Co-ordinated by ECMWF

CO_2 Human Emissions

ENSEMBLE SIMULATIONS TO INFORM DATA ASSIMILATION -WP1+WP2+WP3 LINKAGES

March 2019

Nicolas Bousserez, Joe McNorton, Anna Agusti-Panareda, Margarita Choulga, Gianpaolo Balsamo and more

ECMWF

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776186.

The CHE project and its constraints

- Prototype **operational** CO_2 source inversion system (future Copernicus service). Requirements:
  - Support environmental decision-making and implementation of mitigation policies→posterior fluxes and their uncertainties, inform observational system requirements (e.g., satellite instrument design).
  - Computational efficiency (high-resolution and fast).
  - Multi-scale approach (from global to regional to local)→seamless integration of heterogeneous inversion systems.

Operational constraints will drive the methodological choice for the prototype and coordination across CHE's research activities.

ECMWF in the CO_2 community landscape

- Integrated system: online DA capabilities (i.e. joint meteorology/chemistry).
- Near real-time atmospheric composition DA and forecasting (CAMS).
- Ensemble DA system → can provide accurate transport fields together with their uncertainties as inputs for offline CO_2 inversions.
- The Object-Oriented Prediction System (OOPS) allows for modular implementation of DA methods → ideal testbed for testing new DA algorithms (Variational, Ensemble-variational).

ECMWF contribution to CHE

- Coordination of the different WP activities toward building a prototype operational CO_2 inversion system (WP5).
- Support for OSSEs → Tier 1, Tier 2 global nature run simulations.
- **Transport error estimation → ensemble of CO_2 forward model simulations (focus of this presentation).**
- Development and testing of a global anthropogenic CO_2 inversion system based on the IFS.

A global CO_2 source inversion prototype for the IFS

- ➢Leverage online capabilities (joint meteorology/chemistry/emission DA).
- ➢Non-intrusive approach →consistency with existing variational framework.
- ➢Computational efficiency→addition of the CO_2 source optimization should not increase significantly the cost of the analysis.
- ➢Maximize the use of existing operational IFS products (e.g., Ensemble Data Assimilation (EDA)).

Formalism and Notations

- Bayesian framework:

$$\underbrace{p(x|y)}_{\text{posterior}} = \frac{\overbrace{p(y|x)}^{\text{likelihood}}\overbrace{p(x)}^{\text{prior}}}{\int p(y|x)p(x)dx}$$

- Maximum a posteriori (normal distribution):

$$\underbrace{\mathbf{x}^a}_{\text{posterior mode}} \equiv \operatorname{Arg}\max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y})$$

$$= \operatorname{Arg}\min_{\mathbf{x}} J(\mathbf{x})$$

High-dimensional problems → gradient-based iterative minimization using TL/AD models (e.g., 4D-Var).

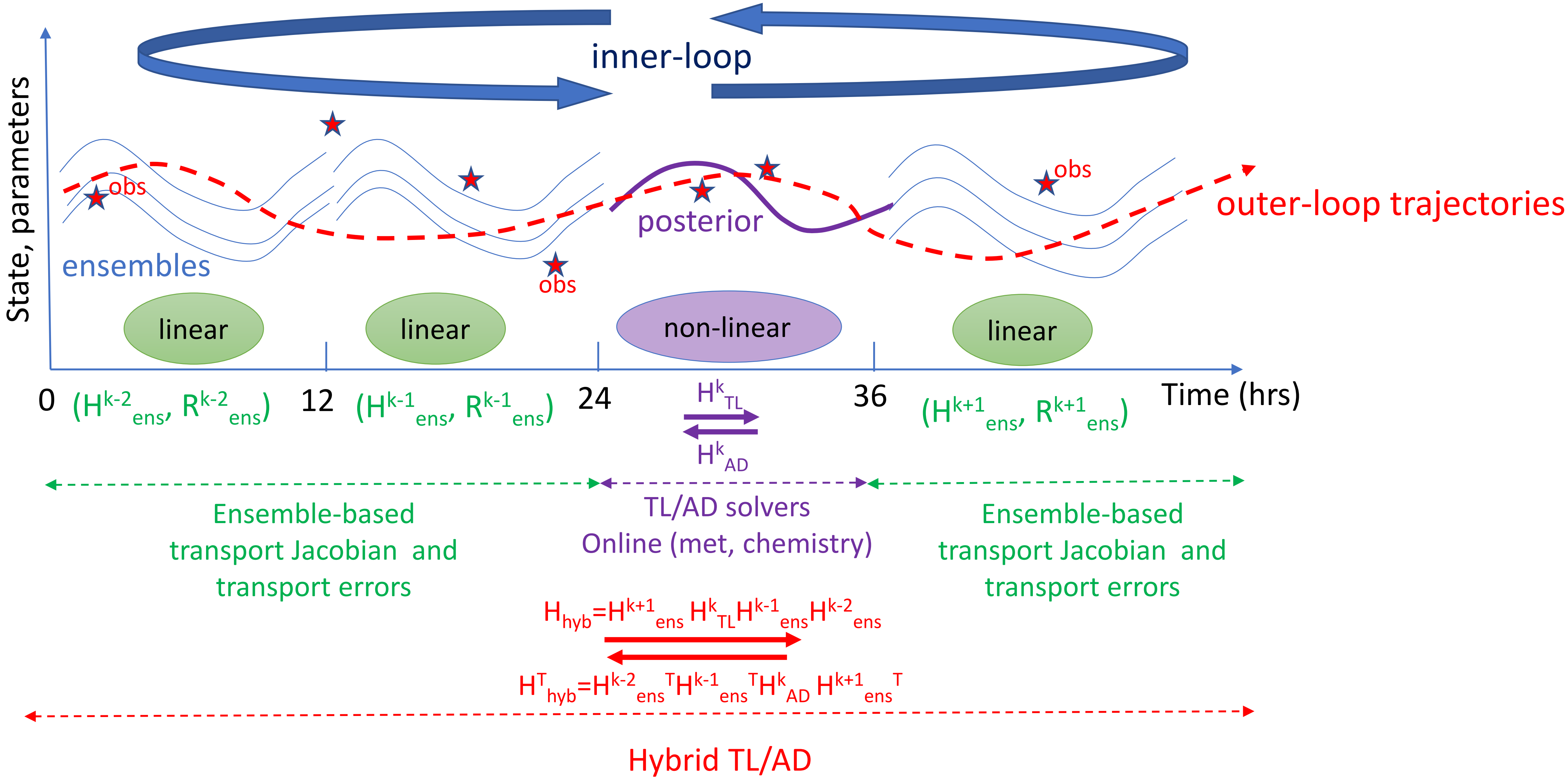
with: $J(\mathbf{x}) \equiv \frac{1}{2}(\underbrace{H}_{\text{forward model}}\mathbf{x} - \mathbf{y})^T\underbrace{\mathbf{R}^{-1}}_{\text{model-data error covariance}}(H\mathbf{x} - \underbrace{\mathbf{y}}_{\text{data}}) + \frac{1}{2}(\mathbf{x} - \underbrace{\mathbf{x}^b}_{\text{prior}})^T\underbrace{\mathbf{B}^{-1}}_{\text{prior error covariance}}(\mathbf{x} - \mathbf{x}^b)$

$$\underbrace{\mathbf{P}^a}_{\text{posterior error covariance}} = (\nabla^2 J)^{-1}(\mathbf{x}_a) = (\mathbf{B}^{-1} + \underbrace{\mathbf{H}^T}_{\text{Adjoint (AD)}}\mathbf{R}^{-1}\underbrace{\mathbf{H}}_{\text{Tangent linear (TL)}})^{-1}$$

Dimensions:

$\mathbf{B}$	: nxn
$\mathbf{H}$	: mxn
$\mathbf{R}$	: mxm
$\mathbf{x}^b$	: n
$\mathbf{y}$	: m

Hybrid long-window 4D-Var (Kalman smoother)



The transport error problem

- Online DA system(i.e., joint meteorology assimilation): transport error implicitly accounted for in **B** matrix→nothing to do!
- Offline DA systems: transport error is represented in **R** matrix (observational error)→ requires ensemble of transport simulations.
- Sampling noise due to small ensemble (a filtering approach is necessary).
- An overlooked issue in long-window offline 4D-Var approaches:
  - Online DA is non-linear →4D-Var window necessarily short (~1 day).
  - Offline DA systems for CO_2 inversion use long-window 4D-Var (weeks to months to multi-years).
  - 4D-Var assumes gaussian transport error distributions→ for a long integration time this might not be valid!

Configuration of Ensemble Simulations

Simulations performed at:

- TCo3099 (**~25km**) resolution with 137 vertical levels.
- 1 month (January 2015).
- **1** hourly output.
- Cycling 24 hour forecasts with analysis initialisation.
- **50** Ensemble Members.

Fluxes:

- EDGAR_2015 anthropogenic emissions gridded onto TCo399 (using **2010 monthly scaling factors**).
- GFAS fire emissions and Takahashi ocean emissions.
- Biogenic fluxes are online generated using CTESSEL.
- A biogenic CO_2 flux adjustment scheme (BFAS) is used to mitigate large-scale model bias.

Perturbations:

- Anthropogenic emissions perturbed using **WP-3** uncertainties per sector & country for 7 ECMWF sectors (see Margarita Choulga's poster).
- Biogenic emissions are member specific, forced by meteorology.

Ensemble Simulations

Initial Concentration

Informed from high resolution (~9km) nature simulation

Prior Emissions

Perturbed fluxes based on uncertainties

Meteorology

Introduce tracers to current IFS-EPS

Perturbed Ensemble Simulations

Three simulation types:

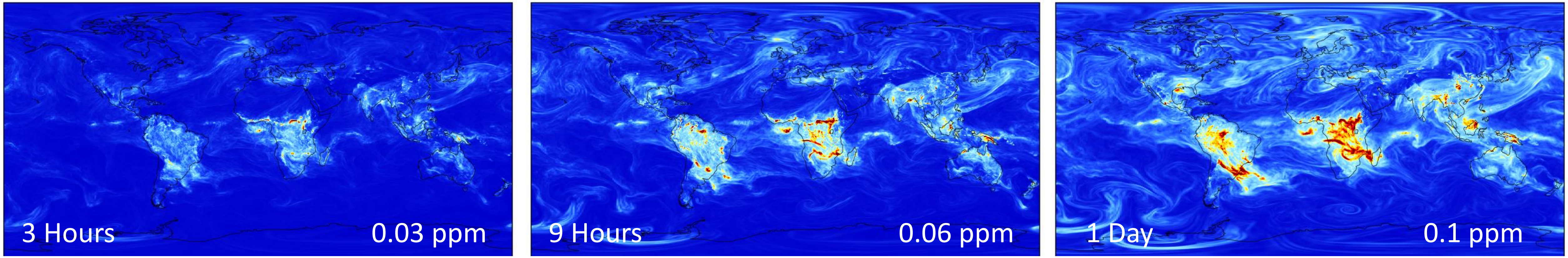
Meteorology	Emissions	Resulting Information
Perturbed	Fixed	Transport Error
Fixed	Perturbed	Transport Jacobian
Perturbed	Perturbed	Transport Jacobian with Noise

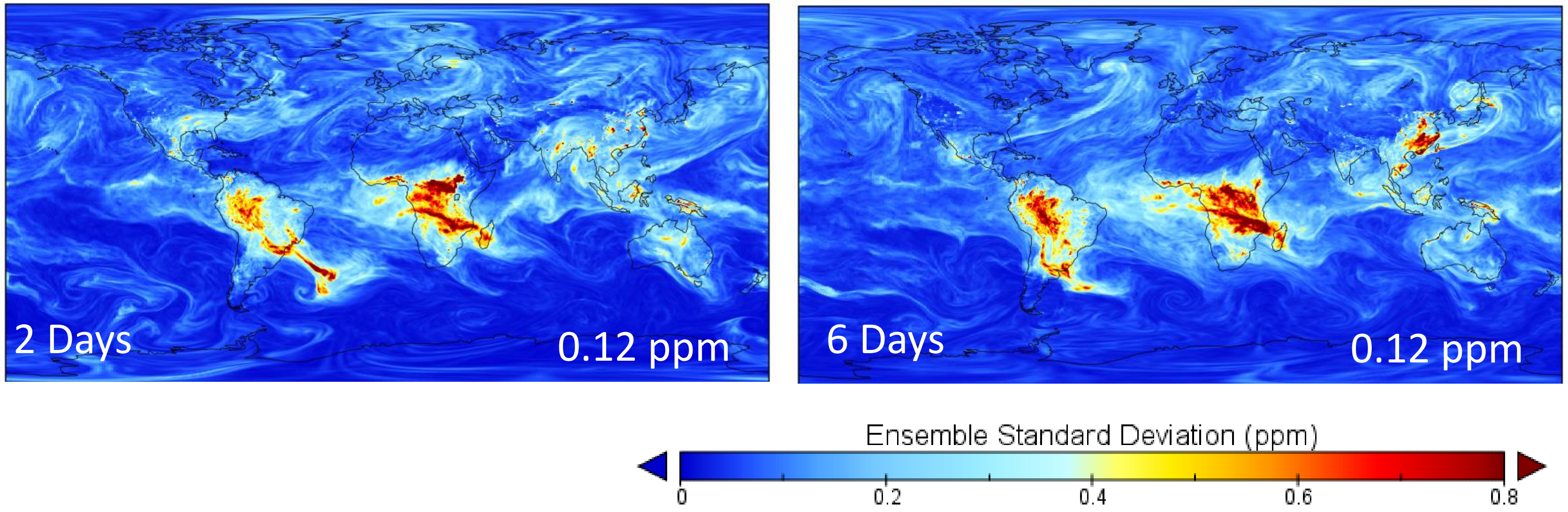
Ensemble-based Inversion System

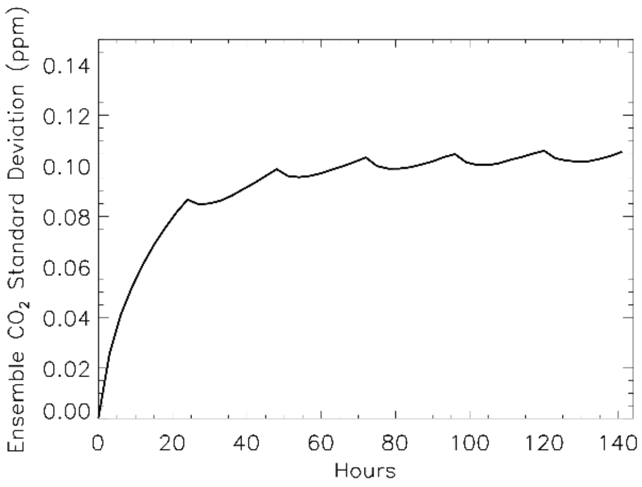
Test multiple inversion systems to estimate sector/national posterior fluxes and errors

Model Transport Error – global

Meteorology	Emissions
Perturbed	Fixed



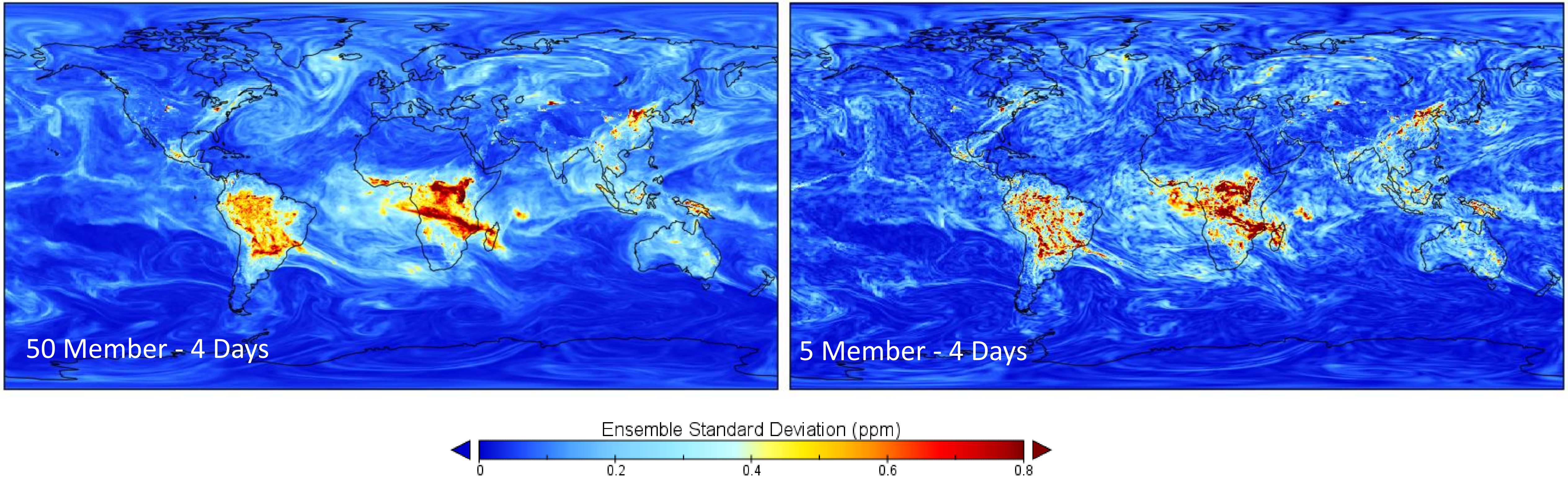




- Global transport error in XCO_2 stabilises after 2-3 days.
- Stabilization results from the impact of diffusion over longer time-scales.

Meteorology	Emissions
Perturbed	Fixed

Model Transport Error – ensemble size

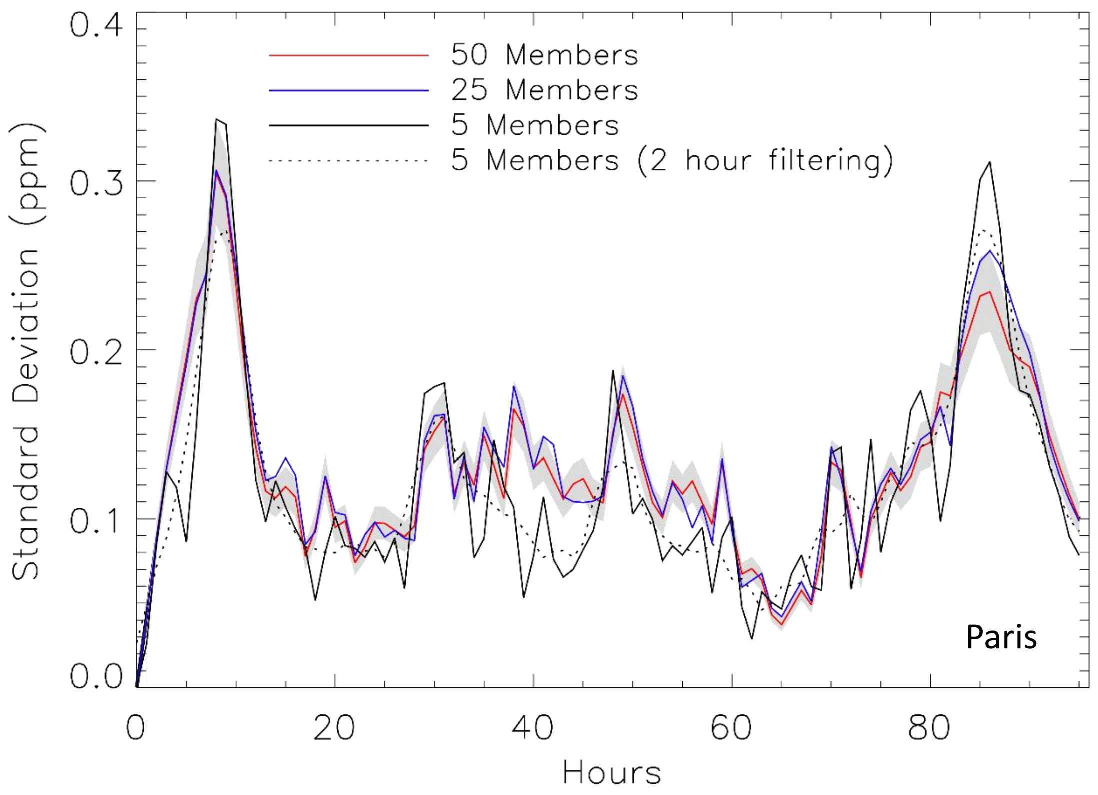


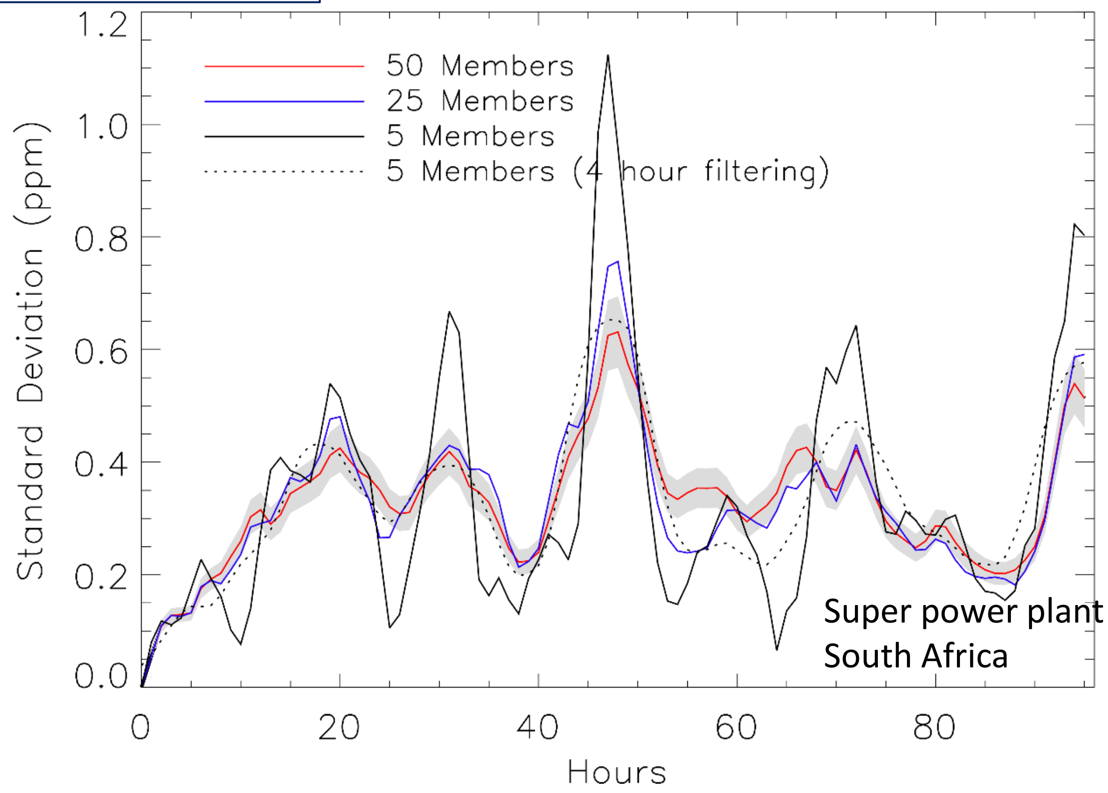
➢Reducing ensemble size increases significantly transport error noise.

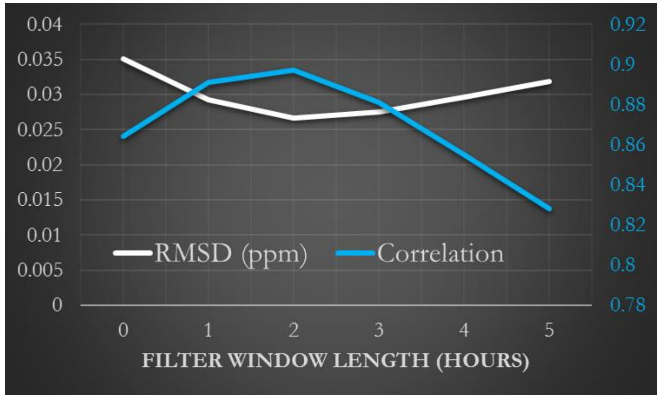
Model Transport Error – temporal filter

Meteorology	Emissions
Perturbed	Fixed

Filter= *n* hours centered average

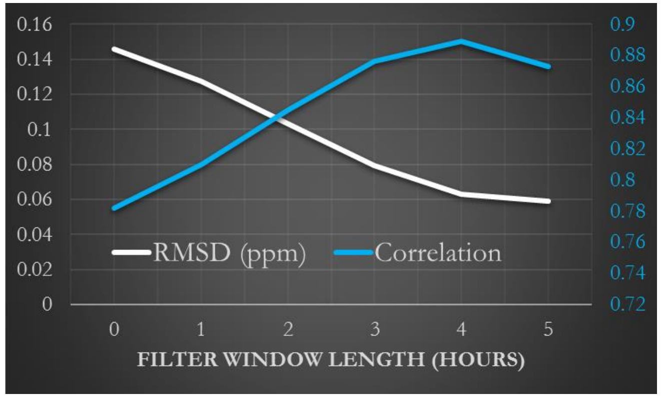






Agreement between 50 (reference) and 5 member ensemble using different time-windows for the filter.

Optimum :
2 hour filter

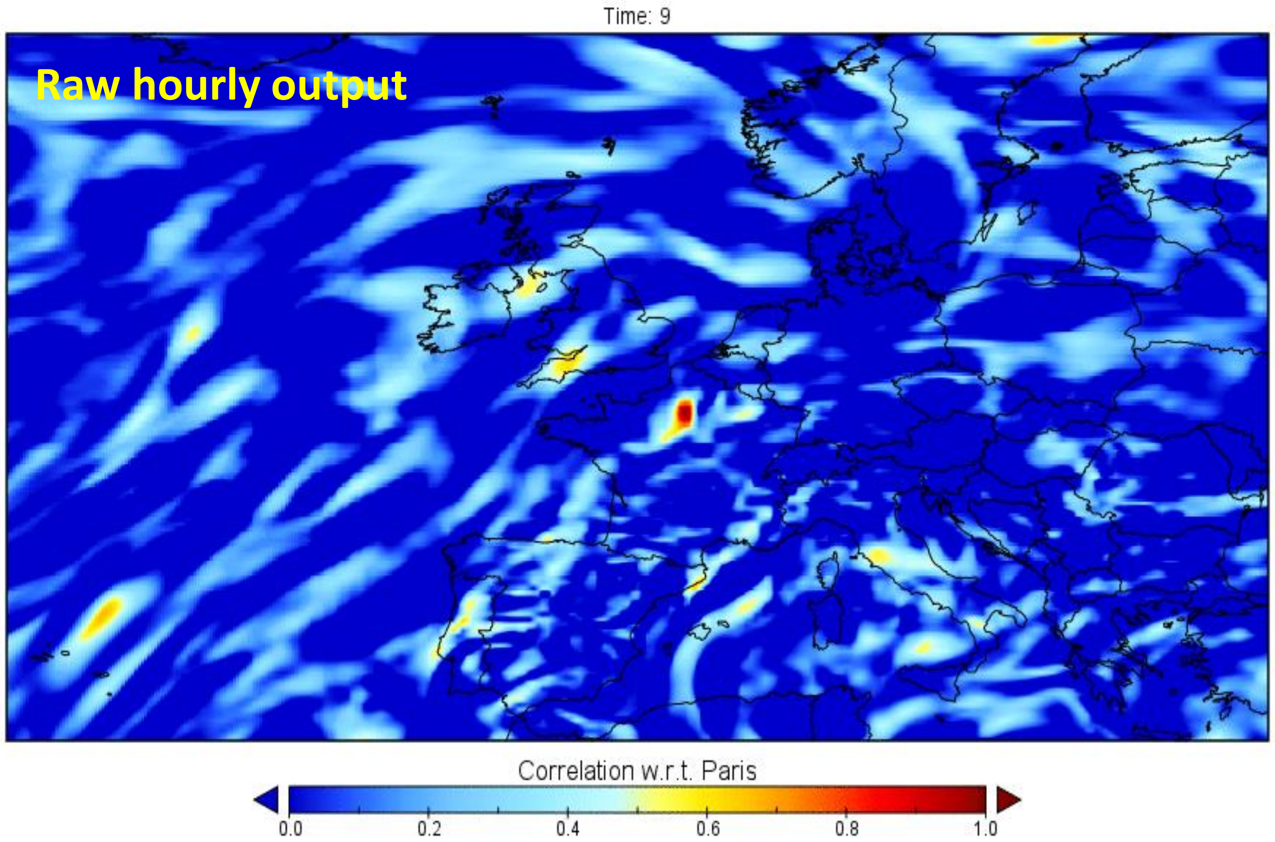


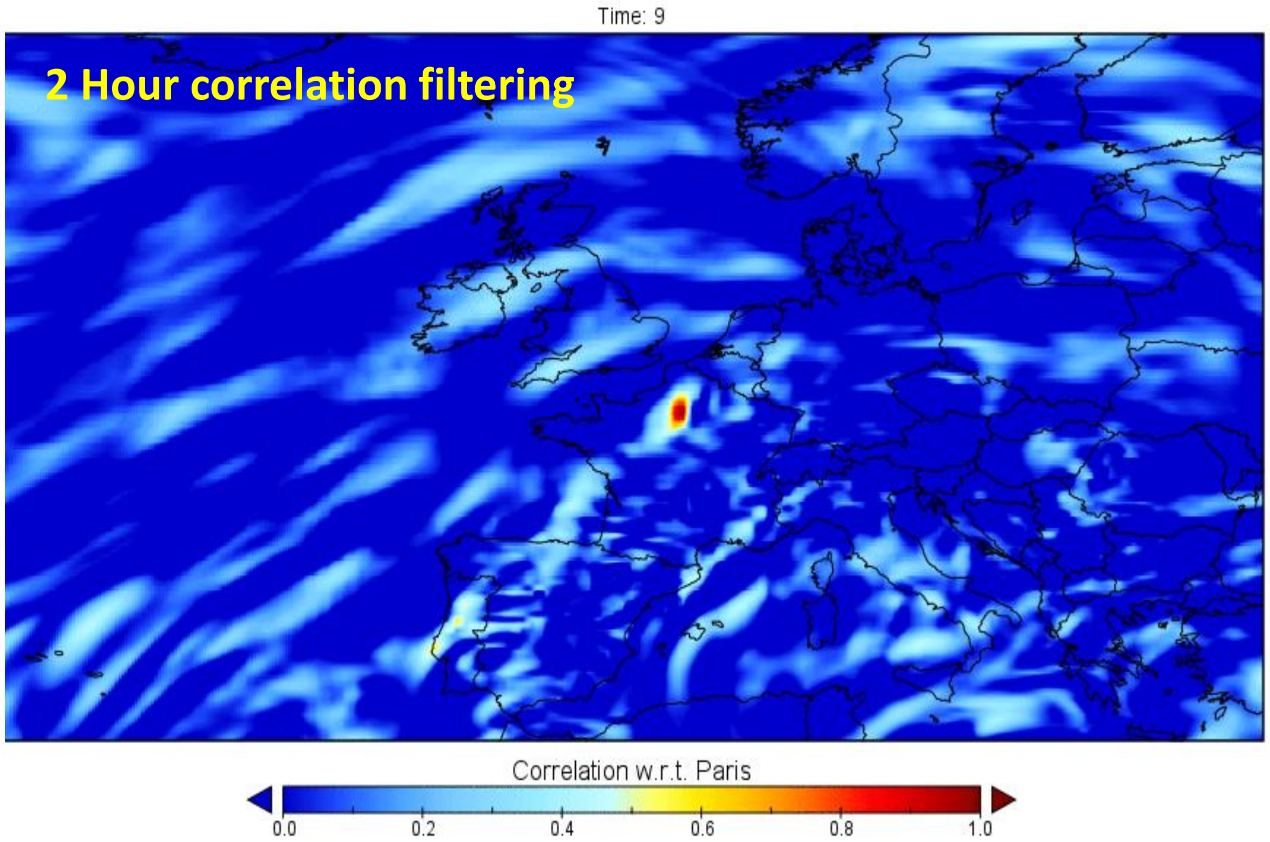
Optimum :
4 hour filter

Meteorology	Emissions
Perturbed	Fixed

Model Transport Error Correlations

Animations showing XCO_2 correlations with respect to Paris XCO_2 for 50 member ensemble.





- ➢ The spurious noise in the raw correlations is removed with a 2-hour window smoothing.
- ➢ 2-hour window smoothing maintains heterogeneous structures.

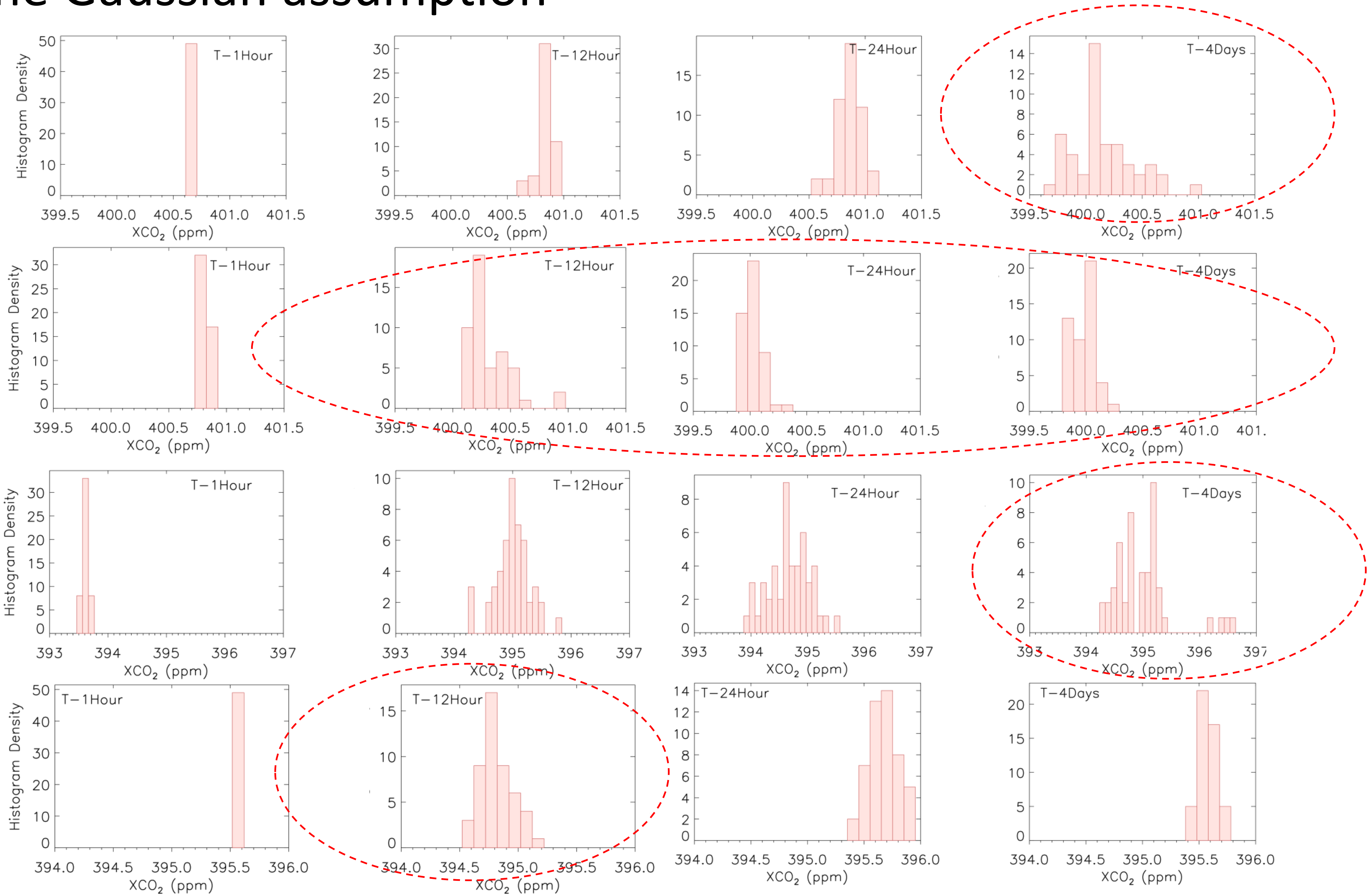
Validity of the Gaussian assumption

New York

Paris

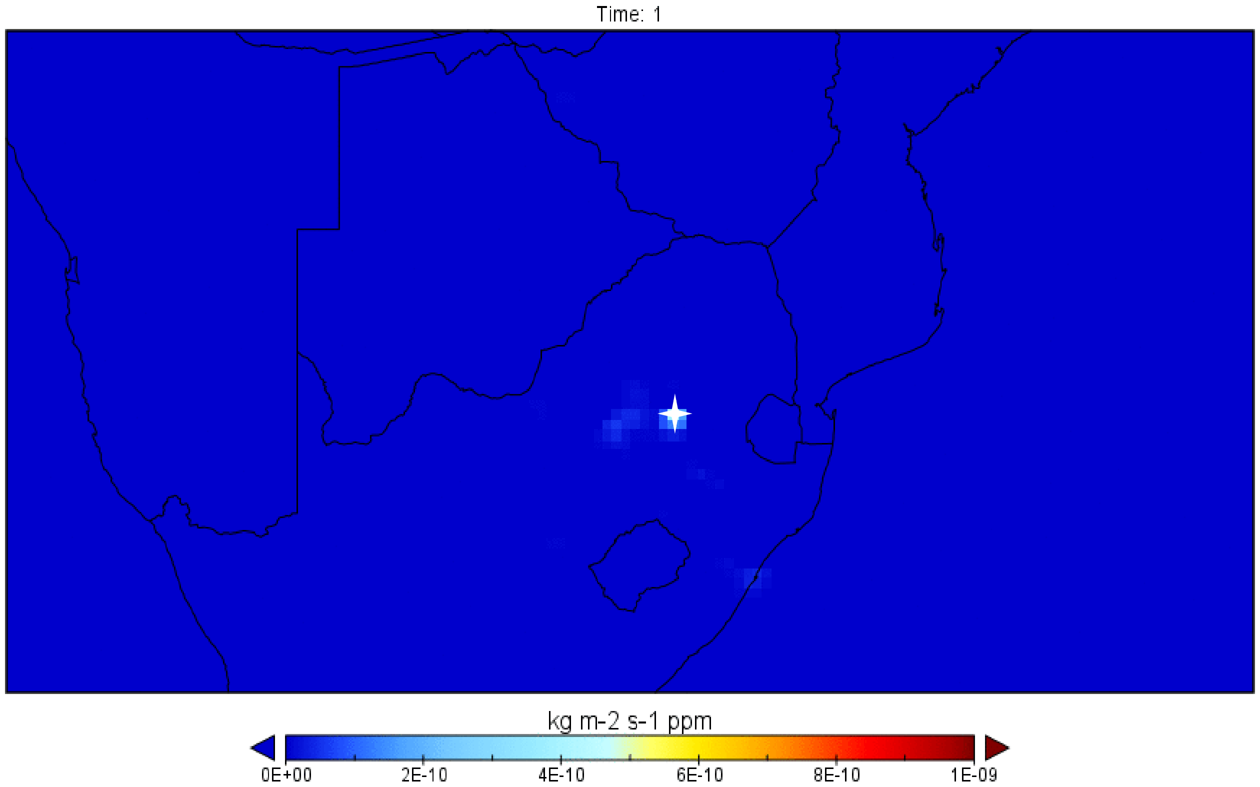
South Africa Power station

Sydney

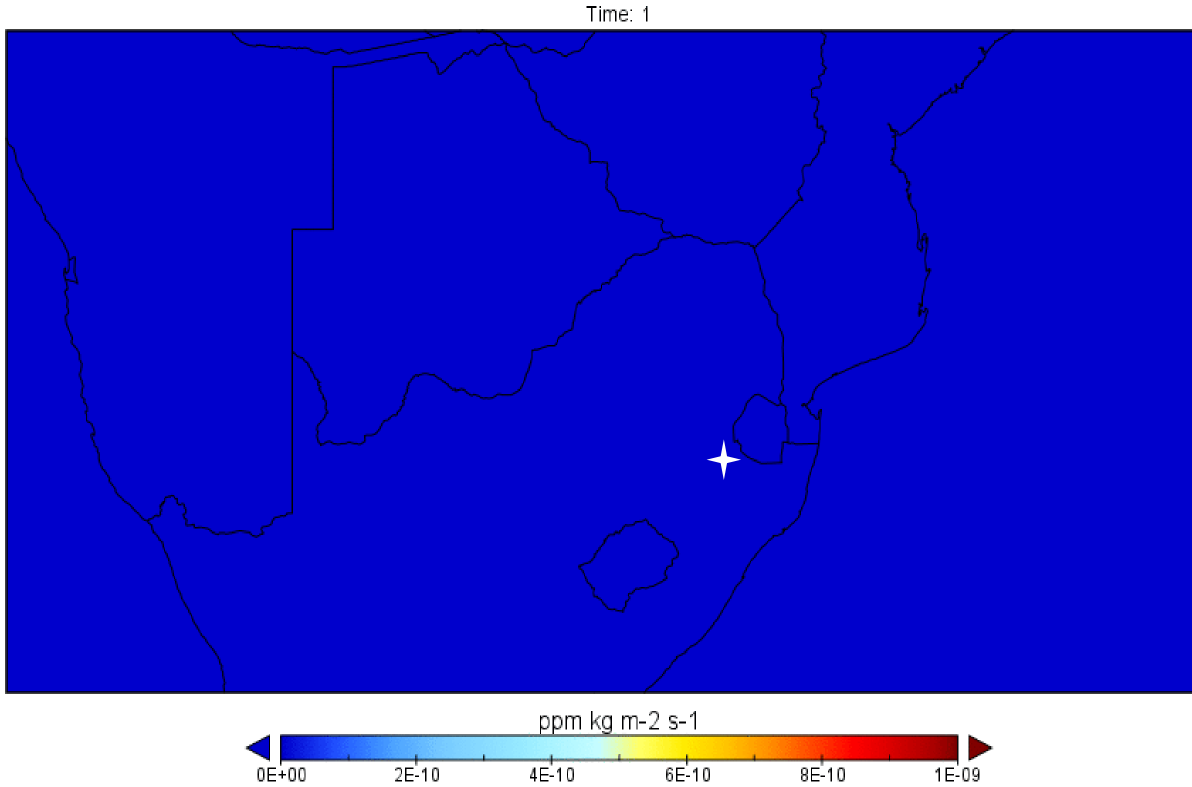


Meteorology	Emissions
Fixed	Perturbed

Flux-column relationship



Covariances between point source and XCO_2 columns
(~forward sensitivities)



Covariances between point XCO_2 column and sources
(~backward sensitivities)

Improved inversion diagnostics and optimization performance

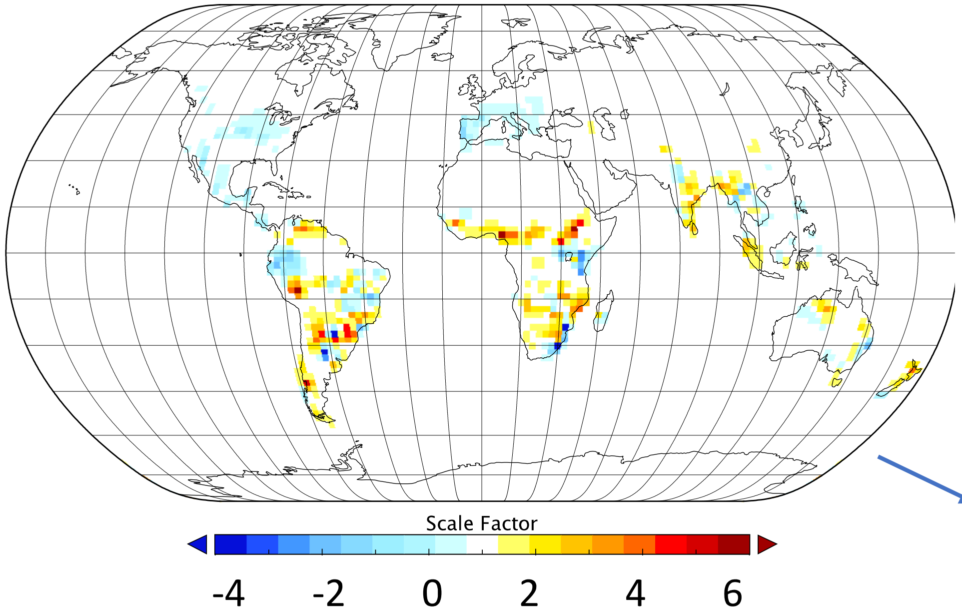
- Posterior diagnostics will include:
  - Posterior error covariances.
  - Effective resolution of the inversion (i.e., what are the spatiotemporal flux patterns being constrained by the observations?)
- New highly parallel stochastic optimization algorithm will be implemented and tested in the IFS to improve computational performance.
- This framework is currently tested in NASA's JPL CMS-Flux system (CO_2) and has been successfully applied to CH_4 GOSAT-based inversions (Bousserez and Henze, 2018).

What are the observational constraints on the CO_2 fluxes?

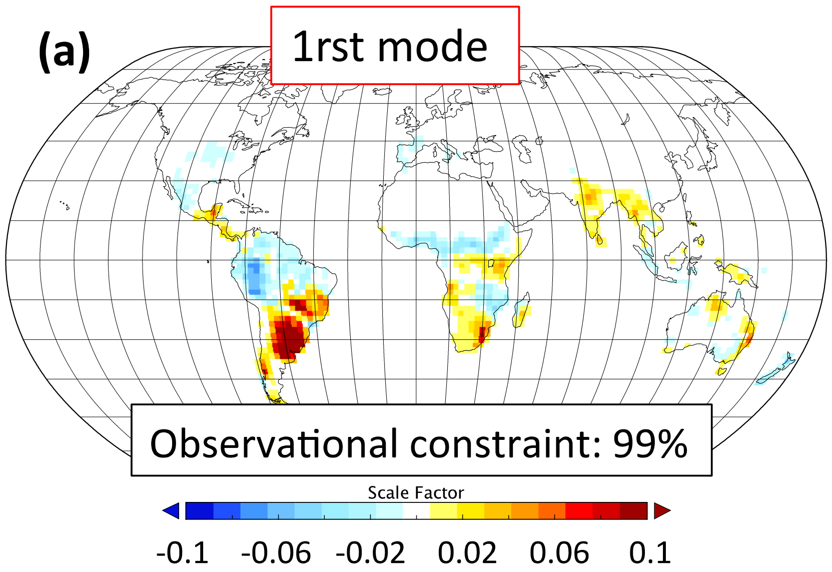
Test Case: OCO-2 data, January-only, 2015

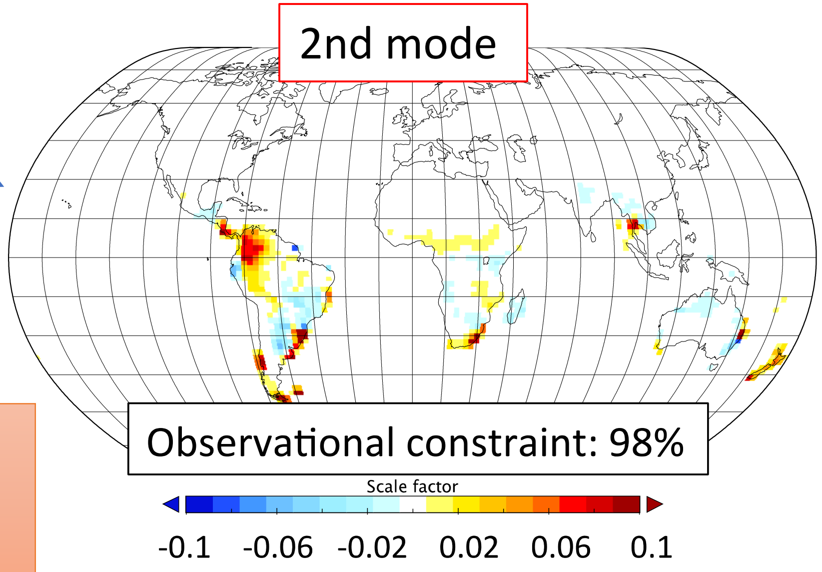
- Flux estimates are computed for ocean, land, and fossil fuel using FRODO with 150 ensembles.
- A priori fluxes follow Liu et al., 2017 and Bowman et al., 2017.
- Flux uncertainty is simply 50%.
- What patterns are constrained by the inversion?

Posterior mean scaling factors



Independently constrained flux patterns

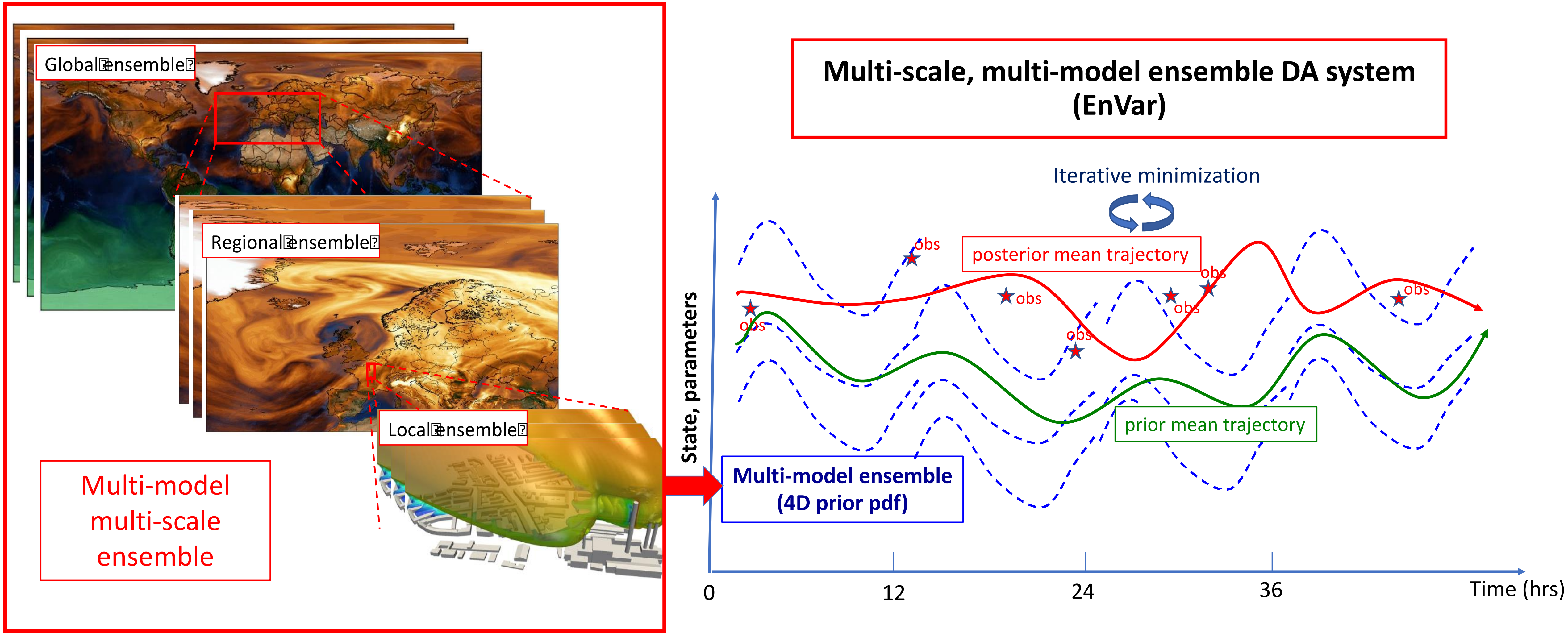




Ongoing collaboration with NASA-JPL (Carbon Monitoring System (CMS) project) using the Fast Randomized and Optimal method for Diagnostics and Optimization (FRODO) (Bousserez and Henze, 2018).

Courtesy Kevin Bowman

How can we optimize the synergy between CHE activities toward building a multi-scale CO_2 inversion system?

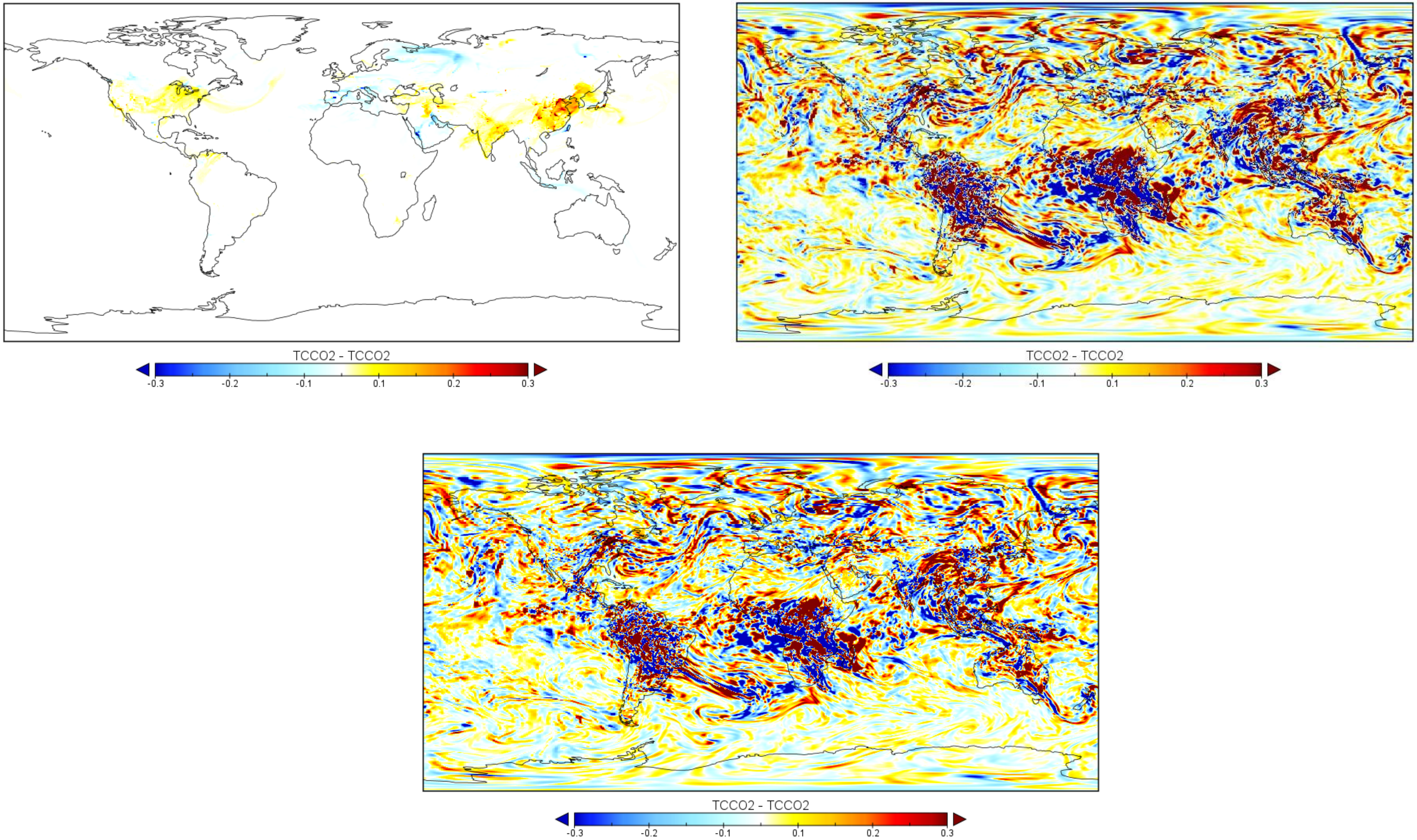


- Each model provides an ensemble of prior forward simulations.
- Prior error flux/state covariance matrix is built from the set of all (multi-model) ensemble forward simulations.
- Posterior mean fluxes are linear combinations of prior ensemble fluxes → multi-model, multi-scale.
- Conceptually simple, provides an efficient framework for collaborations (partners can work on their ensemble of simulations (almost) independently).

Conclusion

- ECMWF is working on a hybrid ensemble-variational system for global anthropogenic CO_2 source inversion.
- IFS ensembles have been used to quantify transport errors and flux-observations sensitivities.
- Transport errors stabilize after 2-3 days.
- Ensemble-based transport error covariances require sampling noise filtering→preliminary tests with a temporal filter show encouraging results.
- Gaussian assumption on transport errors→ errors can become quickly (~day) non-gaussian.
- A paper on transport errors in global CO_2 inversions is in preparation.
- Ensemble-based sensitivities would provide a computationally tractable approach for network design studies (adjoint-based methods are costly).
- Improved diagnostic and computational tools are being tested in a global inversion system (NASA JPL's CMS-Flux) before implementation in the global IFS prototype.
- Volunteers to participate to an experiment with a multi-model ensemble-variational approach? Collaboration would only require to provide an ensemble of perturbed (flux/transport) forward CO_2 simulations. Come talk to us if interested…

Backup slides


TCCO2 - TCCO2
-0.3 -0.2 -0.1 0.1 0.2 0.3
TCCO2 - TCCO2
-0.3 -0.2 -0.1 0.1 0.2 0.3
TCCO2 - TCCO2
-0.3 -0.2 -0.1 0.1 0.2 0.3

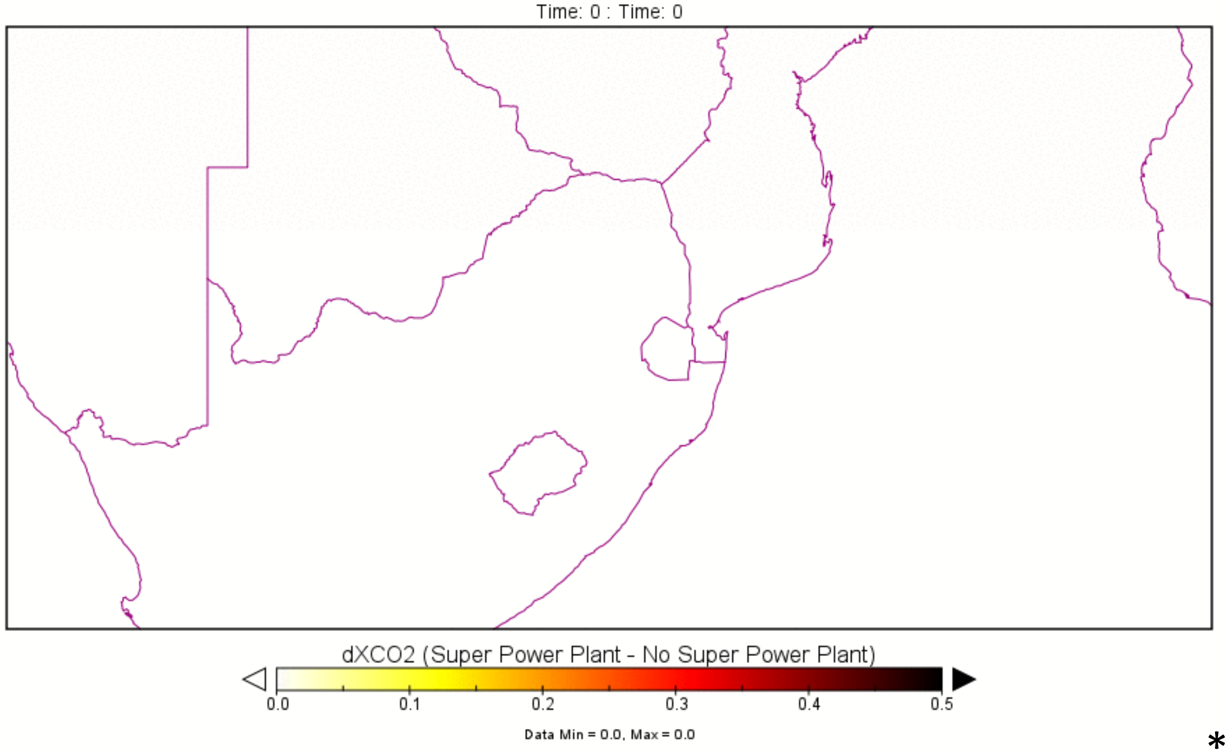
Testing the XCO2 atmospheric signal from the super power station.

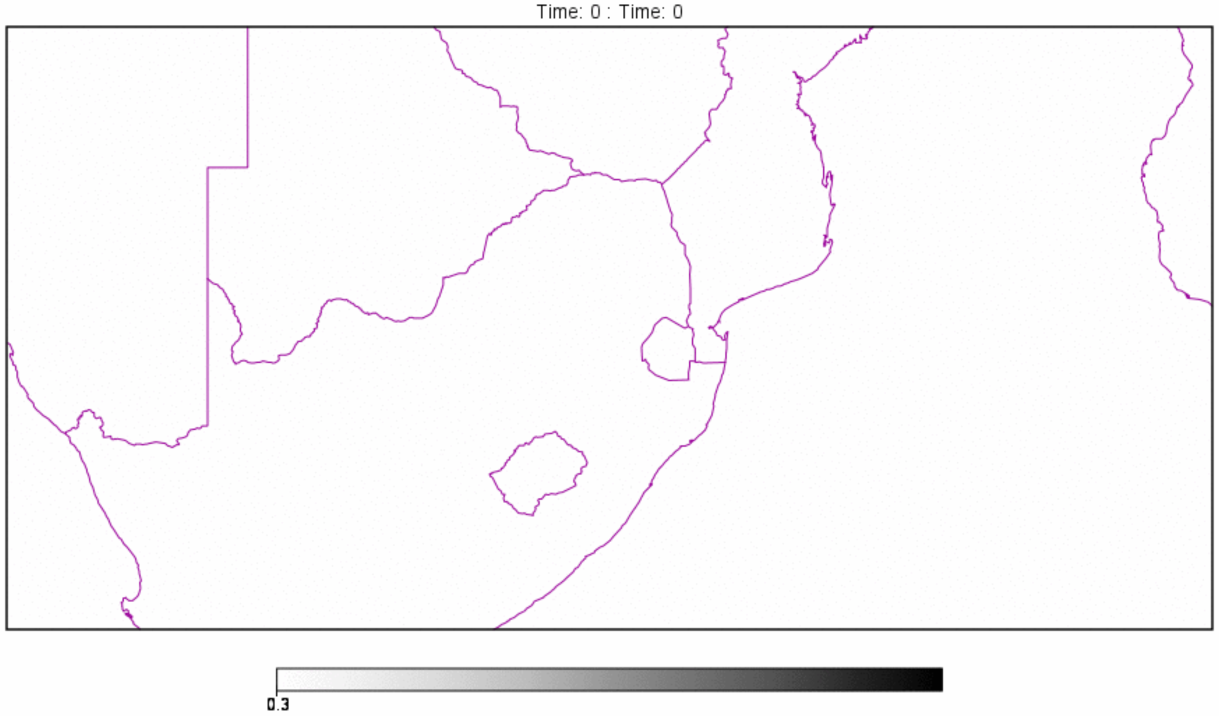
To establish an idealised detection threshold limit two simulations are compared with the following configuration :

- ~25 km resolution
- 137 levels
- 3 hourly output

(i) Using fully resolved EDGARv4.3.2-2015 monthly emissions (EDGARv4.2-2010 scaling factors applied).

(ii) Using the same emissions with "super" emitters (>8.3×10^{-6} kg m^{-2} s^{-1}).

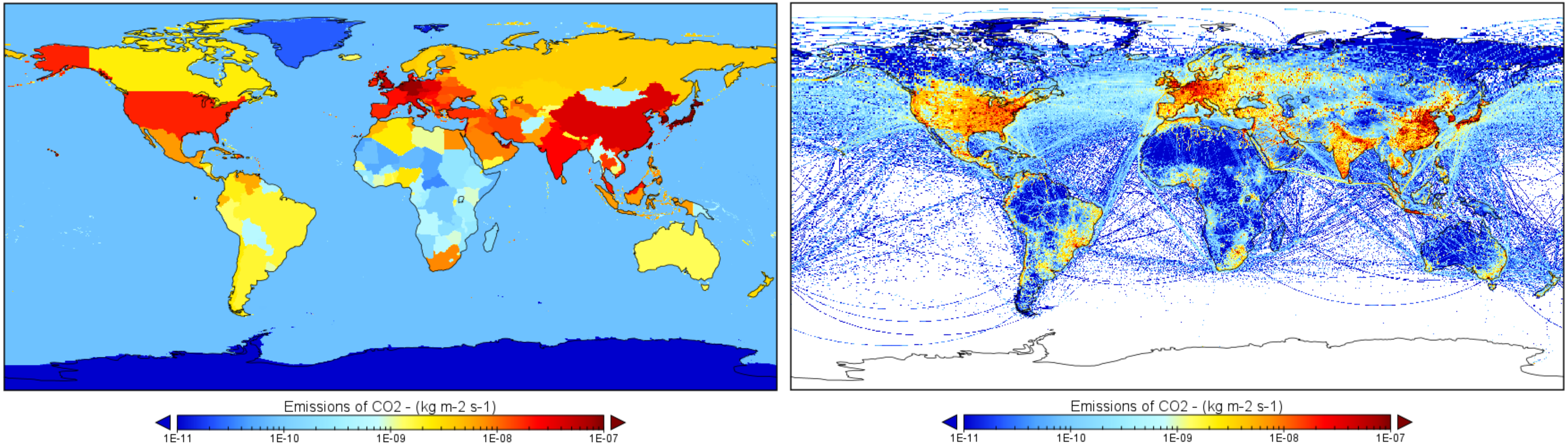


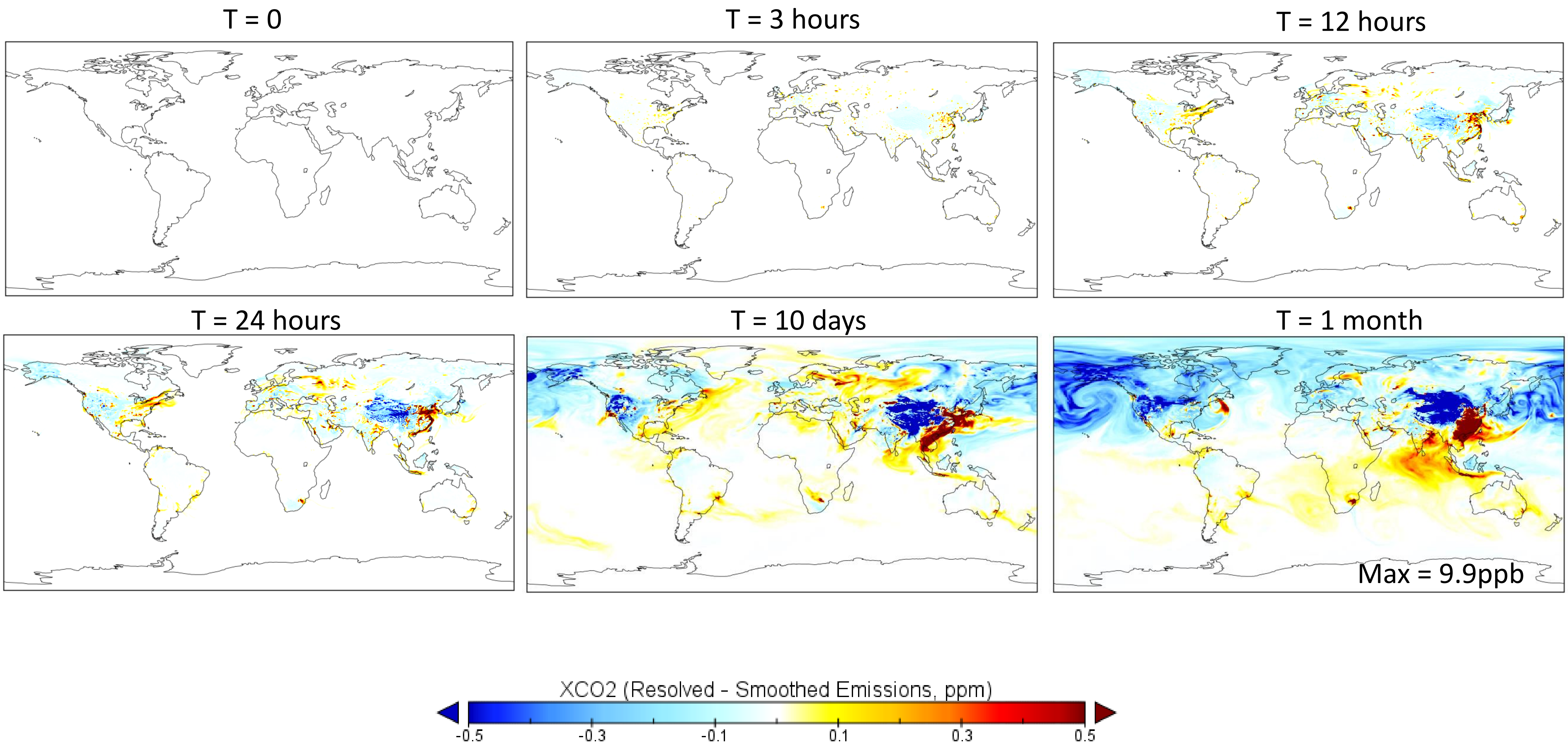


* Total column difference that exceeds 0.3 ppm as an idealised constant observation limit.

Representativeness error in the prior model

To assess the importance of accurate spatial distribution in the prior two simulations performed with (i) uniformly smoothed emissions per country and (ii) EDGARv432-2015 resolved emissions.




T = 0
T = 3 hours
T = 12 hours
T = 24 hours
T = 10 days
T = 1 month
Max = 9.9ppb
XCO2 (Resolved - Smoothed Emissions, ppm)
-0.5
-0.3
-0.1
0.1
0.3
0.5