Implementation of a hybrid ensemble-variational inversion system in the Integrated Forecast System (IFS)

Nicolas Bousserez, L. Cantarello, CAMS, CoCO2 & CORSO team

Atmosphere Monitoring

PROGRAMME OF THE EUROPEAN UNION

Copernicus
Europe's eyes on Earth

IMPLEMENTED BY
ECMWF

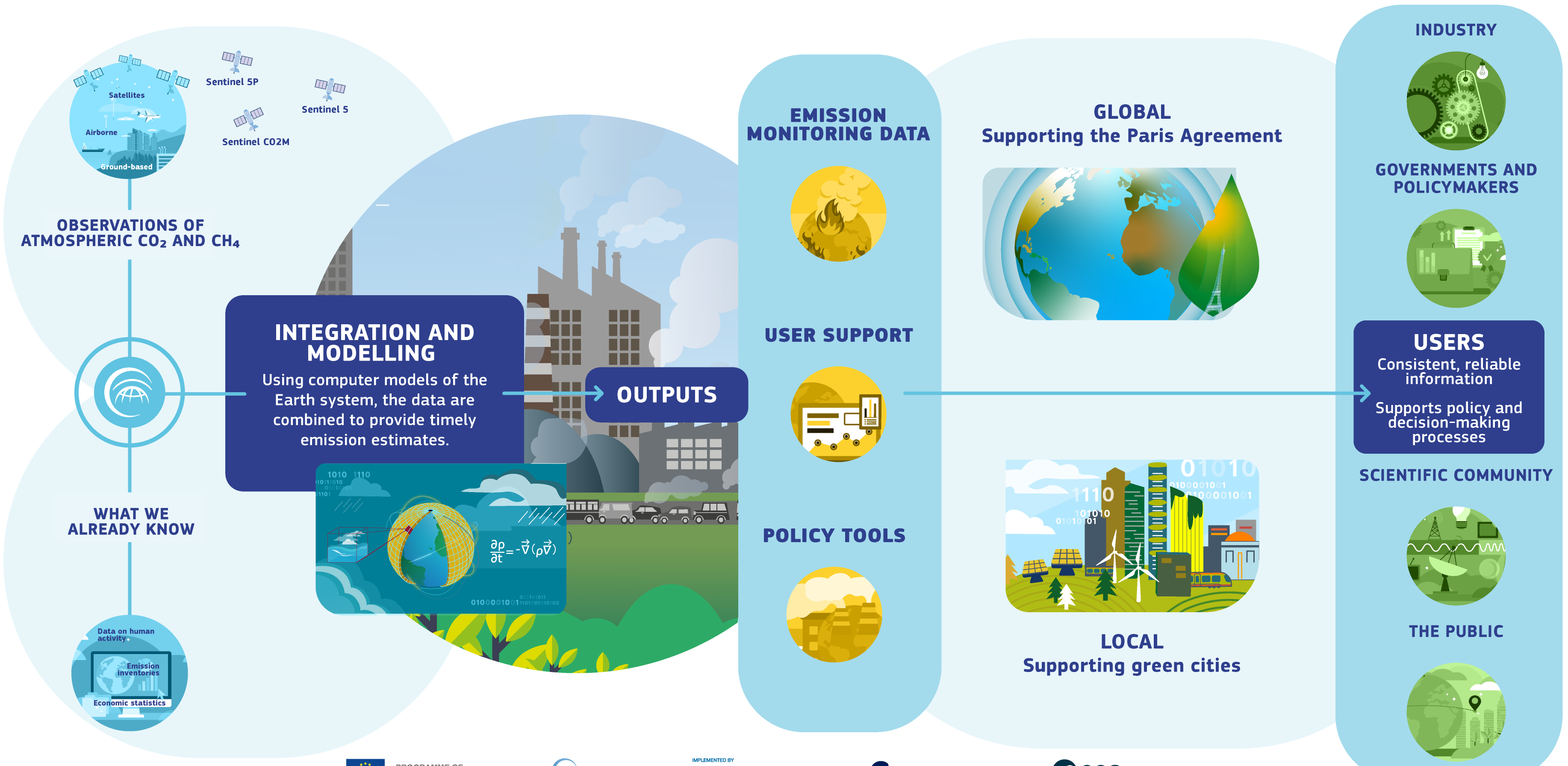
Why an operational emission inversion system?

Atmosphere Monitoring

- **GHG monitoring :**
  - Atmospheric monitoring of GHG together with accurate source & sink attribution are key to address climate change and design efficient mitigation policies.
  - Spatio-temporal scales of bottom-up inventories (country, annual budget) limit our ability to understand the complex interplay of processes that drives the observed atmospheric GHG growth rate.
  - Top-down inversion methods based on near-real time in situ and remote sensing measurements provide timely updates on sources & sinks controlling the temporal variability of GHG
- **Air quality application:**
  - Air quality forecast primarily driven by emissions of pollutants and precursors.
  - Significant day-to-day variability is not well captured by traditional bottom-up inventory.

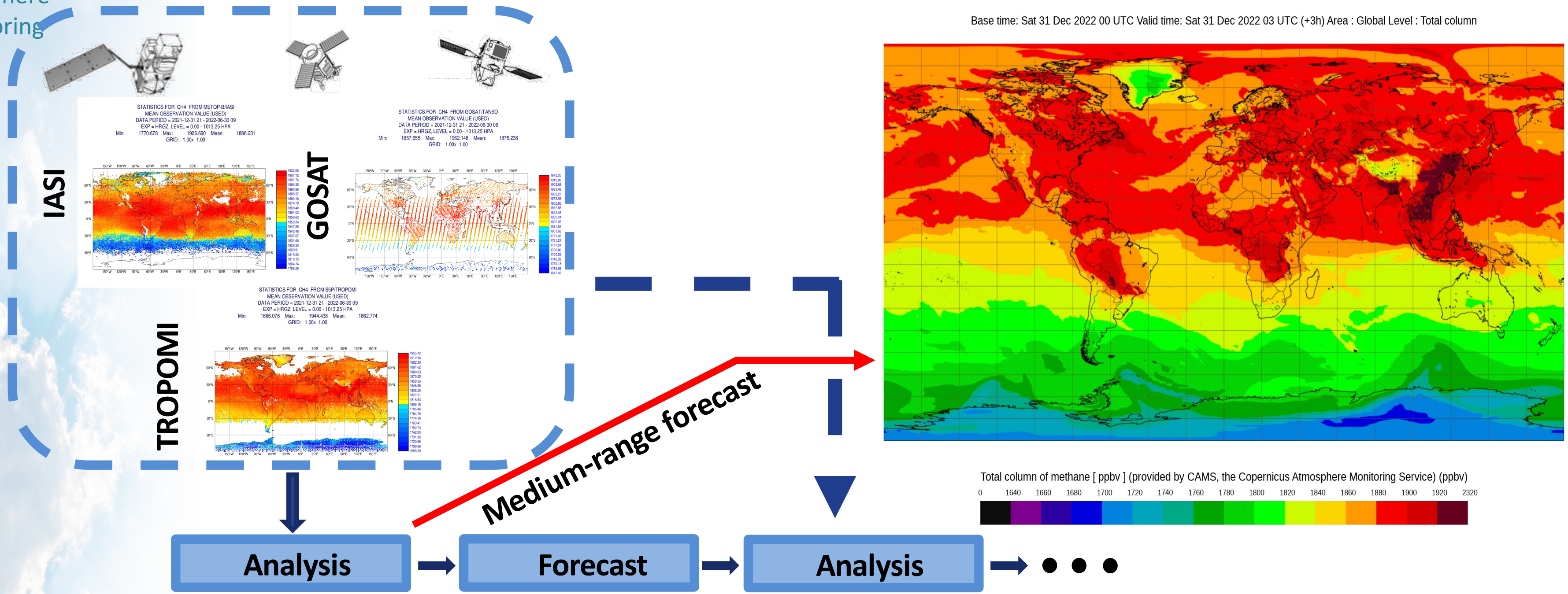
Greenhouse gas emissions monitoring capacity

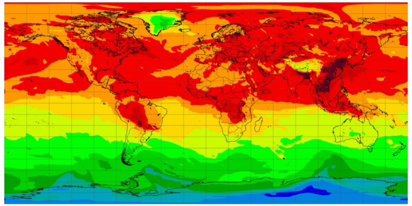
Atmosphere Monitoring Service
atmosphere.copernicus.eu



PROGRAMME OF THE EUROPEAN UNION
Copernicus Europe's eyes on Earth
IMPLEMENTED BY ECMWF
EUMETSAT
esa

CAMS current products: Analyses and forecasts





4D-Var daily 3D state analyses at ~25km and forecasts at ~9km resolution

PROGRAMME OF THE EUROPEAN UNION

IMPLEMENTED BY ECMWF

Atmosphere Monitoring

Towards a joint state/emission optimization system: general approach

- ***Integration***: A global operational emission inversion system leveraging the IFS DA capability & extending it while maintaining its computational efficiency for timely delivery of products.
- ***Multi-species***: A methodology exploiting synergies between different types of atmospheric observations→NO_2 and CO constraints on CO_2 emissions.
- ***Multi-scale, multi-model***: A global inversion system generalising the assimilation of traditional observations to the assimilation of high-resolution top-down regional & point sources estimates based on different models & approaches.
- ***Uncertainty quantification:*** A robust posterior error estimation product leveraging current IFS approaches (Ensemble of Data Assimilation (EDA), 4D-Var Hessian approximation).

4D-Variational inversion

$$J(\mathbf{x},\mathbf{p}) = (\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}_{\mathbf{x}}^{-1}(\mathbf{x}-\mathbf{x}_b)+(\mathbf{p}-\mathbf{p}_b)^T\mathbf{B}_{\mathbf{p}}^{-1}(\mathbf{p}-\mathbf{p}_b)+(\mathbf{y}-h(\mathbf{x},\mathbf{p}))^T\mathbf{R}^{-1}(\mathbf{y}-h(\mathbf{x},\mathbf{p}))$$

state (prognostic) — parameter (e.g., emission scaling factors) — observations (meteorology, atmospheric composition)

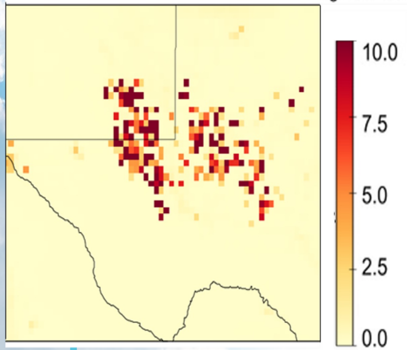
- <u>Characteristics of current system:</u>
  - **Online system**: joint meteorology & chemistry state/fluxes 4D-Var optimisation
  - 12-hour or 24-hour window
  - Emissions: CO_2, CH_4, NO_x, CO
  - Biogenic CO_2 fluxes (GPP and respiration): process-based online prior
  - Observations: OMI NO_2; TROPOMI NO_2, CO, CH_4 ; IASI CH_4; GOSAT CO_2, CH_4; OCO-2 & OCO-3 CO_2
  - **B** model: spatial error correlations
- Tangent linear and adjoint models of simplified chemistry mechanism
- Posterior error covariance estimation based on ensemble of data assimilation (EDA) approach (i.e., Monte-Carlo)
- <u>Limitation</u> : optimization of GHG fluxes requires much longer assimilation window (~months to years)→ building a long-window hybrid ensemble-variational system

Atmosphere

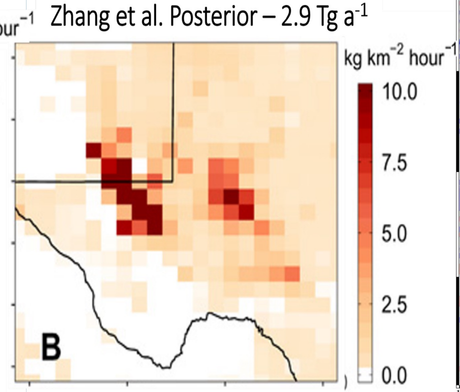
CH_4 inversion over Permian Bassin

IFS posterior

(Prior – 1.9 Tg a^{-1})



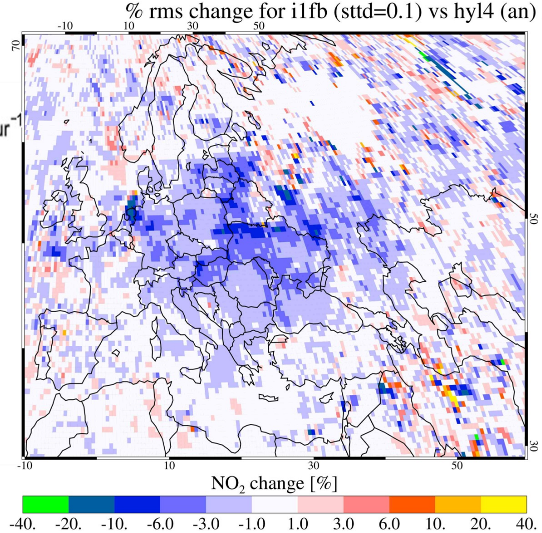
Zhang et al.

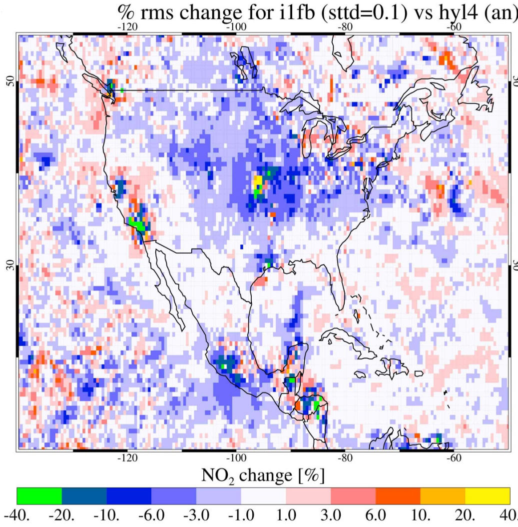


- CH_4 inversion results in agreement with previous studies.

NO_x inversion (04/2020)

inversion vs state only (prior error=10%)

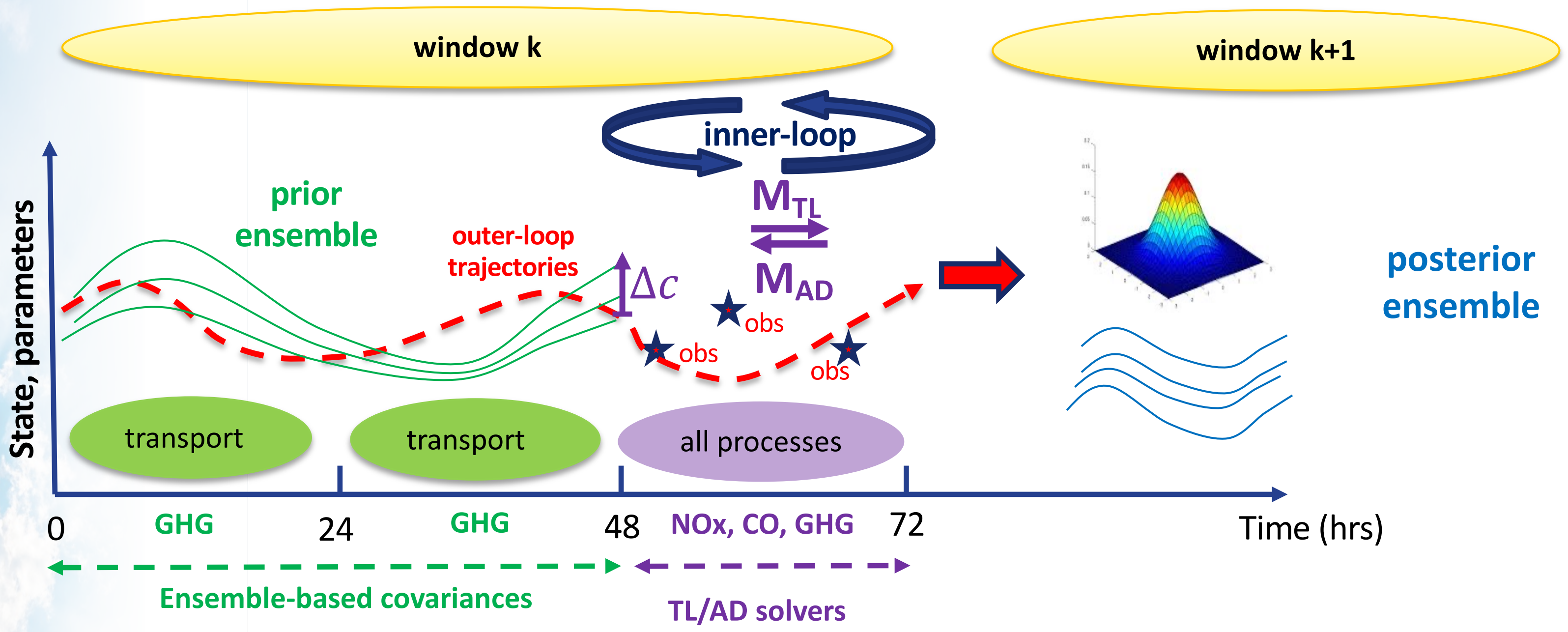




% rmse change

- Evaluation of IFS 24h forecasts against TROPOMI NO_2 columns (PAL product).
- Significant improvements (10-50% rmse reduction).

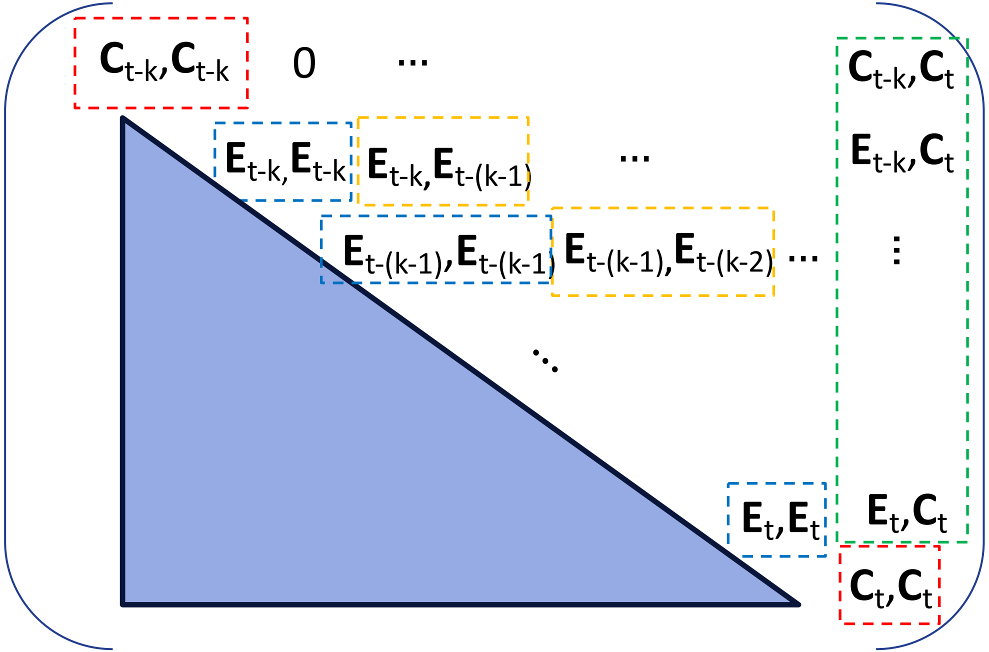
Extended 4D-Var window for GHG inversion



- Tangent-linear/adjoint models used for short-window containing current observations.
- Ensemble-based covariances used for previous days.
- Short-window state increment (Δc) propagated backward to update past emissions.

The **B** Matrix Revisited

subscript=time indice; **E**=emissions; **C**=concentrations



The **B** matrix is reformulated to extend the 4D-Var optimisation to emissions:

$$\delta x = \mathrm{B}\mathrm{H}^{\mathrm{T}}\left(\mathrm{H}\mathrm{B}\mathrm{H}^{\mathrm{T}} + \mathrm{R}\right)^{-1}(y - h(x)) \rightarrow$$ state **δx**

$$\mathrm{B} = \mathrm{L}\mathrm{L}^{\mathrm{T}}$$

inversion of state **B**

$$\delta p' = \mathrm{B}'\mathrm{L}^{-T}\mathrm{L}^{-1}\delta x \rightarrow$$ emissions **δp'**

$$= \mathrm{B}'\mathrm{H}^{\mathrm{T}}\left(\mathrm{H}\mathrm{B}\mathrm{H}^{\mathrm{T}} + \mathrm{R}\right)^{-1}(y - h(x))$$

extended B′

- univariate spatial (wavelet)
- univariate spatial(wavelet) +multivariate cross-species (explicit B)
- transport Jacobian (ensemble)
- temporal correlations (exponential decay)

PROGRAMME OF THE EUROPEAN UNION

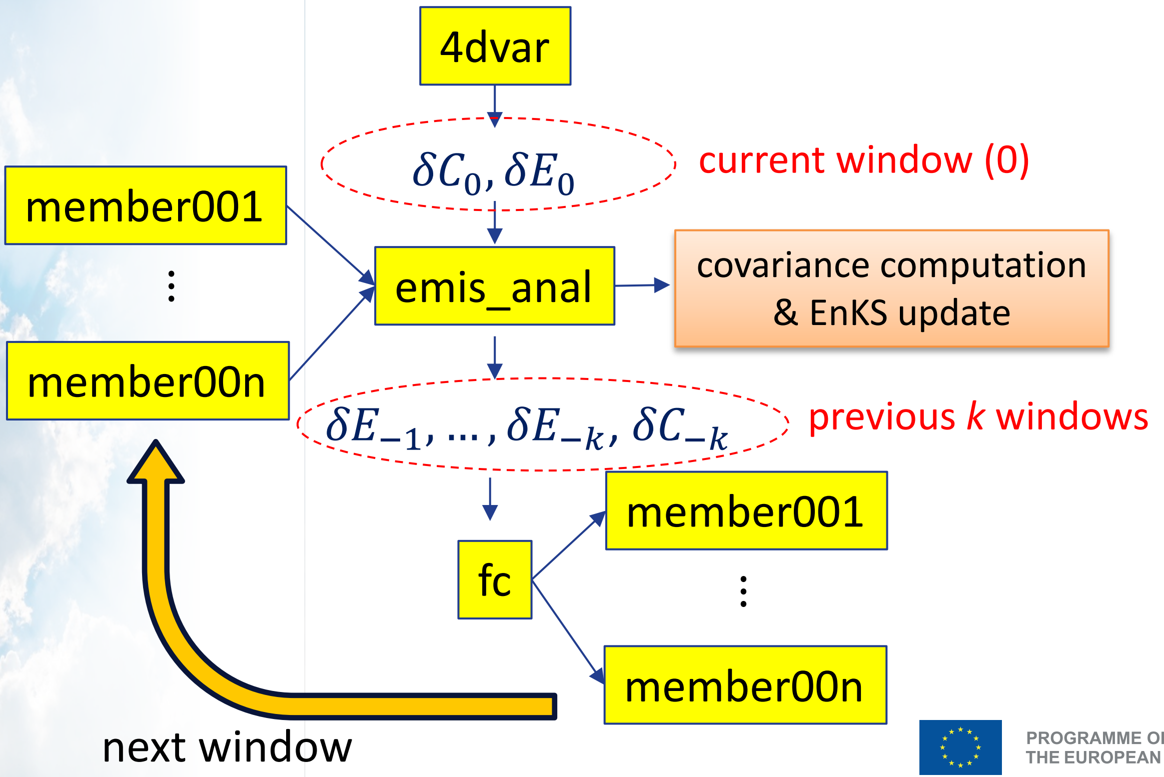
Copernicus Europe's eyes on Earth

IMPLEMENTED BY ECMWF

Atmosphere Monitoring

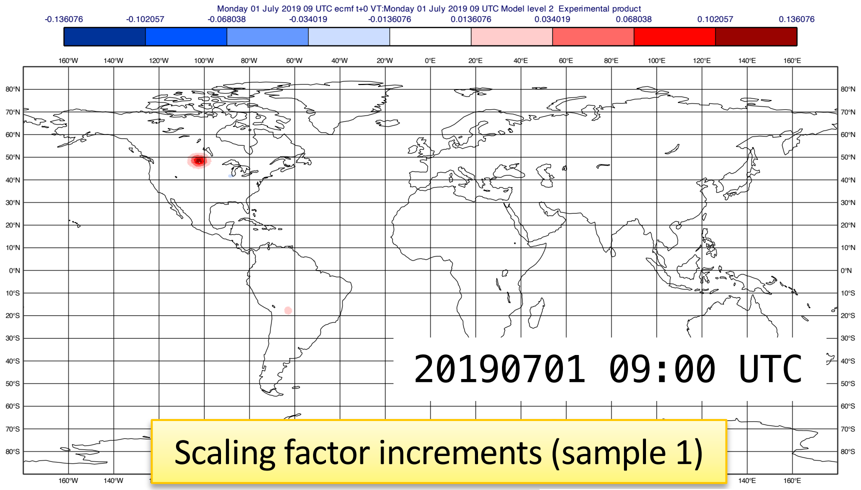
Workflow for long-window inversion

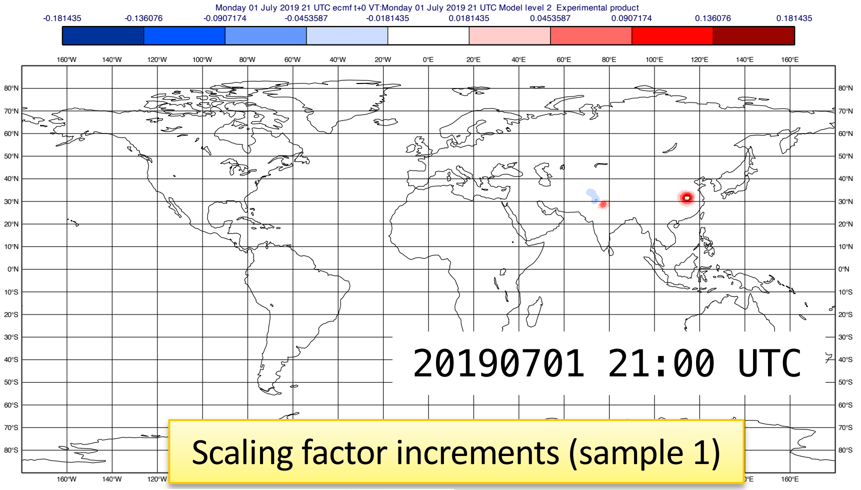
Ensemble of Data Assimilation (EDA)

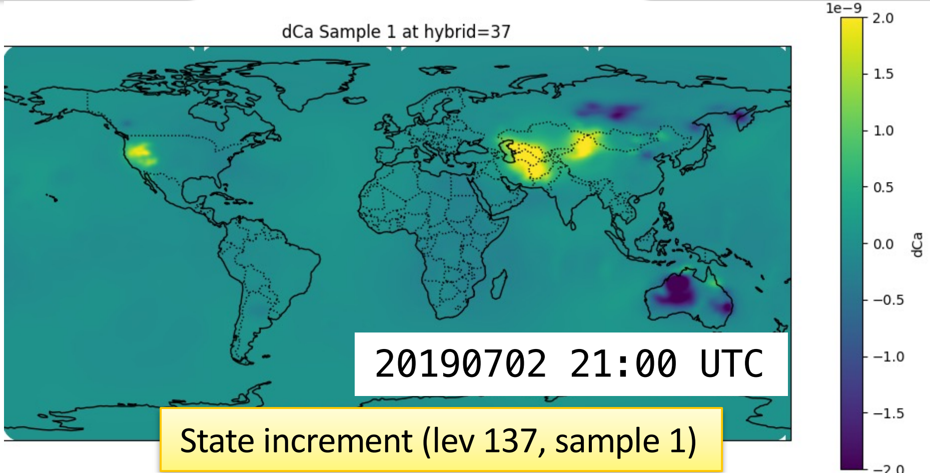


- **Ensemble Kalman Smoother** update of the prior emission ensemble after each 4D-Var minimisation.
- Flow-dependent **B** matrix ensures forward propagation of error covariances

Long-window flux inversion (preliminary results)
CH_4 emissions







…nicus
…'s eyes on Earth

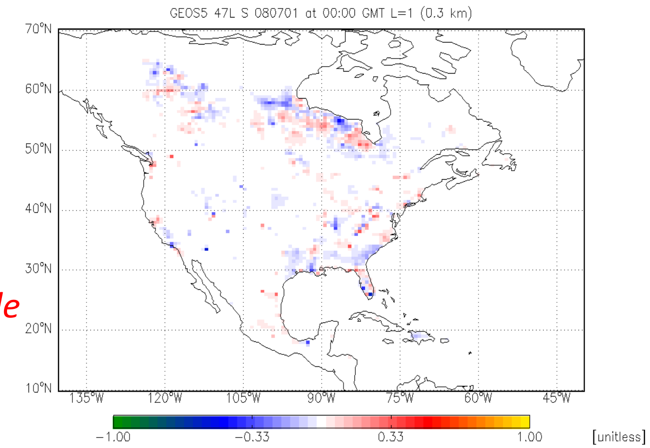
IMPLEMENTED BY ECMWF

Towards affordable posterior ensembles

- **EDA** is expensive to run operationally (prohibitive for the CAMS chemistry configuration)
- **Lanczos-based posterior covariance approximations** can be used for posterior sampling instead of the EDA
- *N* (~50) analyses are replaced by *N* perturbed forward model integrations
- Requires only approximation of eigendecomposition $(\lambda_i, \boldsymbol{v_i})$ of Hessian matrix at the analysis state (x_a):

$$x_a^i = x_b^i + L\left(I - \sum_{j=1}^{k} \lambda_j \left(1+\lambda_j\right)^{-1} v_j^i (v_j^i)^T\right) L H^T R^{-1}(y + \epsilon^i - h(x_b^i))$$

Posterior sample — *posterior covariance matrix* — *observation error sample*



posterior scaling factor sampling for CH_4 inversion (example using the GEOS-Chem model)

- Will be implemented within OOPS

Assimilation of regional/local posterior emission products

Atmosphere Monitoring

Objectives:

- Integrate high-resolution transport information into the global IFS inversion system.
- Two-way flow of information: global posterior atmospheric concentrations can in turn be used as boundary conditions for regional/local inversion systems.

Constraints:

- Method should be computational efficient & non-intrusive
- Inputs from external inversion products should be harmonized

Approach:

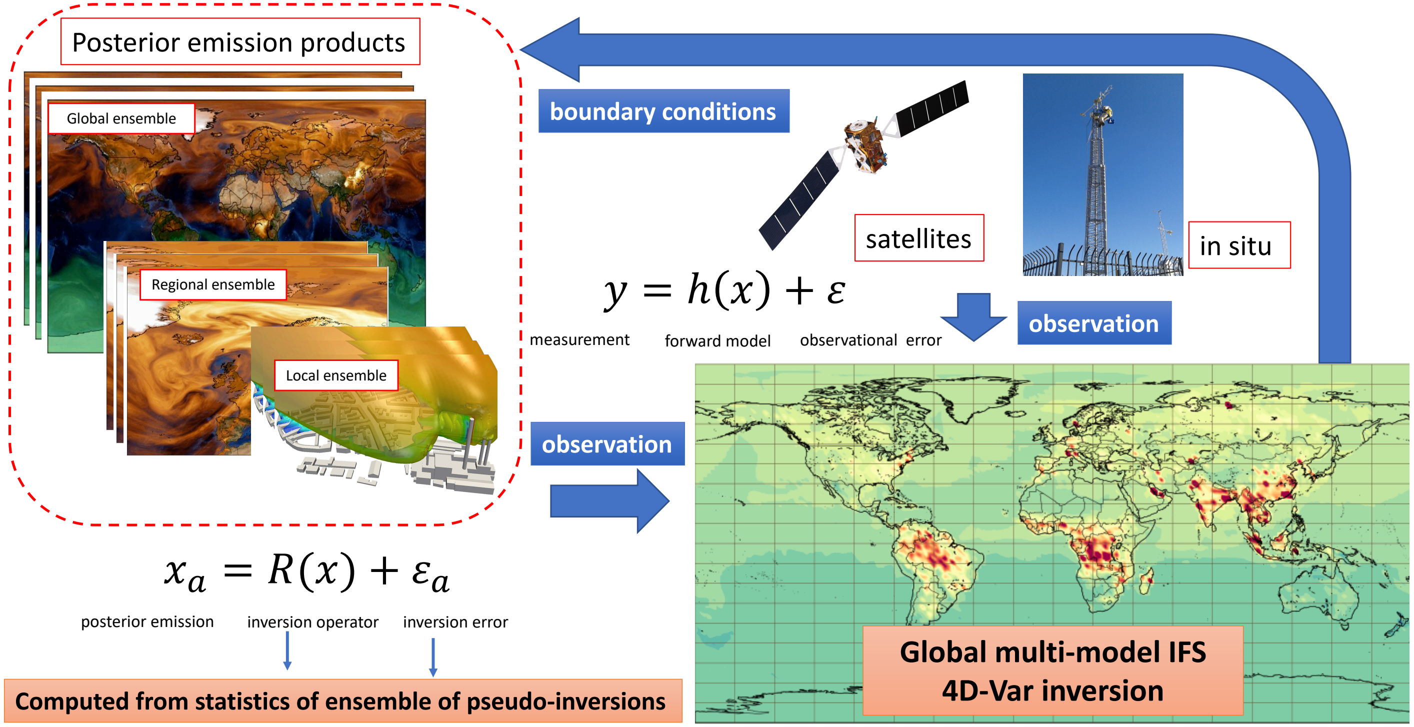
- Use ensemble of perturbed inversions to approximate inversion operator and associated retrieval errors→by-product of EnKF methods; obtained from EDA for 4D-Var algorithms

General Approach

Two-step approach:

1. Assimilate atmospheric observations using IFS 4D-Var system.
2. Use IFS 4D posterior emission statistics to assimilate external posterior emission products.

Multi-model multi-scale approach



Atmosphere Monitoring

CONCLUSION

- ECMWF is implementing a global emission monitoring system in the IFS in preparation for a new Copernicus service.
- Important methodological and algorithmical developments to extend the current capability of the 4D-Var system to enable long-window (several weeks) emission inversions.
- Current hybrid ensemble-variational system implements a Kalman smoother based on the EDA.
- Future developments will focus on:
  - Replacing the EDA by cheaper approaches based on posterior covariance approximations obtained as by-product of the 4D-Var (Lanczos method).
  - Efficient 4D localisation methods to mitigate the sampling noise.