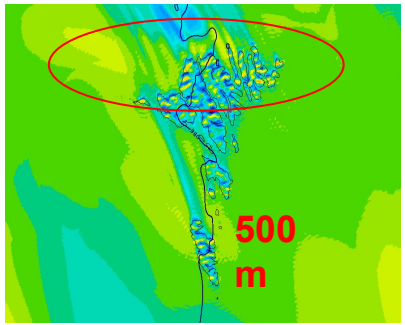


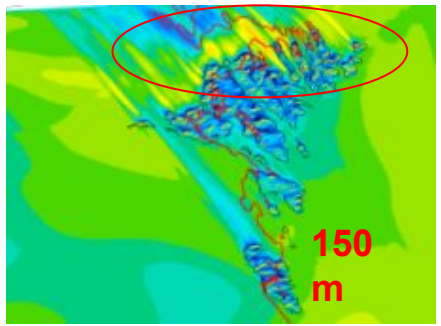
500 m grid



150 m grid





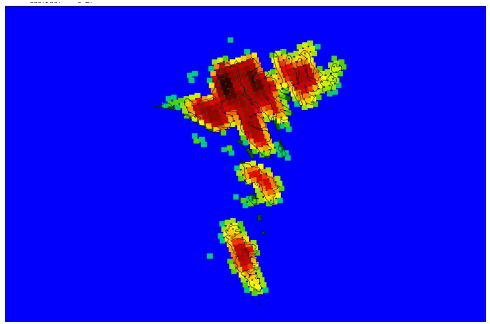


DMI-Harmonie

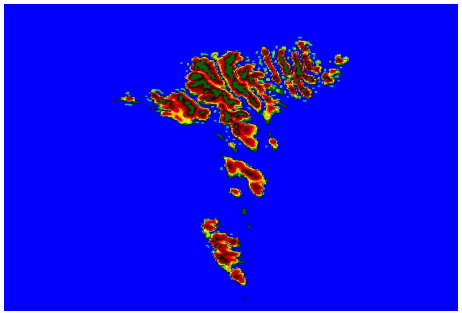
ECMWF

The hurricane-scale wind on 9 March 2021 which escaped operational predictions

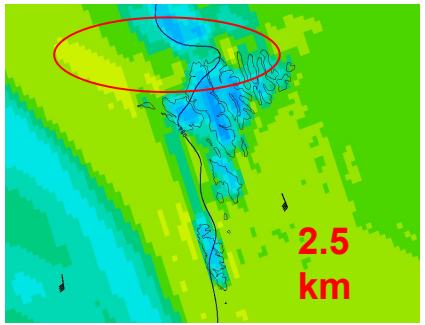
Will DE fast-forward the advances in the NWP capability?

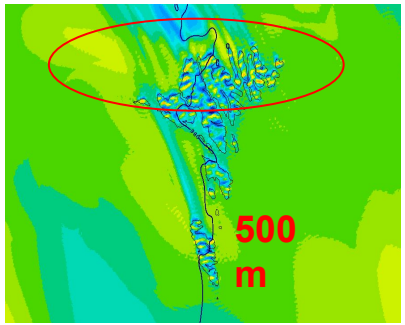


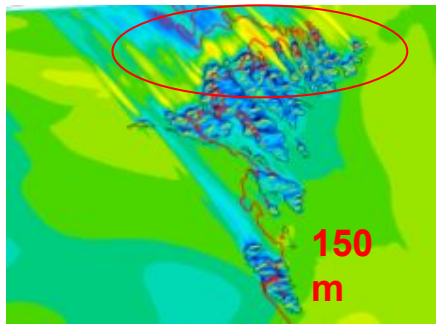
500 m grid



150 m grid



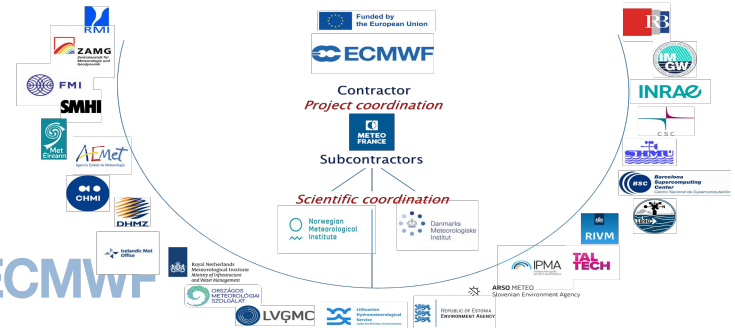


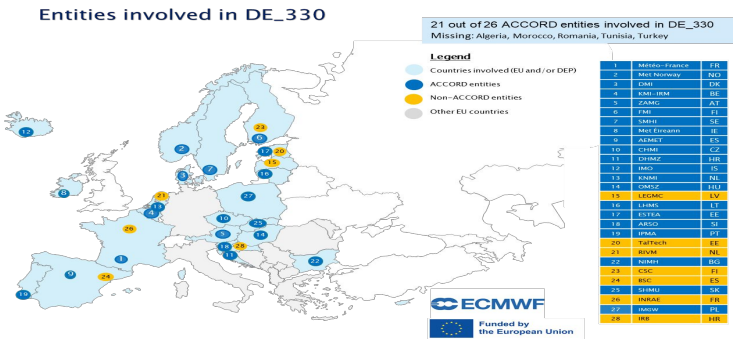


The hurricane-scale wind on 9 March 2021 which escaped operational predictions

Destination Earth On-Demand Extreme DT in short

- Event-driven or user-driven, on-demand weather-induced extreme DT Engine for selected impact sectors
- Brings together 28 European Institutes including 21 National Weather Services
- Development of operational-capable on-demand DTE workflow on EuroHPC platforms
- Showcasing value-added hectometric scale modelling and exploitation of high density observations

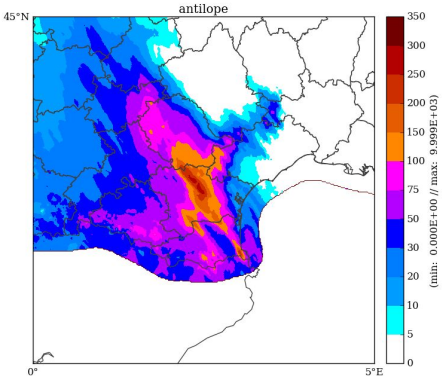




Mission: build a workflow for a hectometric resolution NWP forecast anywhere in Europe within the hour, and couple it with th relevant impact model, on Exascale EuroHPC

Detect a possible extreme event from the continuous extreme DT

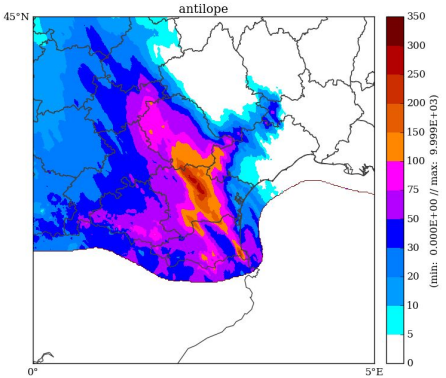
Mission: build a workflow for a hectometric resolution NWP forecast anywhere in Europe within the hour, and couple it with th relevant impact model, on Exascale EuroHPC



The Aude flooding 2018

Activate the appropriate setup over the domain of interest on < 1km grid resolution

Mission: build a workflow for a hectometric resolution NWP forecast anywhere in Europe within the hour, and couple it with th relevant impact model, on Exascale EuroHPC



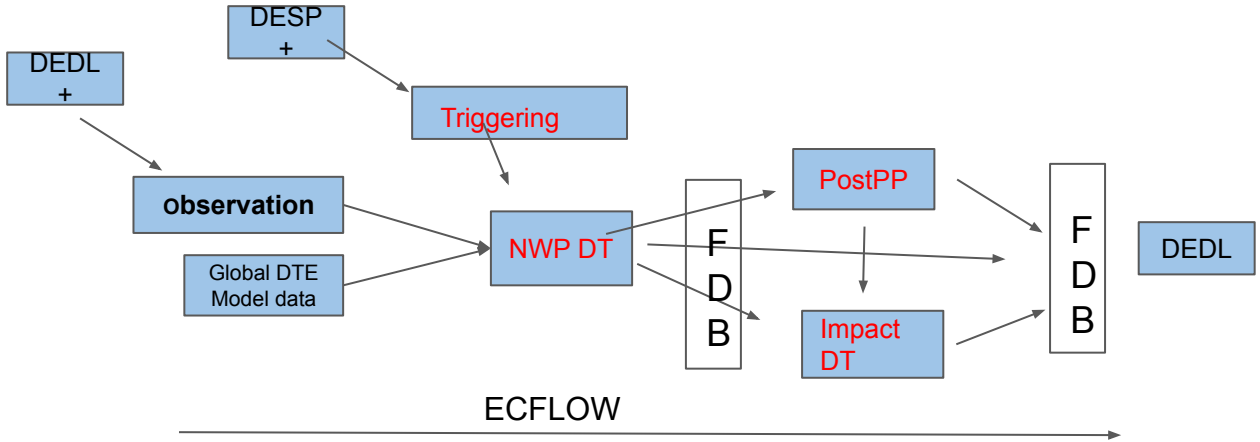
The Aude flooding 2018

Run the relevant impact model

... ... for decision-making support

Technically the On-demand DT is targeted to:

- Build a triggering module to activate on-demand DT
- Construct a framework that interoperate different ACCORD flavours under common code and system base
- Interface it with the (ECMWF) continuous DT workflow for real time production at the EuroHPC-LUMI/Leonardo
- Establish an integrated workflow, using **ecflow**, connecting internally NWP and externally with the selected impact models (air quality, hydrology, renewables)
- Connect input/output to data providers and users via DEDL data-bridge



Major challenges with the On-demand DT workflow

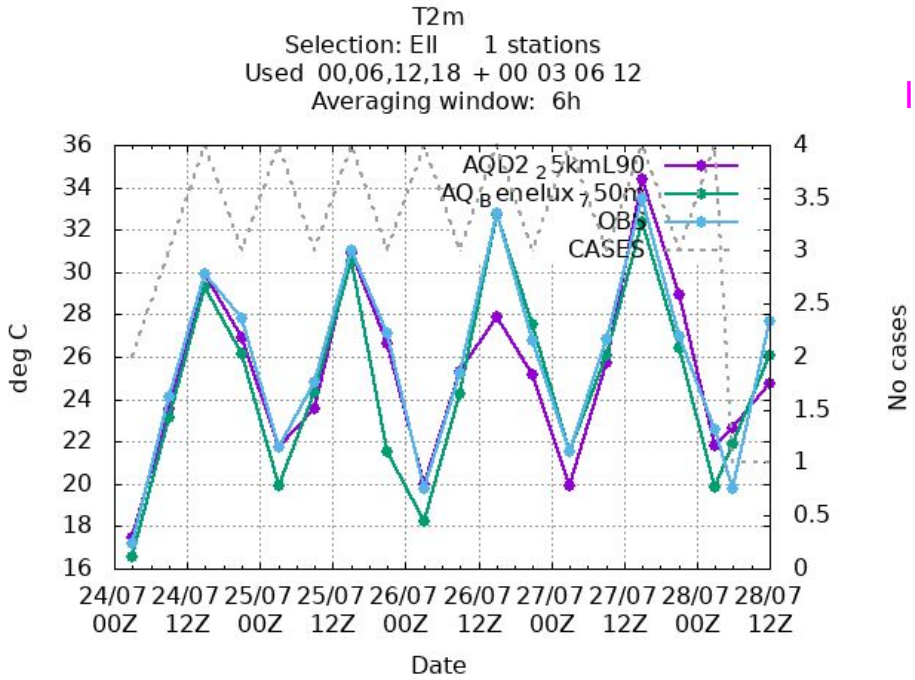
- Adapted to hybrid CPU-GPU platforms (EuriHPC) (Daan Degrauwe talk in MS4A)
- Robust configuration for hectometric scale NWP setup with added values
- Configurability is a special challenge to the on-demand DT workflow
  - Different configurations and domains every day with no history from yesterday
  - More components in the time critical path and higher demands on scalability
    - Ultimately runs will be done on the untested on-demand domains
    - Generation of static data (PGD) has to be efficient
    - Need to Interpolate effectively and reliably surface/soil states to the target grid (PREP)
- Traditional assimilation, ensemble forecast and post-processing methods less applicable
  - No first guess to cycle from
  - Few model data for training
  - The slow adjustment processes near the surface do not have time to act
- Runs on non-operational computation platforms
  - Timeliness and frequent update may be more demanding for extreme related applications
  - Capability to operate on diverse EuroHPC hosts

Special challenge with on-demand DT in cold-start

- For some weather types, cold start is a major challenge to quality of simulation
  - surface/soil quantities go through slow spin-up to reach equilibrium

Illustration: slow convergence with temperature and pressure in cold-start

Special challenge with on-demand DT: cold-start

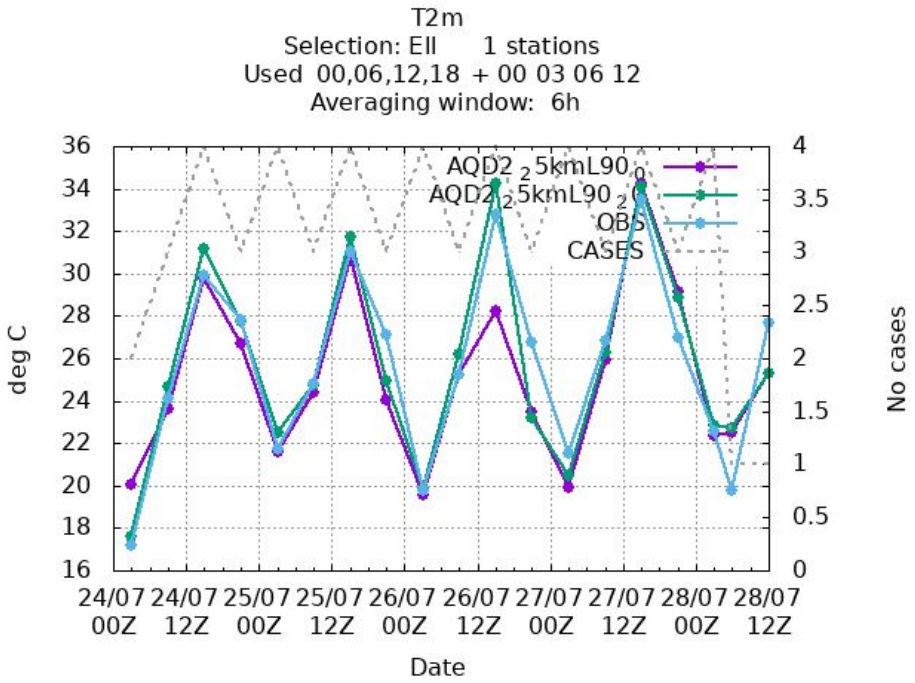


In this example, 2.5km simulation looks inferior

2.5 km
Vs
750 m
Vs
Observation

Simulated 2m temperature time series during the 2018 European heatwaves for station Ell, Netherlands

Special challenge with on-demand DT: cold-start

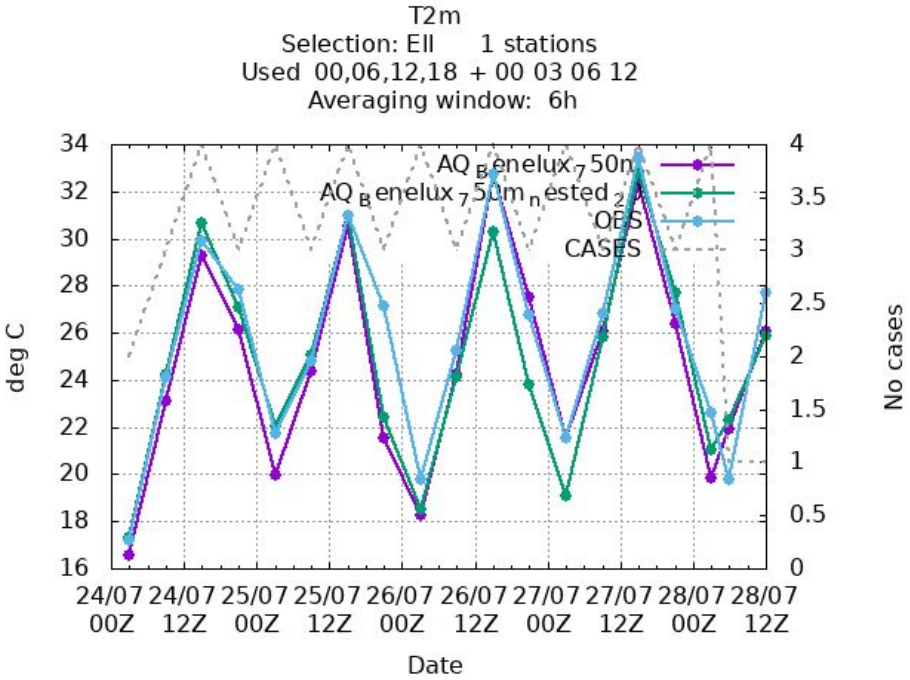


With a 20-day warm up, the simulation clearly improves

Cold-start
Vs
Warm-start (20 day cycled runs)
Vs
Observation

Simulated 2m temperature time series during the 2018 European heatwaves for station Ell, Netherlands

Special challenge with on-demand DT: cold-start



Coupling alternatives make difference...

750m, Downscaling
Vs
750m, Double nesting from continuous
Vs
Observation

Simulated 2m temperature time series during the 2018 European heatwaves for station Ell, Netherlands

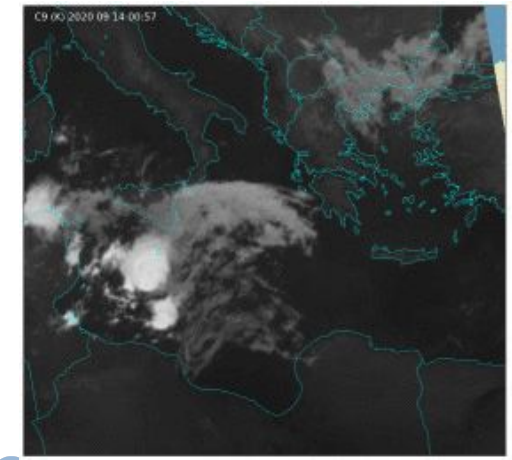
How will cold-start work in the on-demand DTE?

- For some weather types, cold start is a major challenge to quality of simulation
  - surface/soil quantities go through slow spin-up to reach equilibrium
  - coupling to continuous running models with matching surface schemes?
    - Double nesting to a continuous running model?
    - Coupling to continuous running, coarser resolution offline models
      - With own coarser resolution offline cycling using same surface scheme (SURFEX)
      - With Offline many-layer soil scheme with IFS

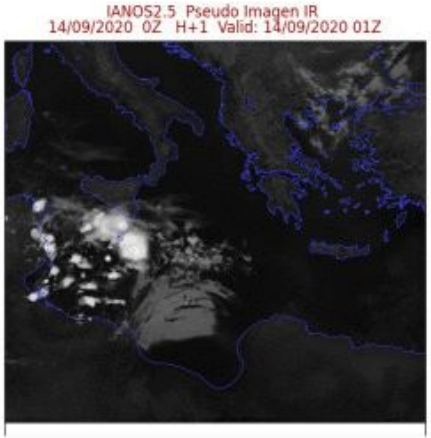
Triggering module

- AII/statistics using continuous DT and observations; HIgh impact? System feasibility? Impact sectors? End-users?
- Configure -> Activate
  - One of application scenario is in setting up domains along storm track

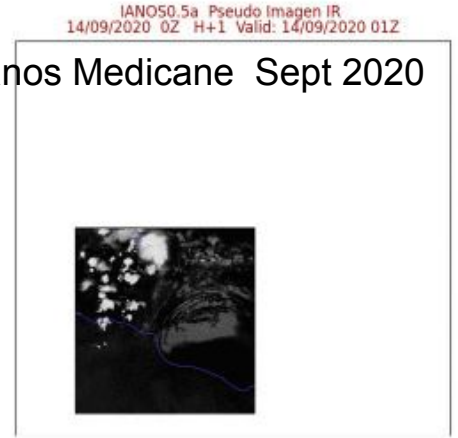
MSG SEVIRI CH. 9 (IR)



ACCORD NWP **2.5 km**



ACCORD NWP **500 m**



(Courtesy Javier Calvo, AEMET)

Tailoring in configuration

- User-needs define different configuration characteristics
  - Situation-dependent horizontal/vertical resolution
  - Air quality has different demands compared to hydrology/renewable energy
  - A stormy case over Iceland has different demands compared to heavy rain in Spain
- The output needs to be tailored to match applications
  - Hydrology requires 2 fields, but wishes sub-hour output and a long calibration period
  - Air quality requires ~30 fields (3D/2D)
  - Renewables requires ~15 minute or more frequent output
- There needs to be integrated downstream workflow to ensure usefulness
  - On-Demand DT needs clear user to provide decision making support

A more data centric approach

We are starting with

- GRIB2 output to files from all components
- Store fields in FDB (Field DataBase)
- Major together with ECMWF on defining GRIB2 template (SURFEX...)
- Preparing for sub hourly output and ccsds packing

- The garden vegetation groupings will be **(TREE + BARE + GRAS)**. TREE will be in {TEBDU, TRDBU, TEBEU, TRBEU, BONEU, TENEU, BONDU}, where {TREE}U is the urban counterpart for a given tree type.
  - Implementation:

```
1034 TEBDU Urban temperate broadleaf deciduous
1035 TRBDU Urban tropical broadleaf deciduous
1036 TEBEU Urban temperate broadleaf evergreen
1037 TRBEU Urban tropical broadleaf evergreen
1038 BONEU Urban boreal needleleaf evergreen
1039 TENEU Urban temperate needleleaf evergreen
1040 BONDU Urban boreal needleleaf deciduous
```

```
1525 G025 Group 025 (NONE + GRAS + TEBDU)
1525 G026 Group 026 (NONE + GRAS + TRBDU)
1525 G027 Group 027 (NONE + GRAS + TEBEU)
1525 G028 Group 028 (NONE + GRAS + TRBEU)
1525 G029 Group 029 (NONE + GRAS + BONEU)
1525 G030 Group 030 (NONE + GRAS + TENEU)
1525 G031 Group 031 (NONE + GRAS + BONDU)
```

  - Is BARE = NONE or BARE = (NONE+ROCK+SNOW)? Currently I have assumed it is only NONE.
- Described new implementations of list of tile attributes, support roof/road/wall temperatures, accumulated parameters, and "Aggregated" covers.

Example of garden vegetation grouping

Courtesy: Matthew Griffith, Sebasiten Villaume ECMWF, Patrick Le Moigne, Søren Borg Nielsen, Mikko Aalto, Patrick Samuelsson, Trygve Aspelien within DE_330

A more data centric approach

We are starting with

- GRIB2 output to files from all components
- Store fields in
- Major together GRIB2 templa
- Preparing for s packing

Exploring

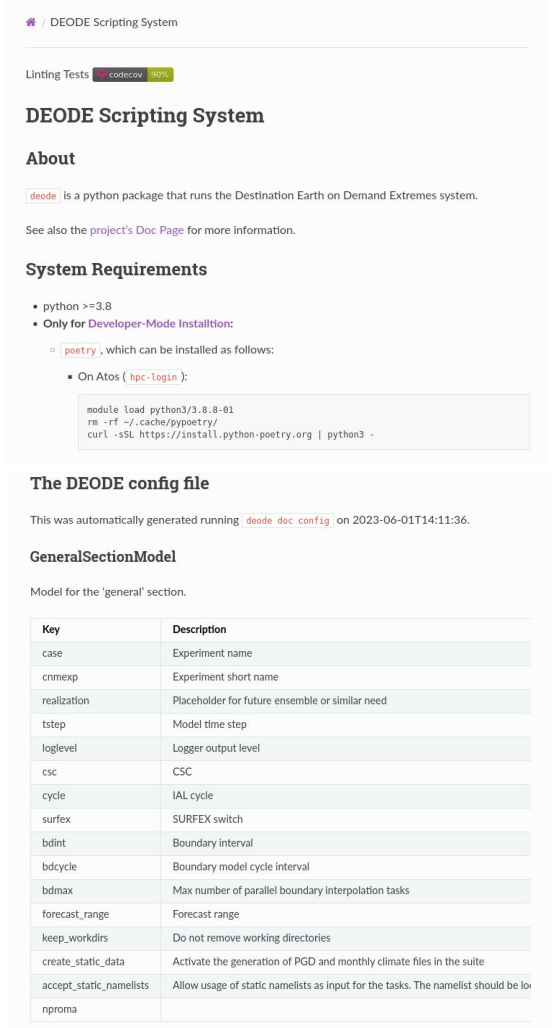
- Writing directly to FDB from th NWP model, and other applications
- Read directly from FDB to downstream applications (https://github.com/ecmwf/earthkit-data)
- Removes a few unnecessary/costly IO steps

Aiming at

- Interface with the new ECMWF multiplex IO server accessing/pushing data runtime
- Do the work while data is in memory!
- For the longer term view

Designing a new scripting environment

- Python based
- Large focus on following standards, unit testing, code coverage from the beginning. Apply github pipelines and CI/CD processes
- Modularity, e.g. separate ecflow from the tasks; interoperability between different ACCORD flavours
- All tasks should be possible to run stand alone for easier development and debugging
- Config file driven (yaml, toml, json)
- Hope to be useful to ACCORD in its harmonisation efforts!



/ DEODE Scripting System

Linting Tests codecov 90%

DEODE Scripting System

About

`deode` is a python package that runs the Destination Earth on Demand Extremes system.

See also the project's Doc Page for more information.

System Requirements

- python >=3.8
- **Only for Developer-Mode Installtion:**
  - `poetry`, which can be installed as follows:
    - On Atos (`hpc-login`):

```
module load python3/3.8.8-01
rm -rf ~/.cache/pypoetry/
curl -sSL https://install.python-poetry.org | python3 -
```

The DEODE config file

This was automatically generated running `deode doc config` on 2023-06-01T14:11:36.

GeneralSectionModel

Model for the 'general' section.

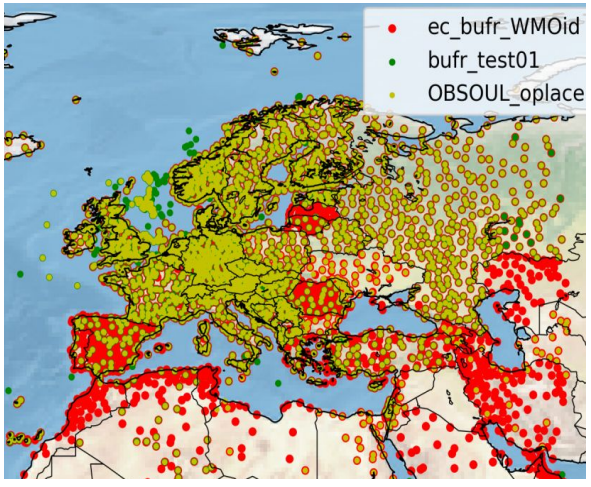
Key	Description
case	Experiment name
cnmexp	Experiment short name
realization	Placeholder for future ensemble or similar need
tstep	Model time step
loglevel	Logger output level
csc	CSC
cycle	IAL cycle
surfex	SURFEX switch
bdint	Boundary interval
bdcycle	Boundary model cycle interval
bdmax	Max number of parallel boundary interpolation tasks
forecast_range	Forecast range
keep_workdirs	Do not remove working directories
create_static_data	Activate the generation of PGD and monthly climate files in the suite
accept_static_namelists	Allow usage of static namelists as input for the tasks. The namelist should be lo
nproma	

ECMWF

https://github.com/DEODE-NWP

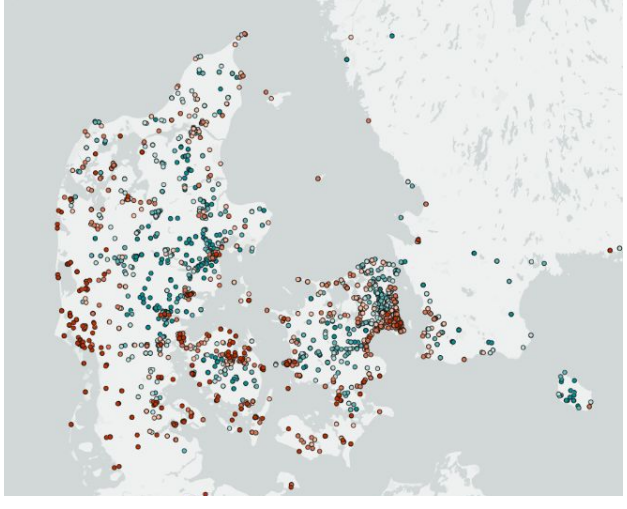
Observations

- Utilisation of high density obs data such as crowd-sourced data is one of the value-adding features in the On-demand , hectometric scale DTE
- The first phase does not focus on implementing full scale data assimilation but setting up real time acquisition takes time
- Possible combination with DEDL and contribution from involved NMSs
  - Granted data from LACE, UWC-W/MetCoOp
  - Investigation of potential input from others
- Collection of high density observation streams such as crowd sourced data
- Now focusing on gathering data for verification and postprocessing



Conventional Obs

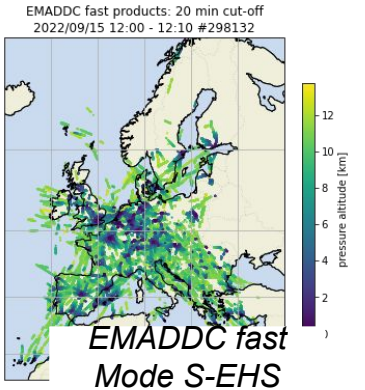
(B Strajnar)



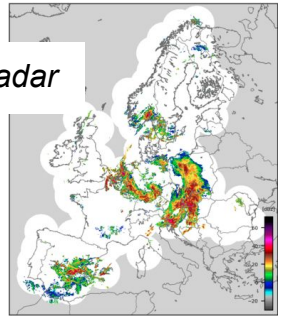
Smartphone Pressure obs

(K Hintz)

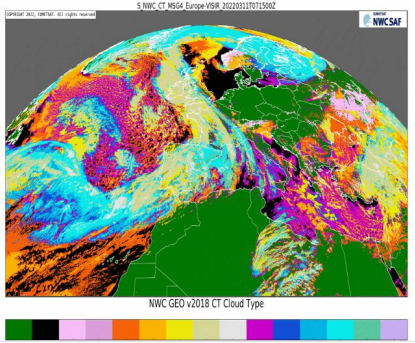
High Density Observations



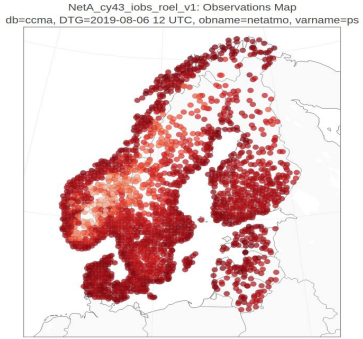
EMADDC fast Mode S-EHS



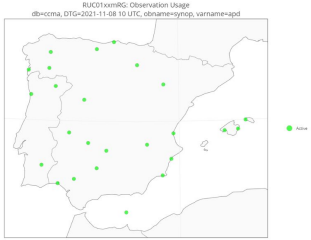
Radar



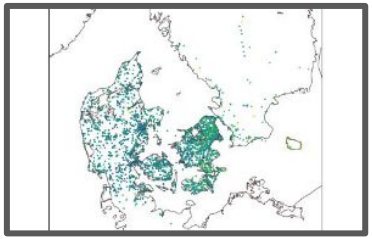
NWC-SAF cloud product



NETATMO/WOW



Fast processing GNSS



Smartphone pressure

(Courtesy
Magnus Lindskog 2022)

Summary

- We're building a flexible, on-demand system targeting on extreme events to be "demonstration ready" on a non-operational EuroHPC platform
- Numerous scientific and technical challenges are yet to be addressed but these are also opportunities for development. Solutions to these challenges highly synergic and beneficial for the NWP activities at the European weather services
    - Hectometric scale modelling; recoding to platform with accelerators; Integrated workflow with impact modellers; triggering modules with increasing sophistication; transparent, harmonised and integrated production workflow; organisation of production on diverse EuroHPC and cloud platforms; data centric workflow
- Not dealt with so far during current phase: upper air data assimilation and probabilistic forecast
- For ACCORD specifically, the project is moving us closer, benefitting partner collaboration and harmonisation

Thank you for your attentions!

Acknowledged co-workers

Paulo Medeiros Kasper Hintz Søren B Nielsen Emy Alerskans Stefan Rethmeier Fabrizio Baordo Tommaso Bennachio Leif Denby Mikko Partio Mikko Aalto Elmeri Nurmi Erik Gregow Christoph Wittmann Florian Weidle Phillip Scheffknecht Adam El-Said Siebren de Haan Trygve Aspelien Eivind Støylen Roel Stappers Ole Vignes Samuel Viana Daniel Martin Javier Calvo Juan Jesus Gonzalez Maria Monteiro Bolli Pálmason Guðrún Nína Petersen Sigurður Þorsteinsson Xiaohui Zhao Maria Derkova Oldrich Spaniel Radmila Brozkova Antonin Bucanek Alena Trojakova Martina Tudor Martynas Kazlauskas Rimvydas Jasinskas Kristina Kryžanauskienė Boryana Tsenova Konstantin Mladenov Milen Tsankov Alex Deckmyn Denise Haumont Jure Cedilnik Neva Pristov Benedikt Strajnar et al