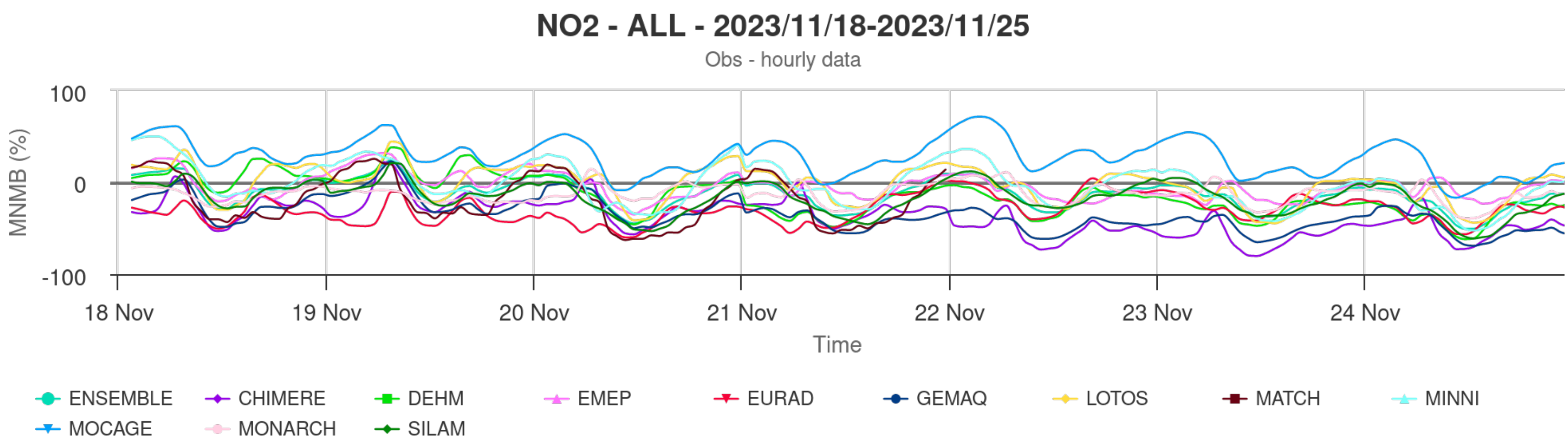
WP4: Added value of the SEEDS assimilation scheme

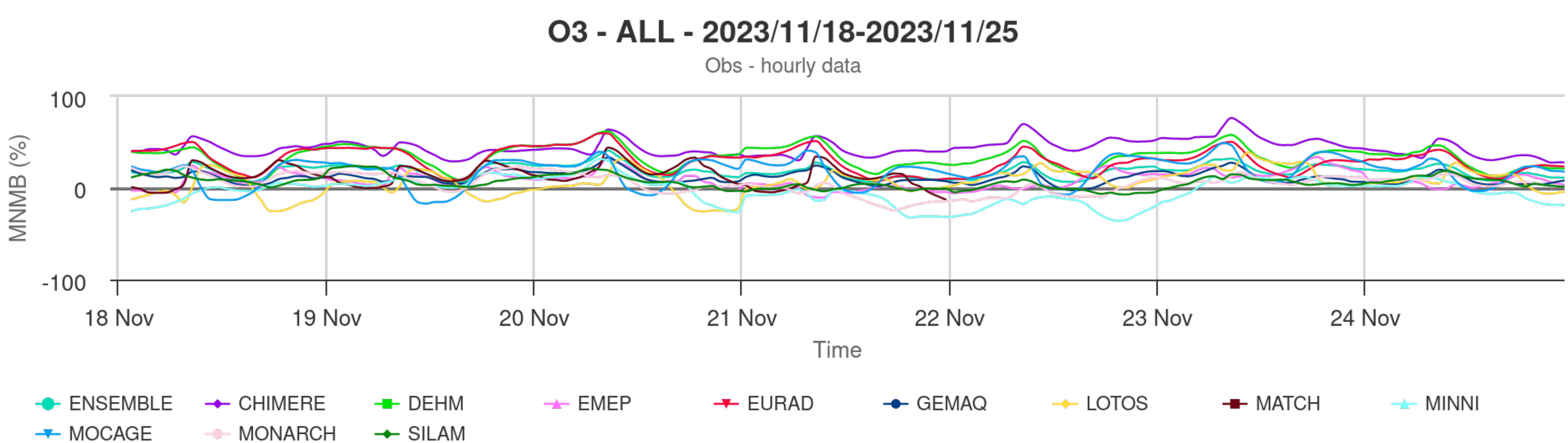
Emanuele Emili*

CERFACS WP4 team: A. Piacentini (Developer), G. Jonville (Developer)
* *Now at the Barcelona Supercomputing Center* (emanuele.emili@bsc.es)
Acknowledgments: B. Ménétrier (Met Norway)

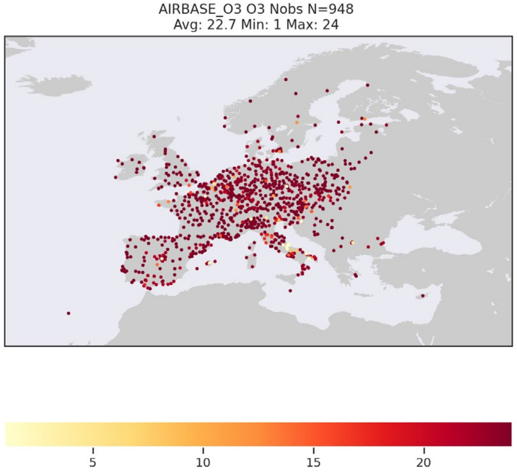
Funded by the European Union

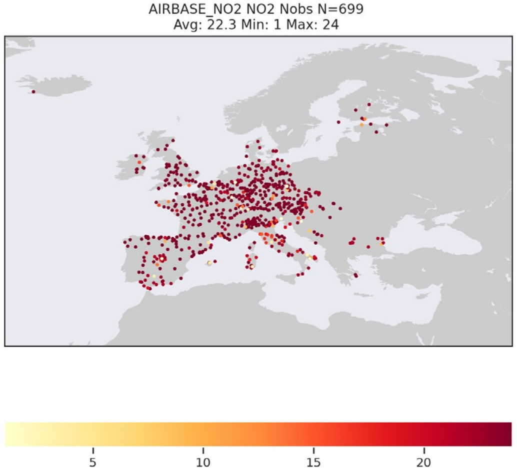
Motivation: air quality forecasts biases

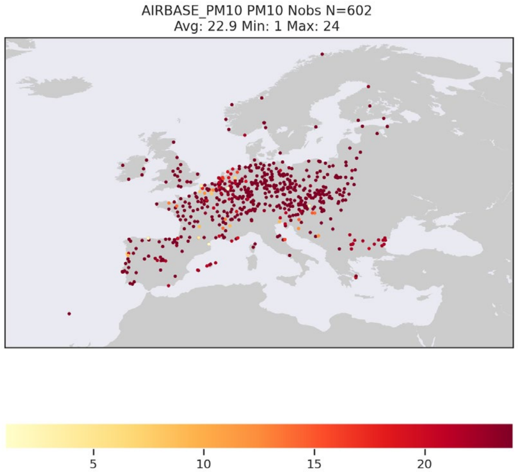


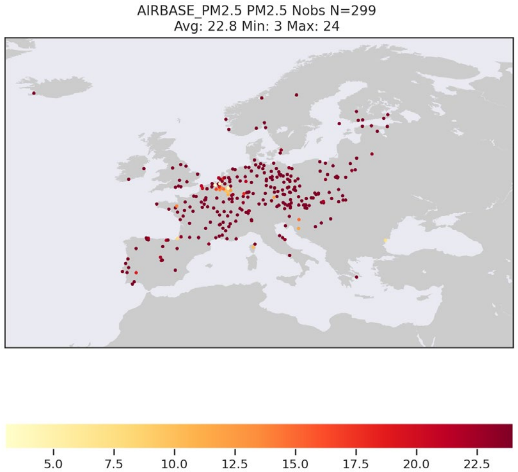


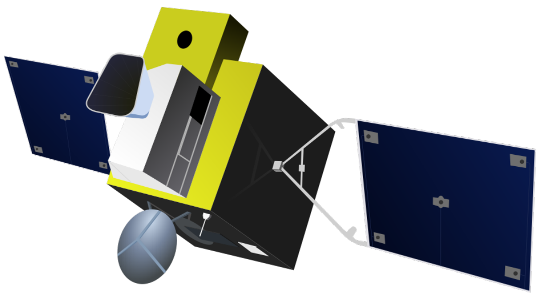
Motivation: Near Real Time observations











Sentinel-4: 8x8 km^2 hourly columns of NO_2, O_3, SO_2, HCHCO and aerosols

Data Assimilation in CAMS models

MODEL	ASSIMILATION ALGORITHM	ASSIMILATED OBSERVATIONS
CHIMERE	Kriging	O_3, NO_2, PM_{10}, $PM_{2.5}$
EMEP	3D-Var	O_3, NO_2, SO_2, PM_{10} + SAT NO_2
EURAD-IM	3D-Var	O_3, NO_2, PM_{10}, $PM_{2.5}$ + SATELLITE NO_2, SO_2, CO
LOTOS-EUROS	EnKF	O_3, NO_2, PM_{10}, $PM_{2.5}$ + SAT NO_2
MATCH	3D-Var	O_3, NO_2, SO_2 CO, PM_{10}, $PM_{2.5}$
MOCAGE	3D-Var	O_3, NO_2, PM_{10}, $PM_{2.5}$
SILAM	3D-Var	O_3, NO_2, SO_2, CO, PM_{10}, $PM_{2.5}$
DEHM	Optimal Interpolation	O_3, NO_2
GEM-AQ	Optimal Interpolation	O_3, NO_2, SO_2 CO, PM_{10}, $PM_{2.5}$
MONARCH	LETKF	O_3, NO_2, SO_2 CO, PM_{10}, $PM_{2.5}$
MINNI	Optimal Interpolation	O_3, NO_2, SO_2 CO, PM_{10}, $PM_{2.5}$

Motivation

Problem is: the impact of DA on forecasts is not very significant

Solutions:

A posteriori

- Improve models processes
- Tune models parameterizations

Near real time

- **Improve models forcings (e.g. emissions) SEEDS WP1-3**
- Joint state and parameters estimation
- **Estimate and correct directly model biases SEEDS WP4**

SEEDS WP4 objectives

Estimate and correct model biases in NRT … with the help of air quality ensembles[1]:

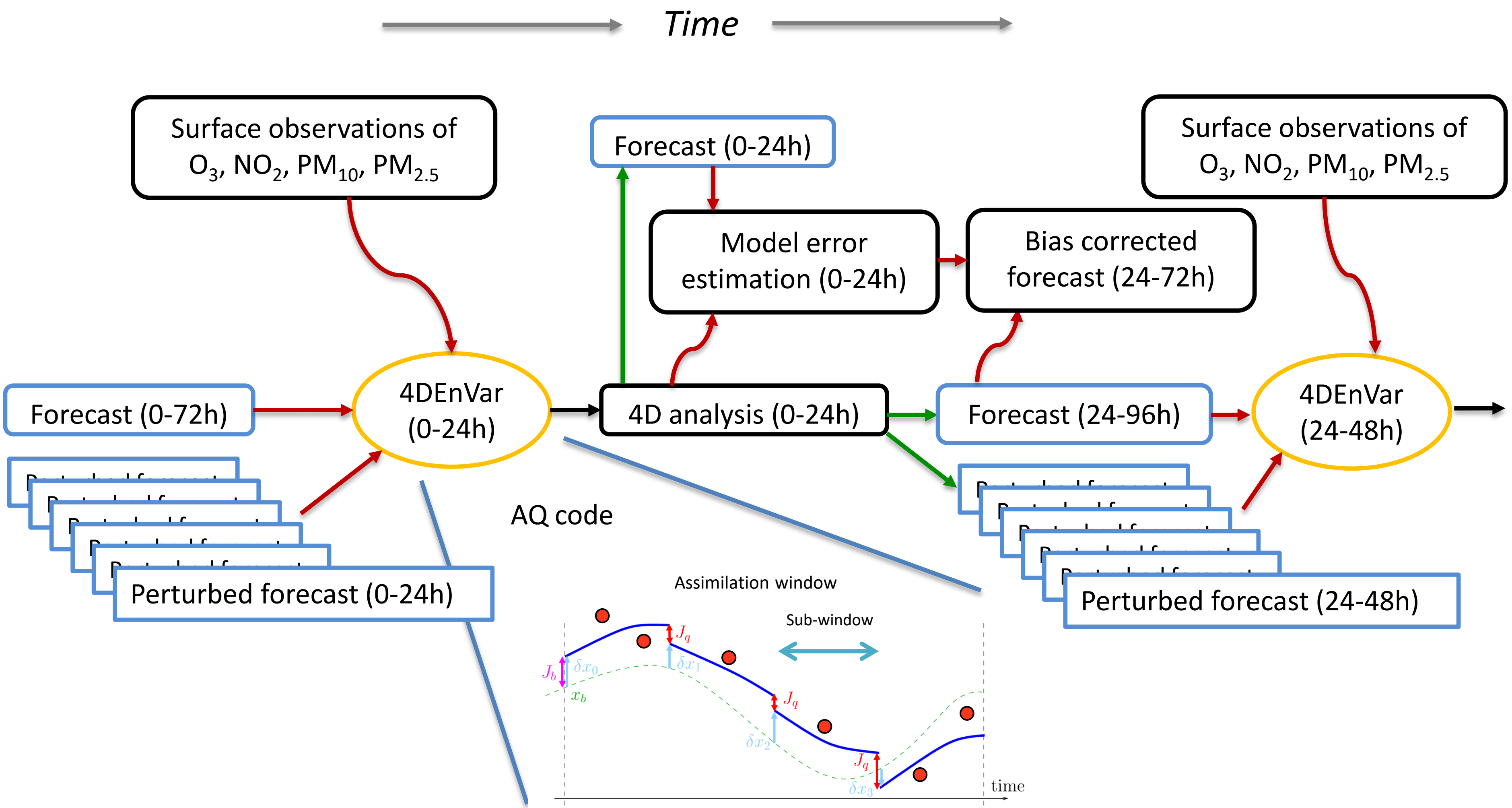
1. Compute ensembles of forecasts based on perturbations of key forcing fields (e.g. emissions, deposition).

2. Assess the added value of an assimilation scheme that includes model error (4DEnVar) with respect to the sequential filters used in CAMS operations (use MOCAGE 3DVar as reference).

3. Provide an open-source assimilation code that can be adapted to other CAMS models equipped with ensemble forecasts, or to the CAMS ensemble itself.

[1]Emili et al. Accounting for model error in air quality forecasts: an application of 4DEnVar to the assimilation of atmospheric composition using QG-Chem 1.0, GMD, 2016

Assimilation and forecast cycles

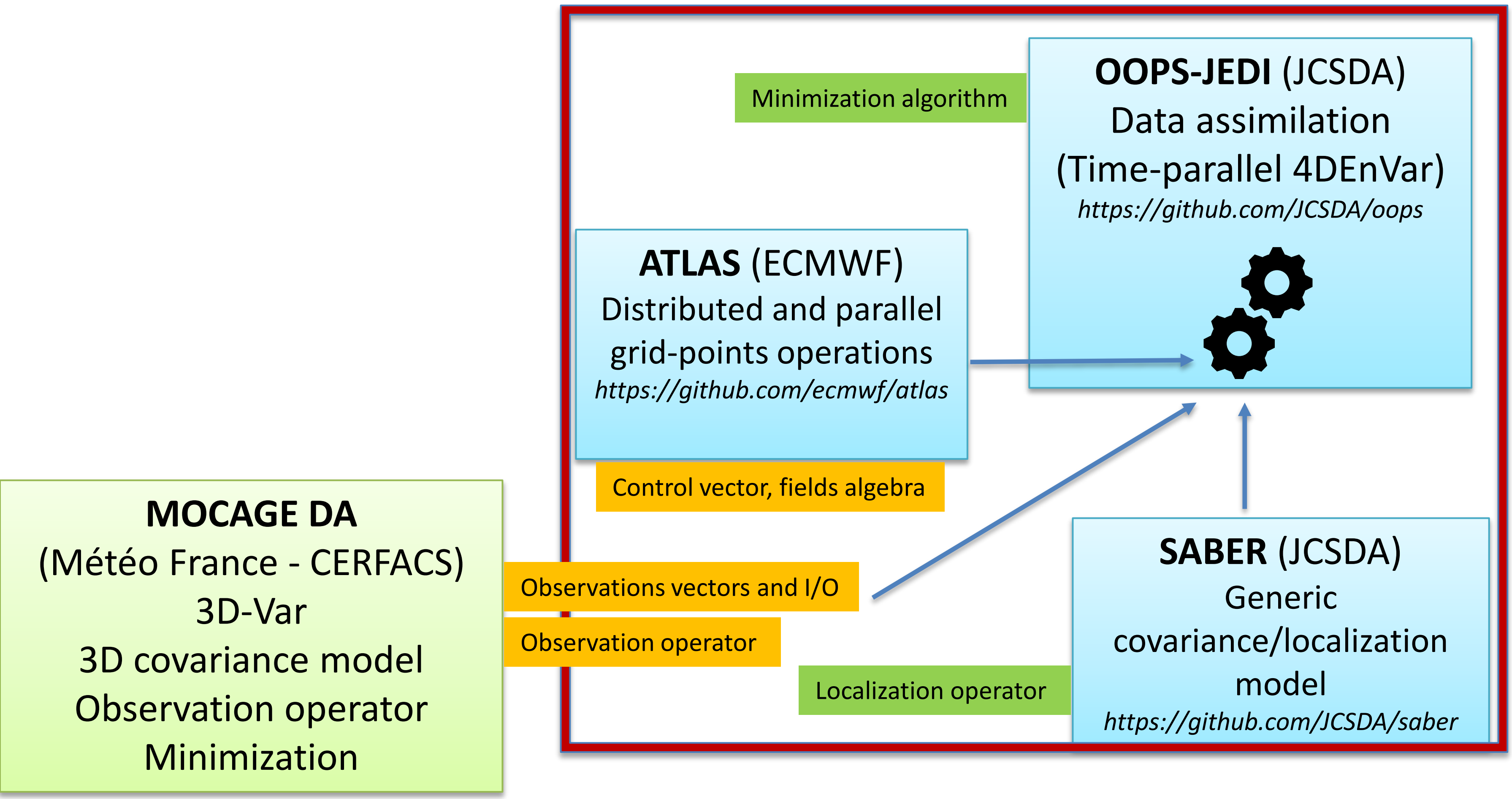


[1]*Emili et al. Accounting for model error in air quality forecasts: an application of 4DEnVar to the assimilation of atmospheric composition using QG-Chem 1.0, GMD, 2016*

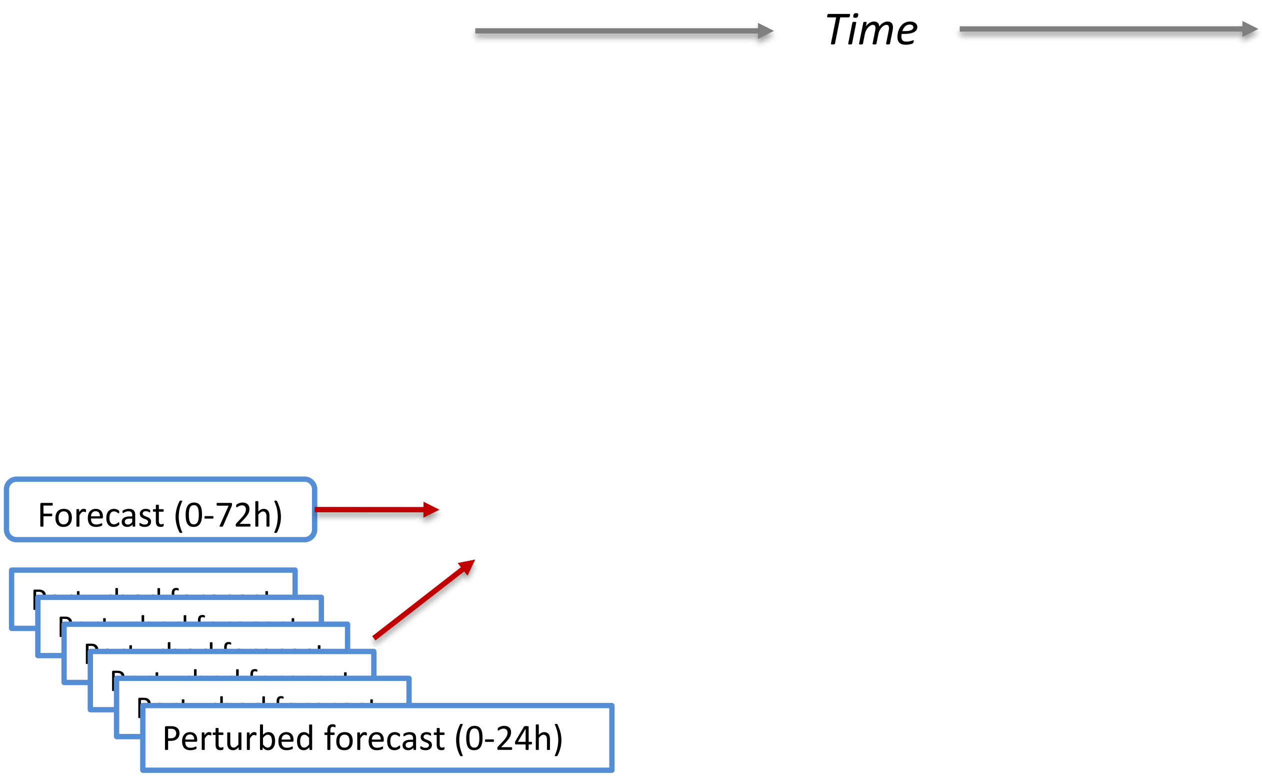
Air Quality (AQ) Data Assimilation

Task 4.2

https://github.com/andreapiacentini/aq/releases/tag/v1.1.0

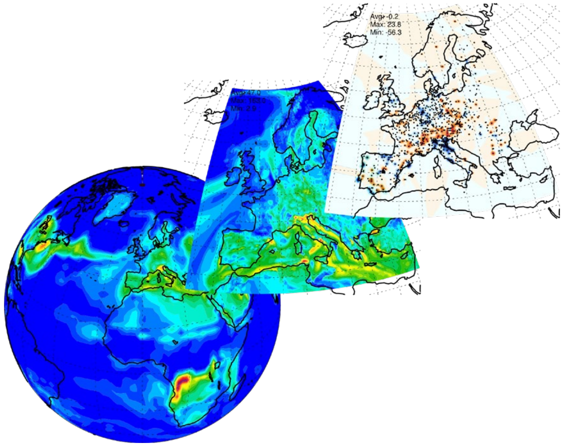


Ensembles of air quality forecasts



Ensembles perturbations

Use MOCAGE chemical transport model:
- RACM tropospheric chemistry
- 3DVar assimilation
- CAMS operational configuration (10 km)



Perturbed quantity	Perturbed Parameters	Species	Probability distribution
Emissions	Total flux	Emitted species	Lognormal (μ=0,σ=0.4) / multiplicative [0.45, 2.19]
Dry deposition	Dry deposition velocities	Deposited species	Lognormal (μ=0,σ=0.5) / multiplicative [0.37, 2.66]
Initial Condition	3D chemical state	All	Lognormal (μ=0,σ=0.2) / multiplicative [0.67, 1.48]
Vertical diffusion	Diffusion coefficient	All	Lognormal (μ=0,σ=0.4) / multiplicative [0.45, 2.19]

Table 1: Configuration used for the ensemble generation. The 95% confidence interval corresponding to the given lognormal distribution is reported in brackets in the 4th column.

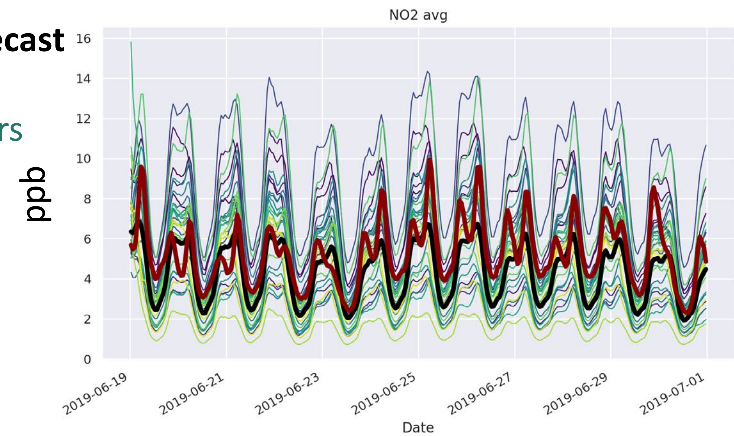
Feedback from satellite emissions estimations WPs in SEEDS:

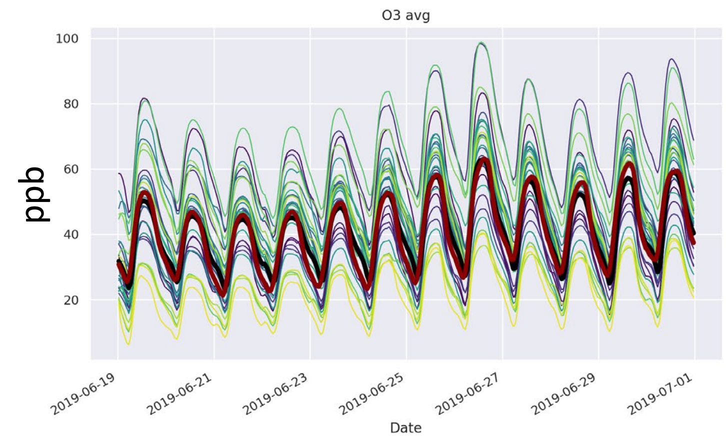
- ✓ NO_x emissions perturbations reduced (-> σ=0.2 [0.67,1.48])
- ✓ Isoprene emissions perturbations increased (-> σ=0.57 [0.32,3.06])

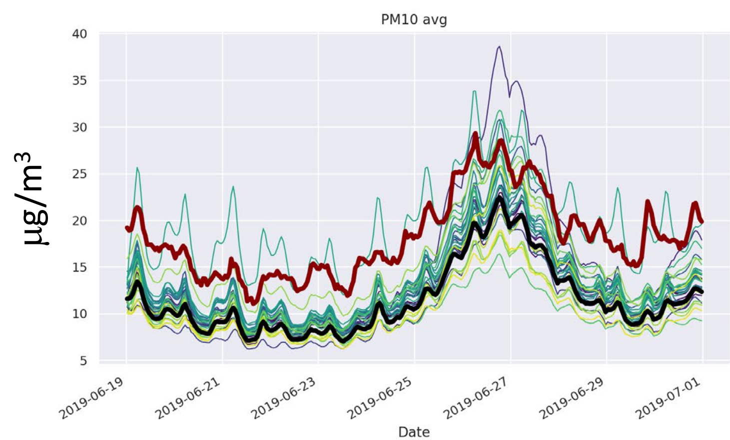
Ensembles verification

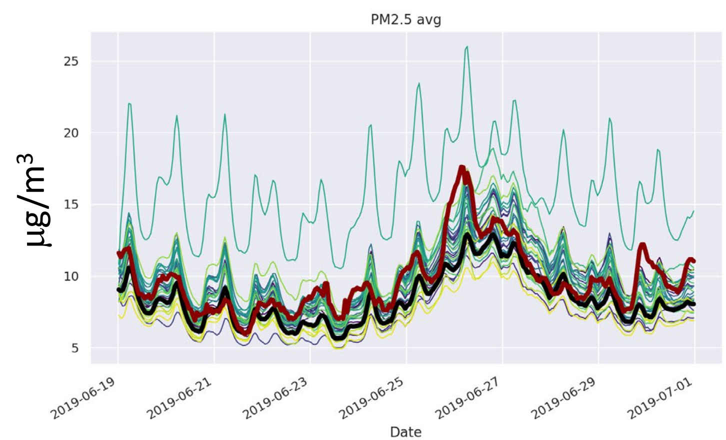
Evaluation against CAMS surface observations (SEEDS D4.1): average over all EU sites

Deterministic forecast
Observations
Ensemble members



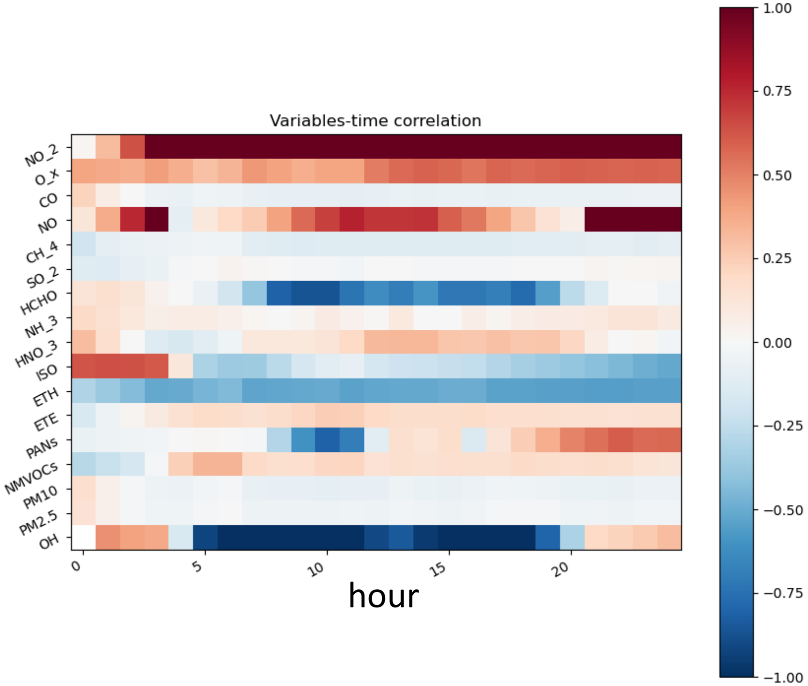






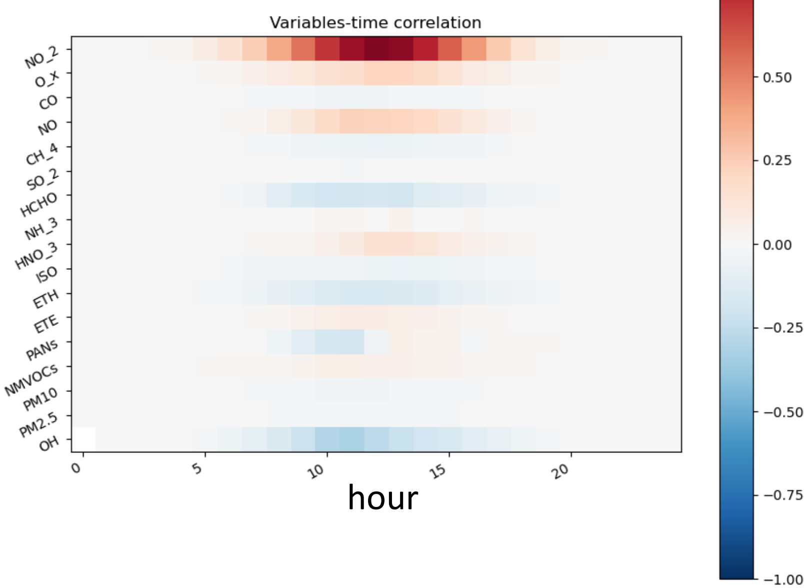
4D Covariance and localization

Raw ensemble correlations

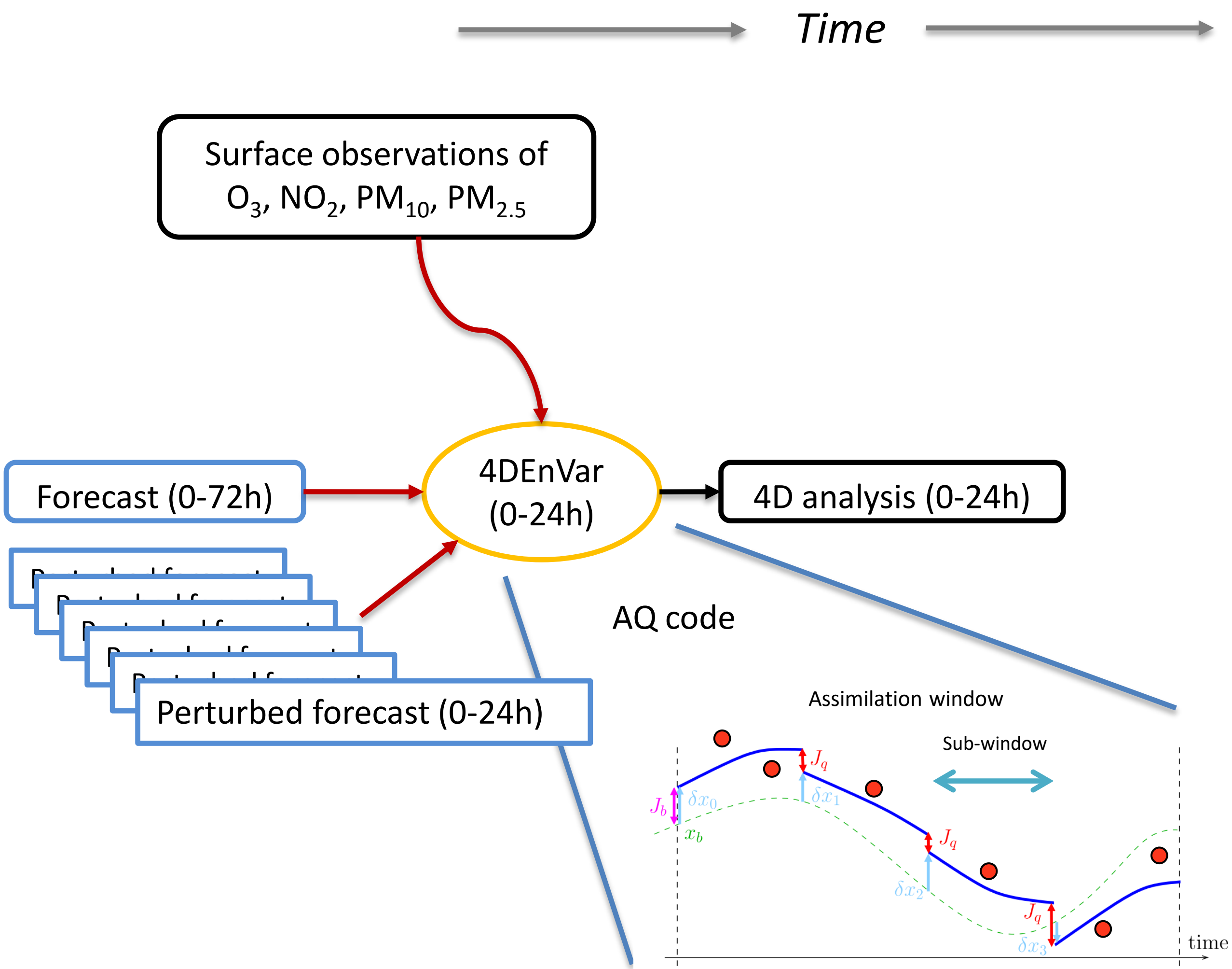


Localized correlations

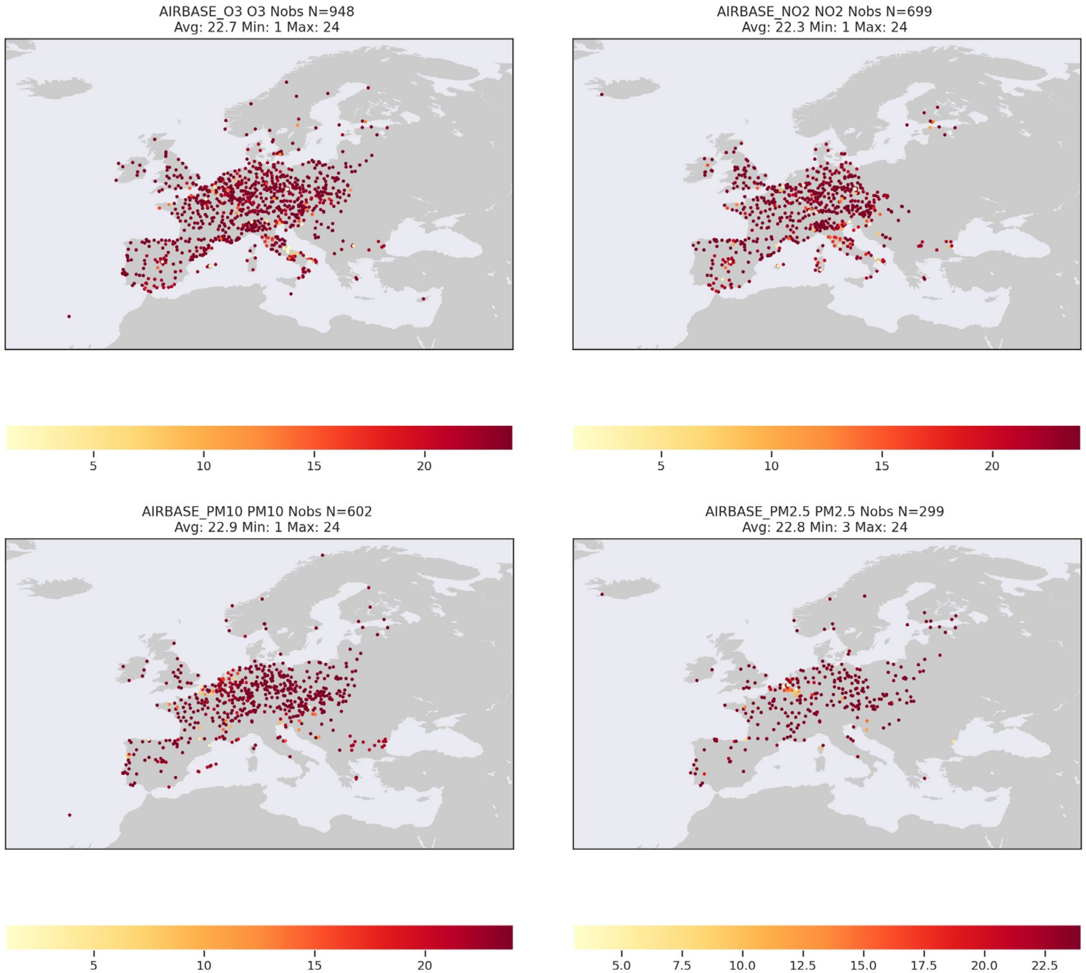
Cross-variable localization: 0.5
Temporal localization length: 3h



4D assimilation

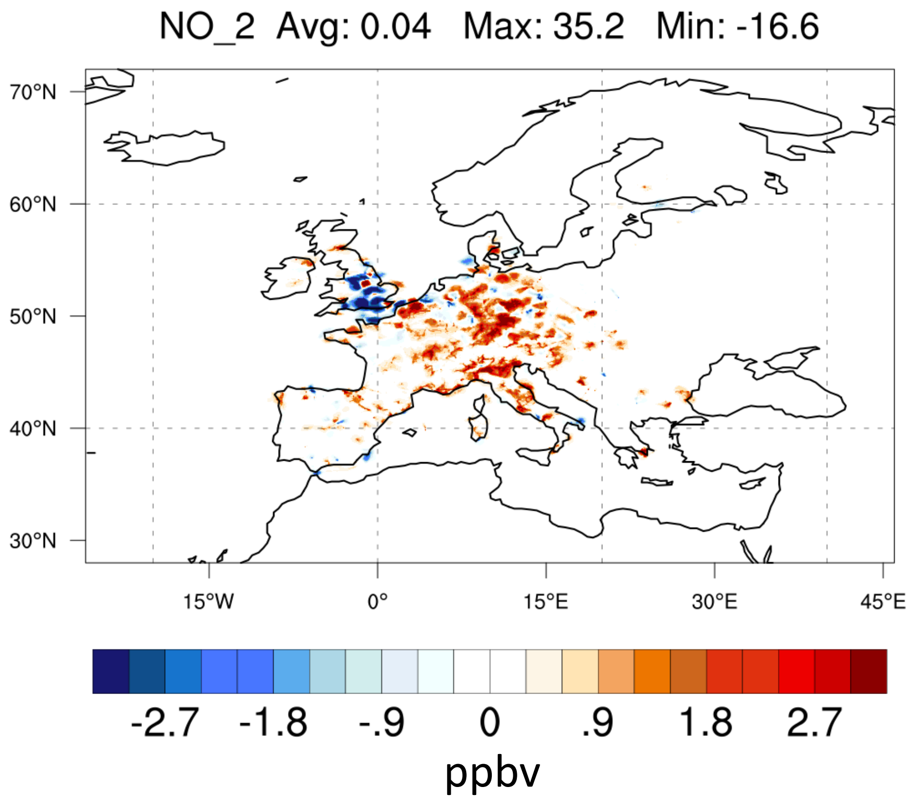


Assimilated network

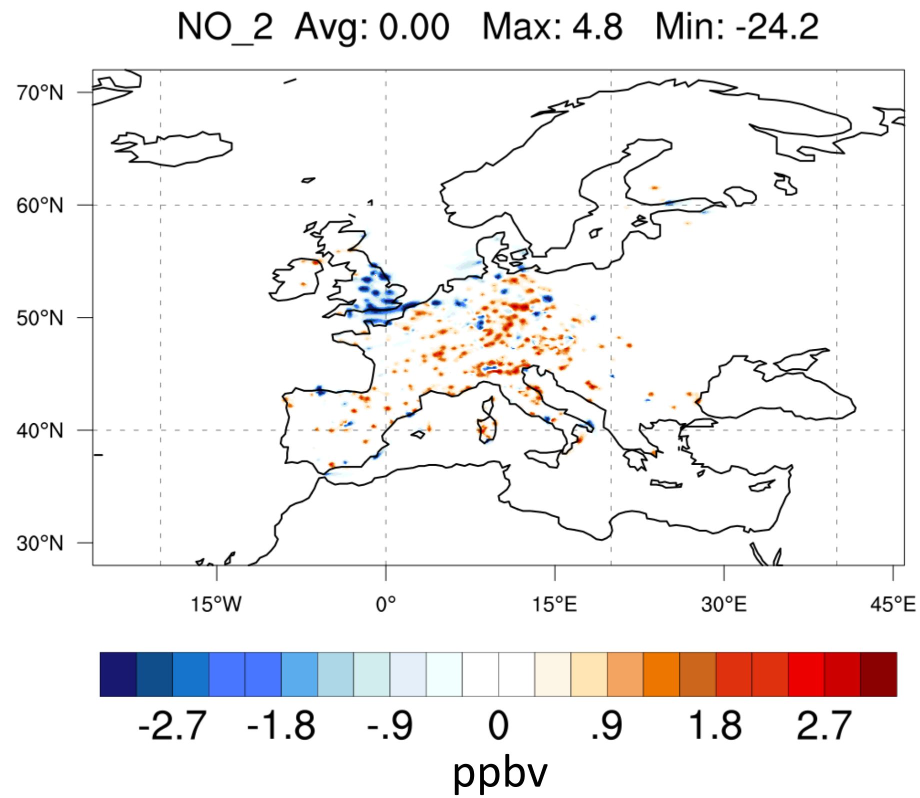


Comparison with 3D-Var

4DEnVar (Analysis-Forecast)

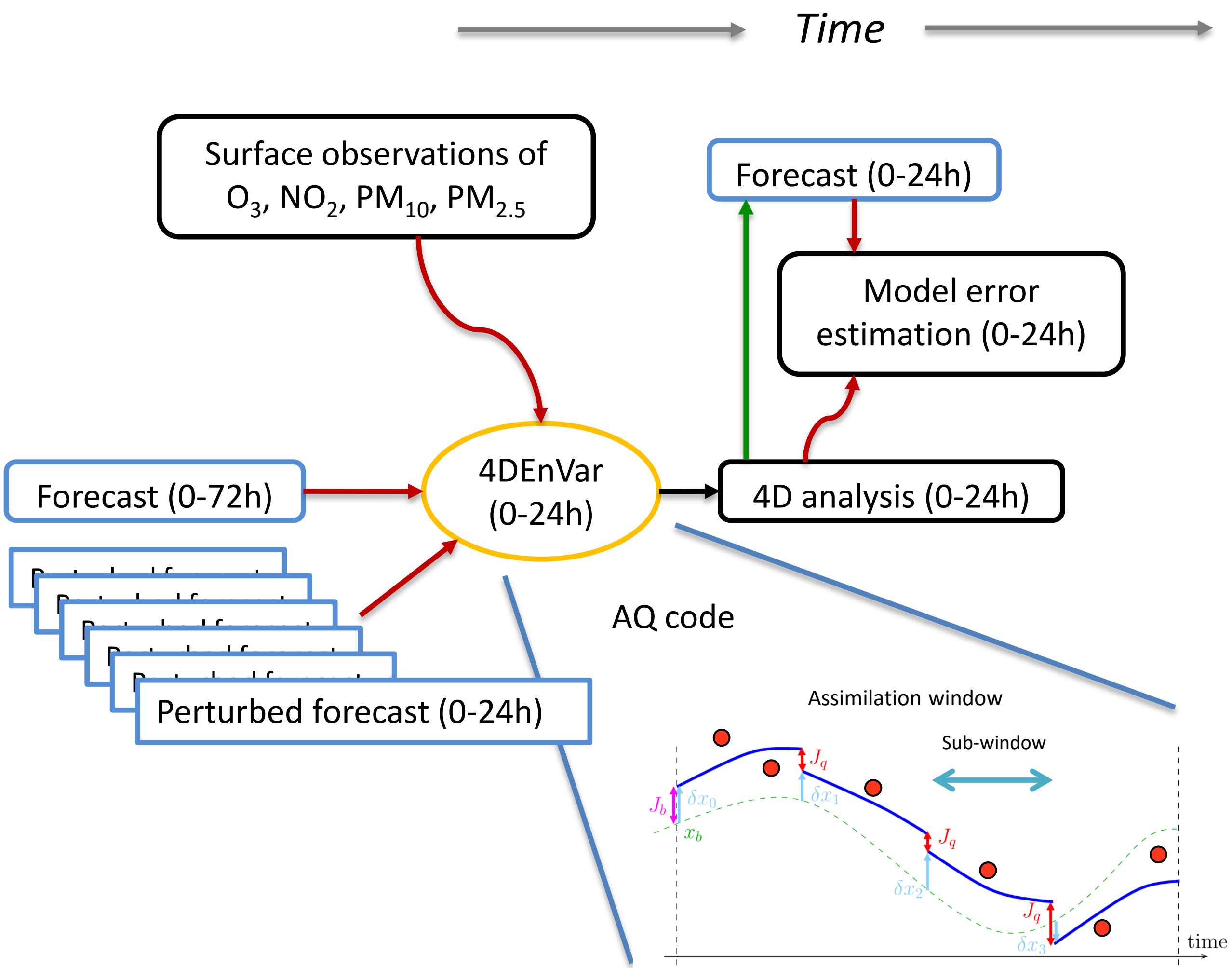


3D-Var (Analysis-Forecast)

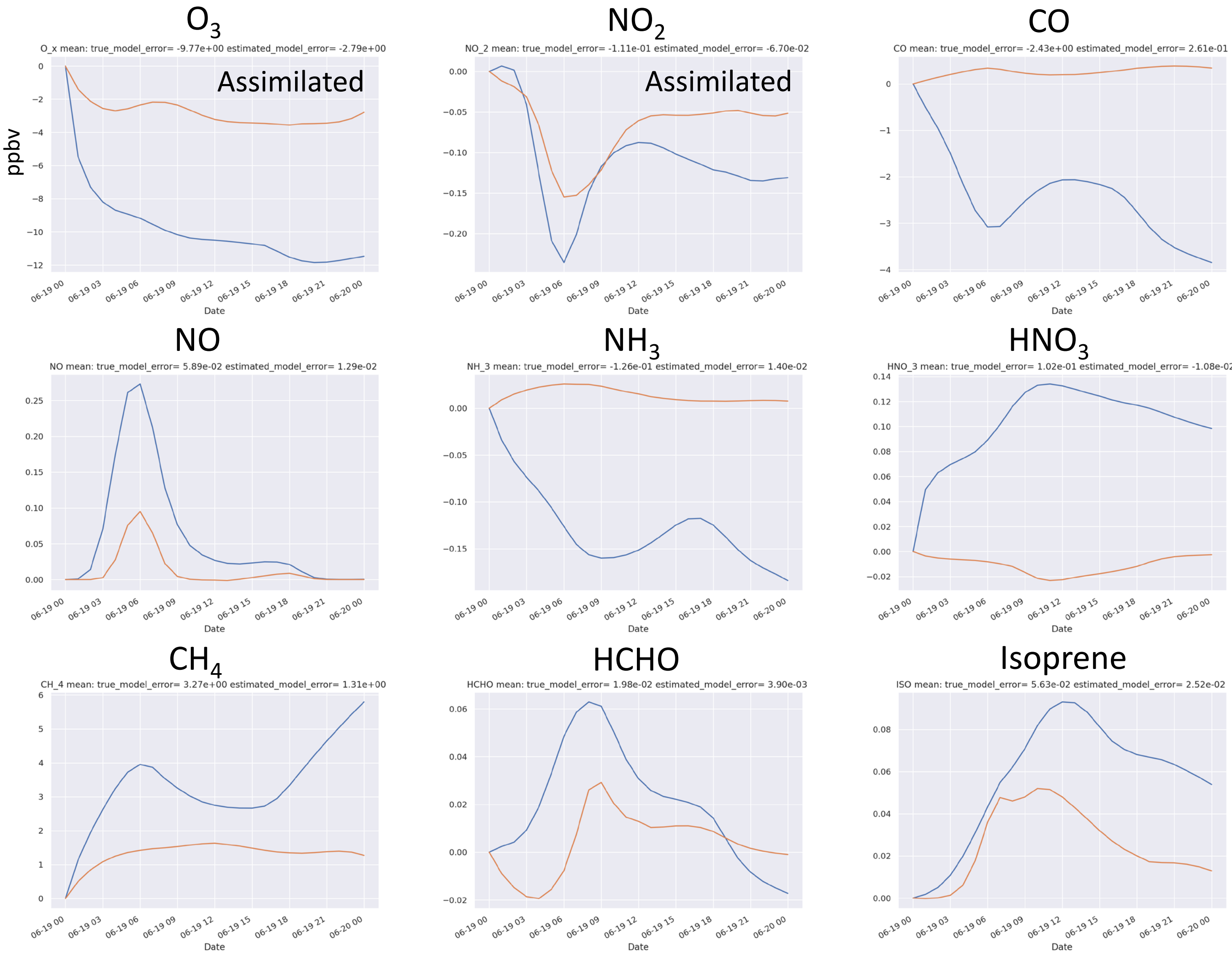


Note: 4DEnVar assimilates logarithms of concentrations, 3DVar absolute concentrations

Model error estimation



Model error estimation: synthetic experiments

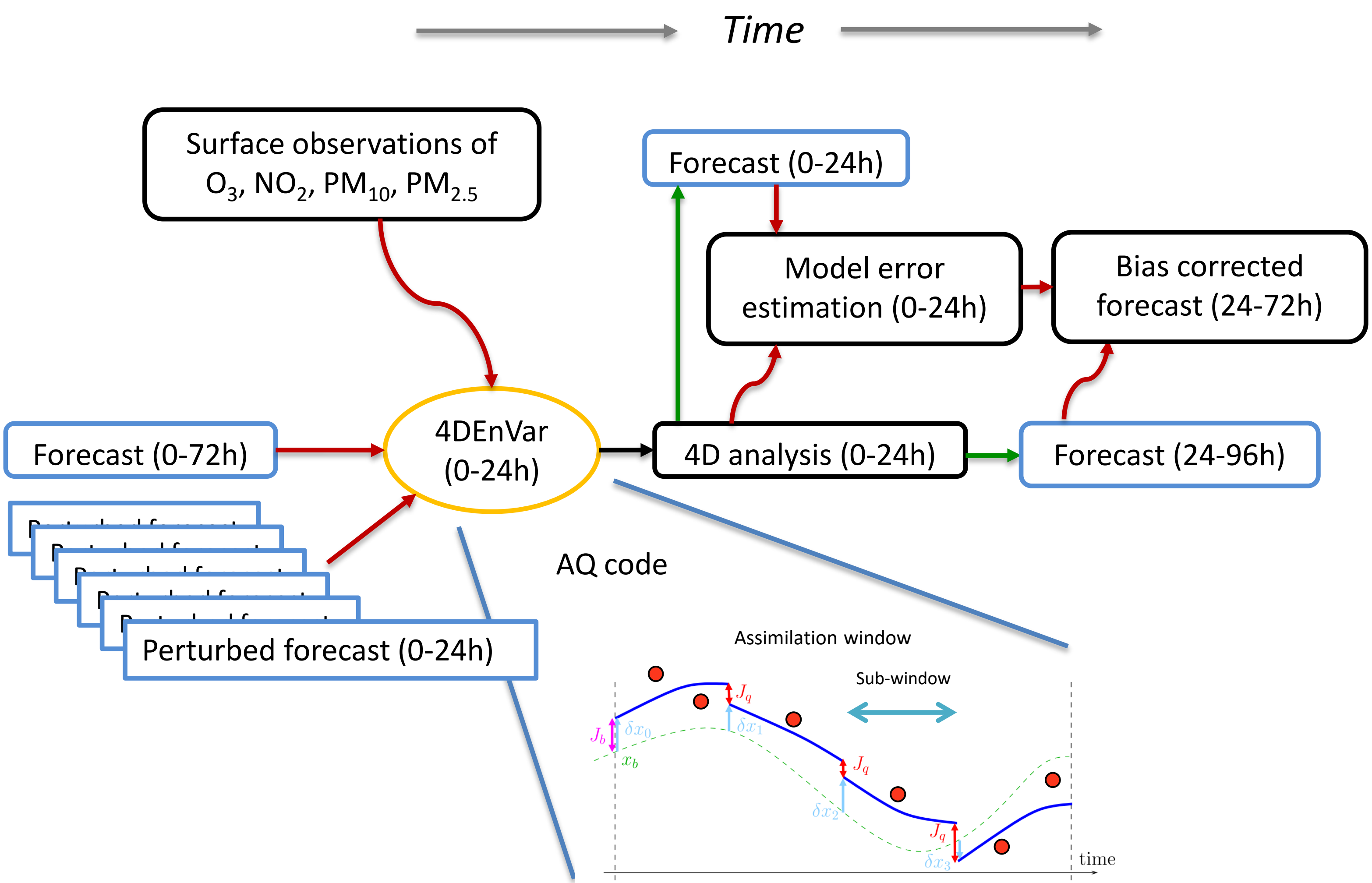


True model error

Estimated model error

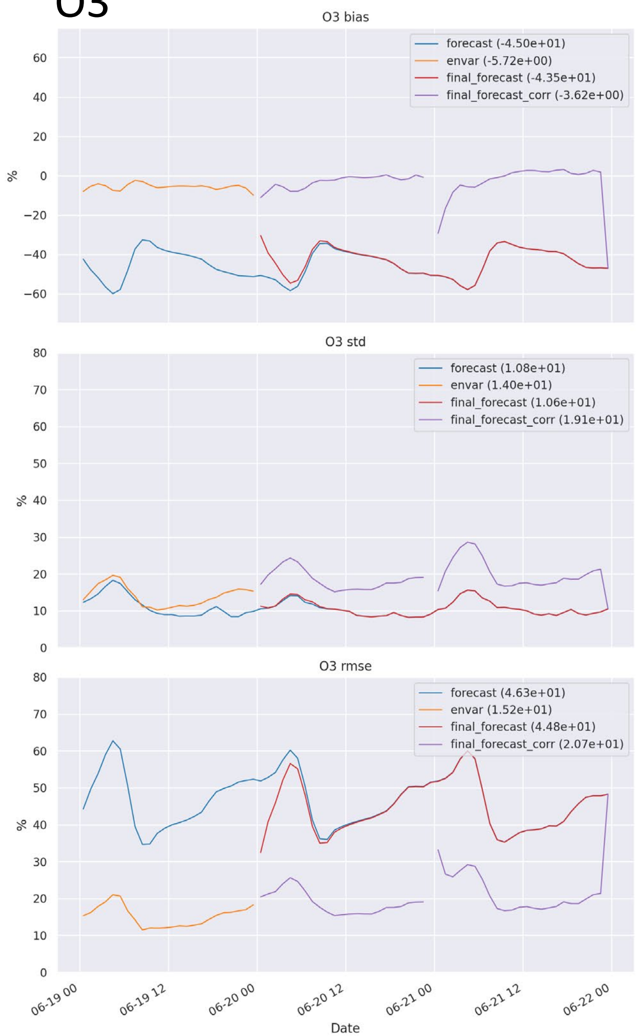
(spatially averaged over continental surfaces)

Bias correction



Bias correction: synthetic experiments

O3

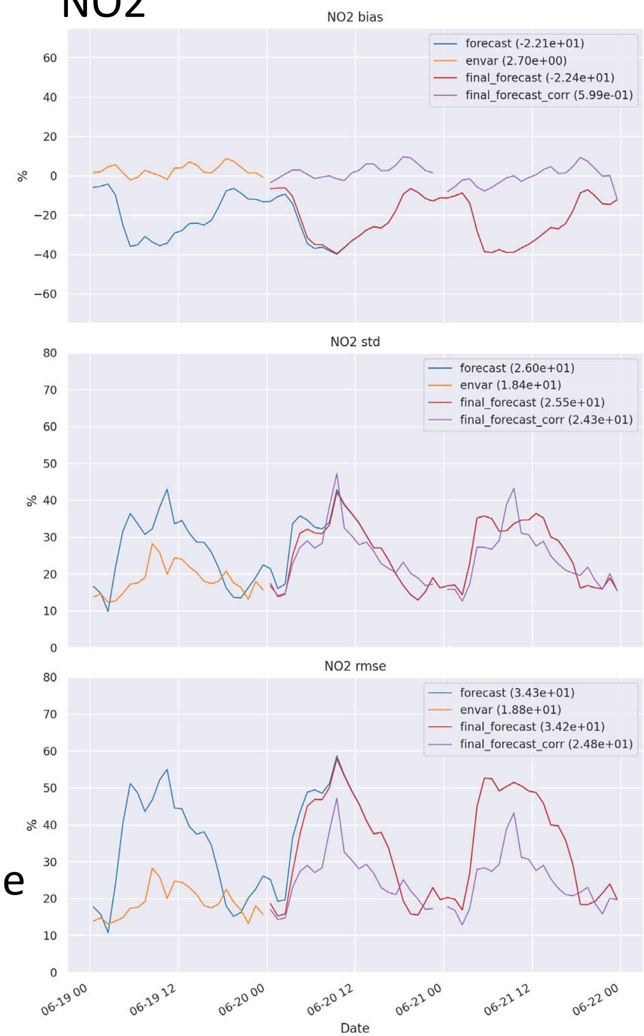


Bias

Std

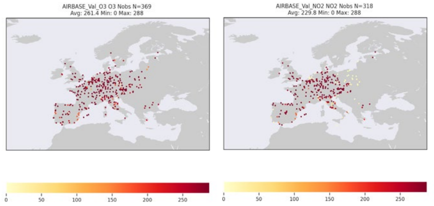
Rmse

NO2

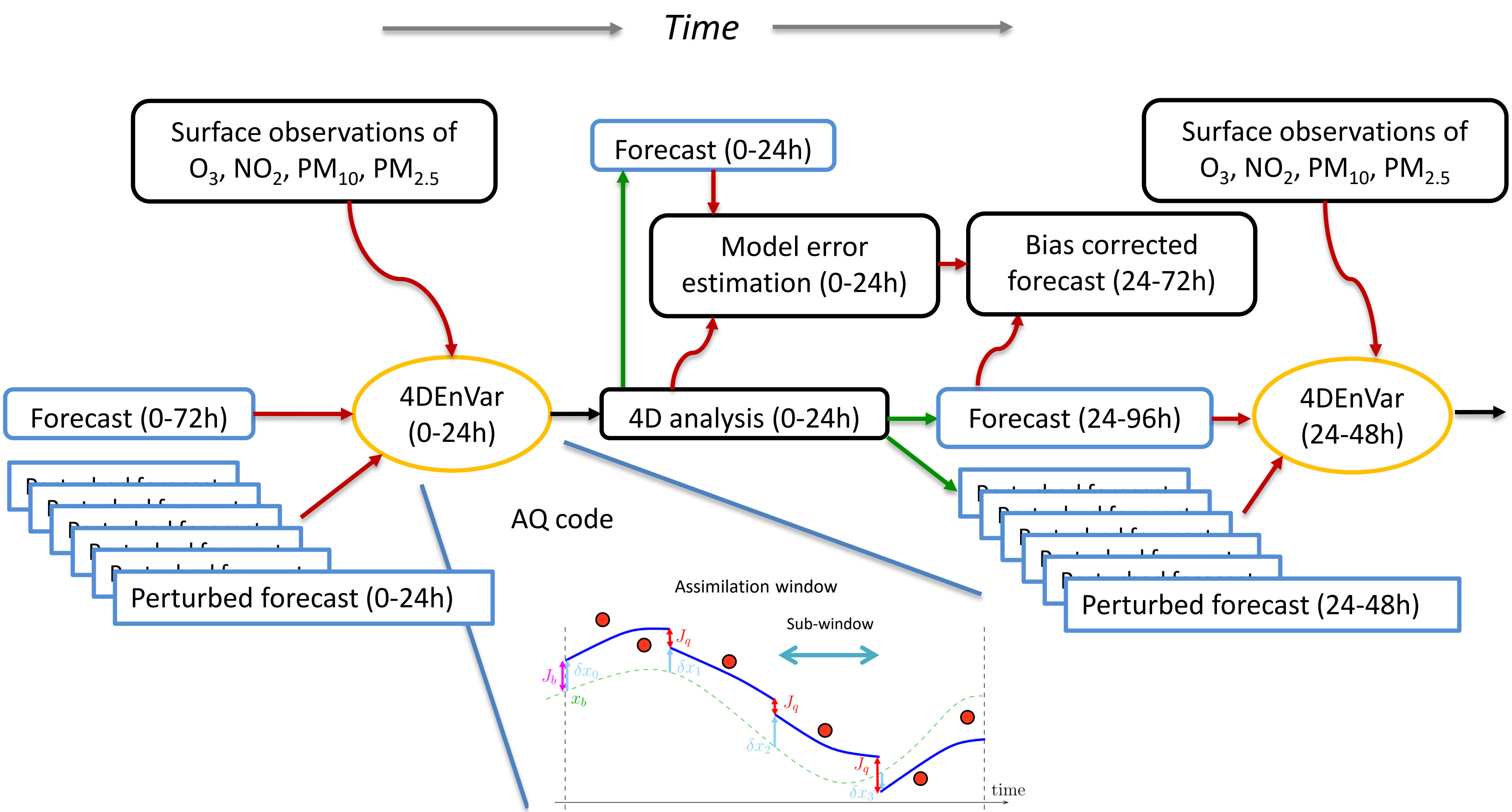


- **Forecast (72h)**
- **4DEnVar analysis (24h)**
- **Forecast initialized with the analysis (48h)**
- **Bias corrected forecast (48h)**

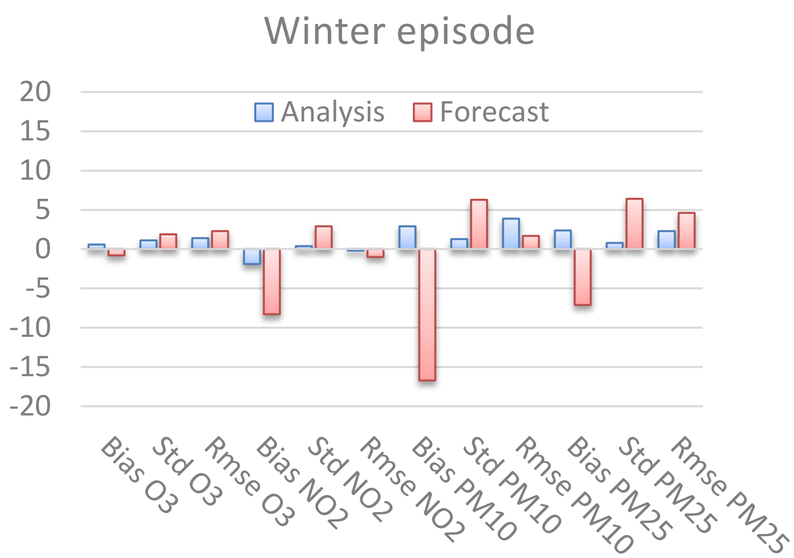
(averaged over all validation stations)



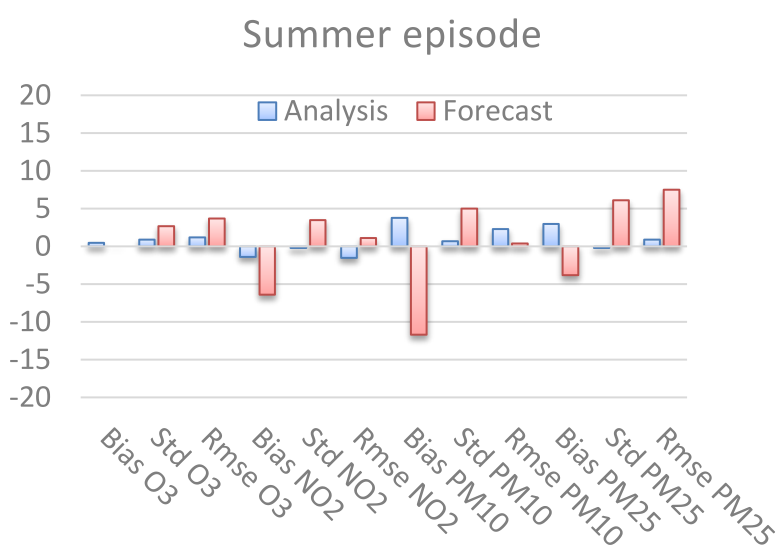
Assimilation cycles with real observations

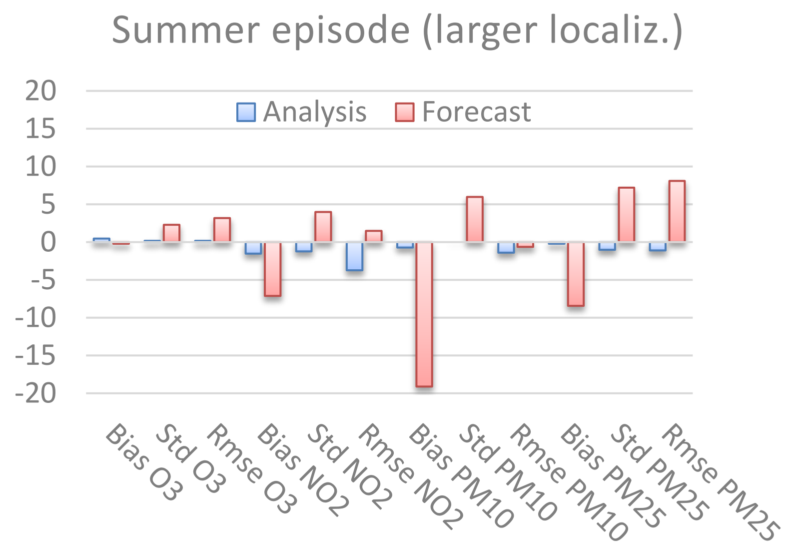


Assimilation cycles: validation statistics



Median improvement of analysis and forecast skills (%) with respect to the operational 3DVar over all validation sites (negative values mean improvement)





Conclusions and perspectives

✓ **Developed a model-agnostic 4DEnVar data assimilation code** (AQ v1.1):

- Only inputs are forecast, ensemble members and observations
- Handles a single analysis cycle with multiple time slots
- Highly optimized (24 hours analysis in about 70 minutes on 200 CPU cores for 10km resolution CAMS domain, 30 ensemble members, 20 species and 7 vertical levels)
- Open source and based on community libraries (JEDI, ATLAS)

✓ **Evaluated the DA and the bias correction strategy within a CAMS operational set-up**:

- Used MOCAGE operational CAMS configuration
- Developed the 4DEnVar workflow for MOCAGE
- Synthetic and real observations experiments
- See the D4.2-4.3 report for more details

Conclusions and perspectives

- 4DEnVar algorithm shows promising results in a controlled environment (better than 3DVar)
- Positive DA results with real observations but not game-changer wrt 3DVar
- Forecast bias correction quite effective but at the cost of degraded standard deviation

- Deeper evaluation of results for unobserved species
- Lumping of coemitted species in ensemble perturbations
- Towards more realistic ensembles that include uncertainties in the chemical scheme (CAMS ensemble) and other key sources of error
- Preparation for the assimilation of satellite data

SEEDS

Sentinel EO-based Emission and Deposition Service

www.seedsproject.eu | info@seedsproject.no

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004318