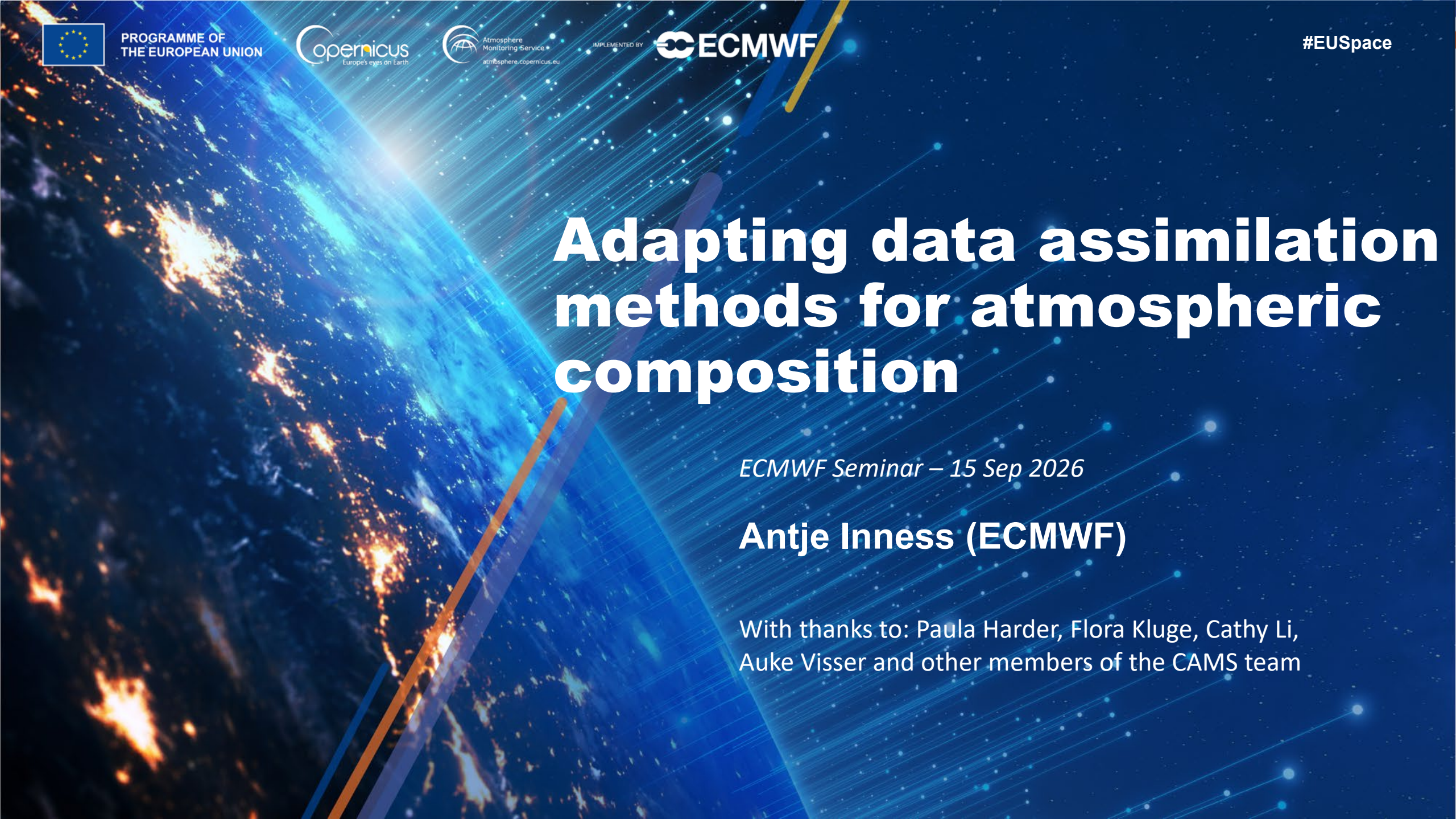



# Adapting data assimilation methods for atmospheric composition

*ECMWF Seminar – 15 Sep 2026*

**Antje Inness (ECMWF)**

With thanks to: Paula Harder, Flora Kluge, Cathy Li, Auke Visser and other members of the CAMS team

# Outline

1. CAMS overview
2. Machine learning for atmospheric composition
3. Emission inversion
4. Impact of Atmospheric composition on NWP forecast scores
5. Summary

# 1. Copernicus Atmosphere Monitoring Service (CAMS)

Within the 6 Copernicus Services, the CAMS mandate is to deliver consistent and quality-controlled information related to **the chemical composition of the atmosphere**, its drivers (**air pollutants and GHG**) and its impact on air quality and health, solar energy, climate forcing, in Europe and everywhere in the world.



Air quality


Policy tools


Solar energy


Ozone layer and UV radiation


Emissions and surface Fluxes


Climate forcing

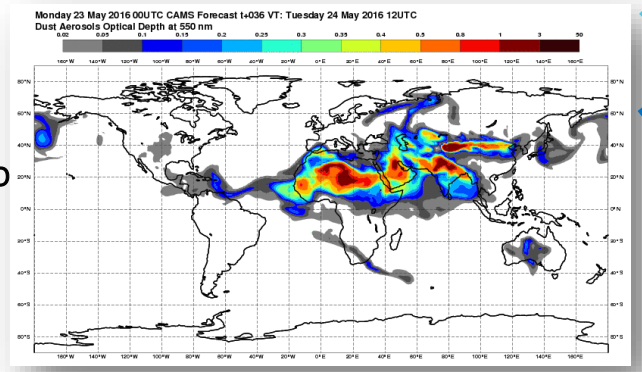
# CAMS integrated framework

Global CAMS Emissions

GFAS wildfire emissions

European CAMS Emissions

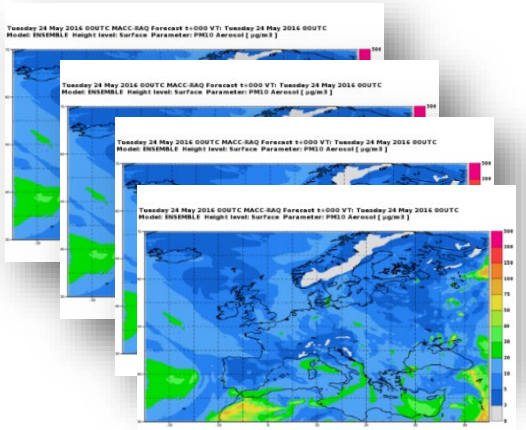
ECMWF IFS-compo global, 40x40 km



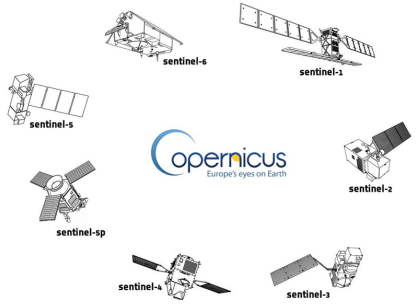

5-day forecasts of atmospheric composition

Boundary Conditions

11 regional CTMs over Europe, 10x10 km

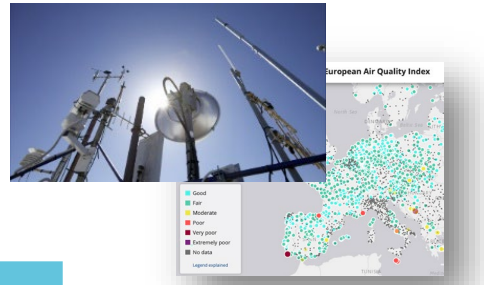


Data assimilation




Satellite Observations (AOD, $O_3$, CO, $NO_2$, $SO_2$,$CH_4$,$CO_2$)

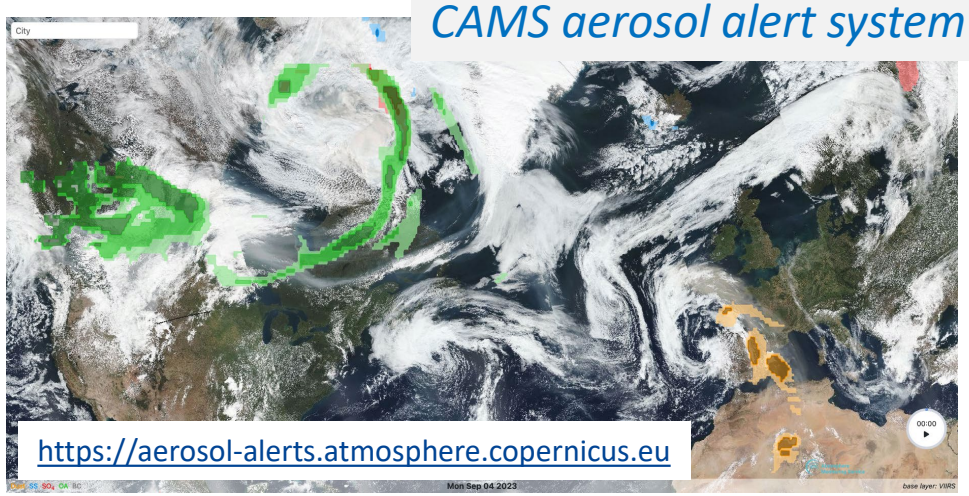
Data assimilation




Surface AQ observations

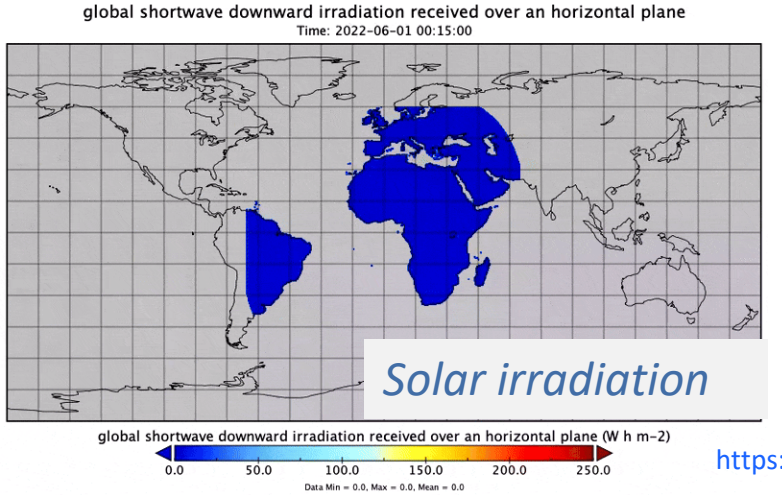
# Examples of CAMS services

## CAMS aerosol alert system




https://aerosol-alerts.atmosphere.copernicus.eu

## Solar irradiation

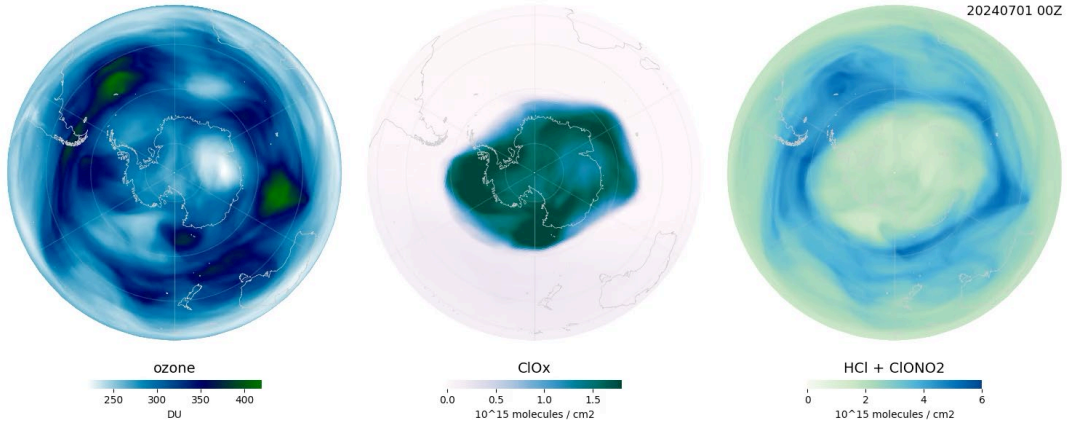



## Wildfires



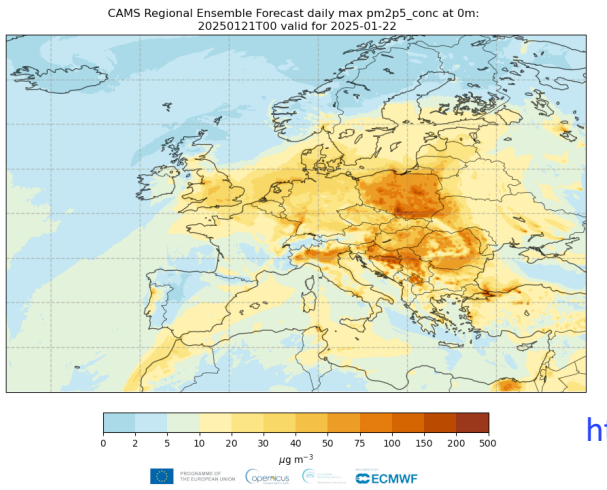

https://apps.atmosphere.copernicus.eu/fire-emissions-watch/

## Ozone hole

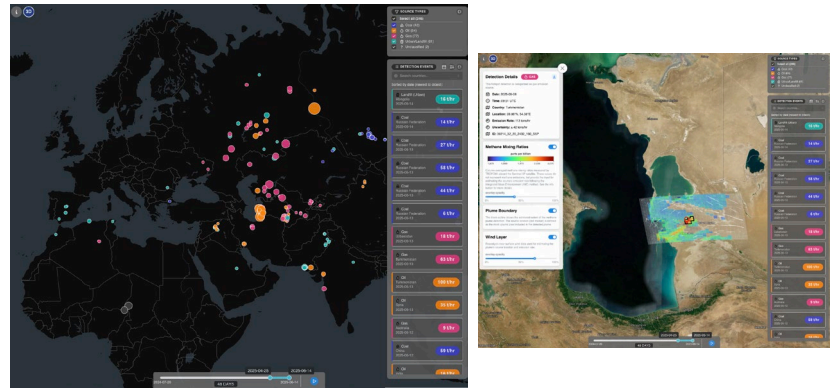



https://atmosphere.copernicus.eu/monitoring-ozone-layer

## European air quality




## Methane Hot-spots



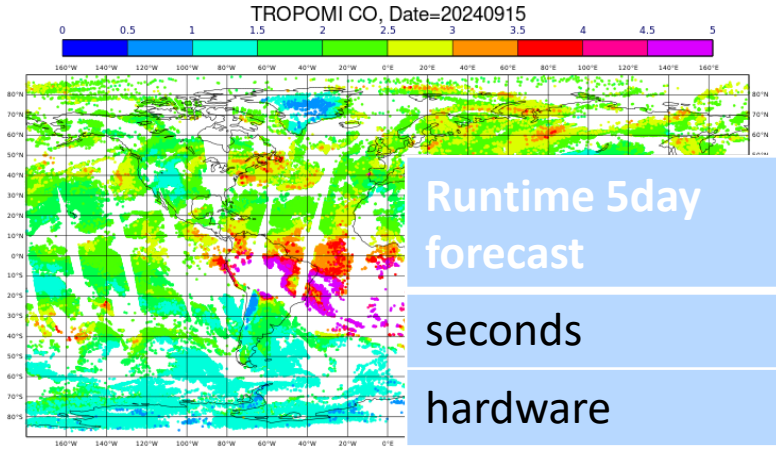
https://apps.atmosphere.copernicus.eu/methane-explorer

PROGRAMME OF THE EUROPEAN UNION

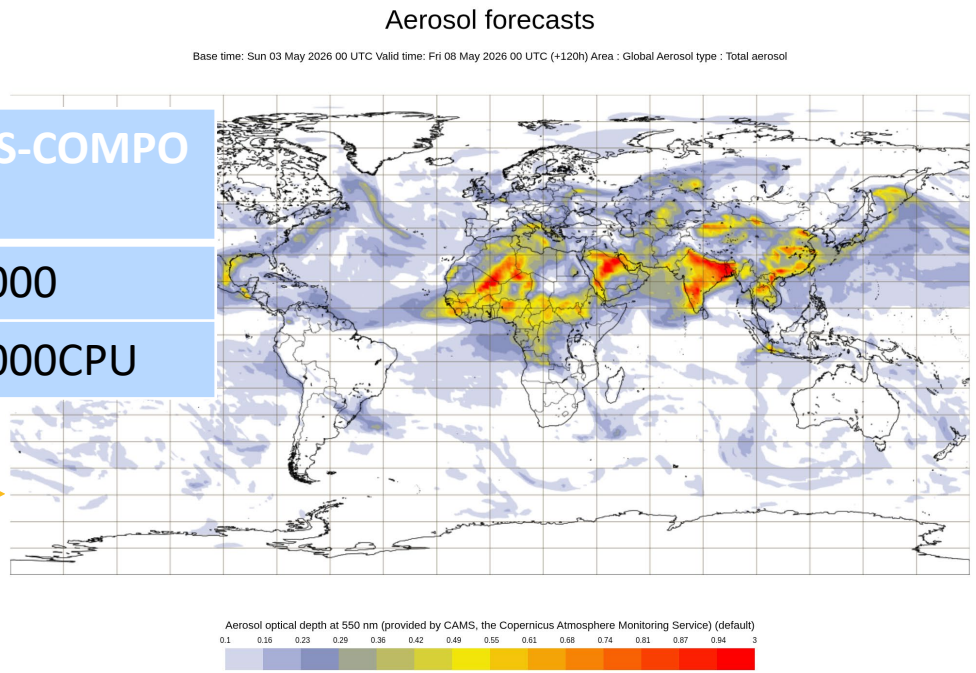
Thanks to Paula Harder for her contribution

# 2. Machine learning for Atmospheric Compostion

3-hourly 5-day AC forecast with AI at a fraction of the computational cost




| Runtime 5day forecast | AIFS-COMPO | IFS-COMPO |
|---|---|---|
| seconds | 50 | 1000 |
| hardware | 1GPU | 8000CPU |




**Satellite observations**

**AIFS-COMPO**

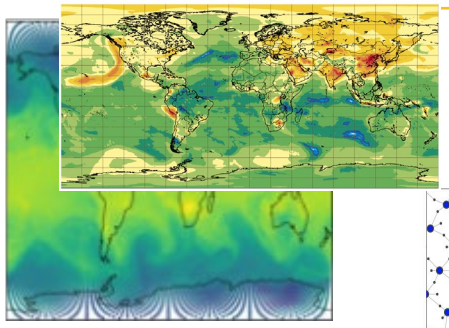
**Forecast** → **Analysis** → **Forecast** → **Analysis** → • • •

PROGRAMME OF THE EUROPEAN UNION

Credit: Paula Harder

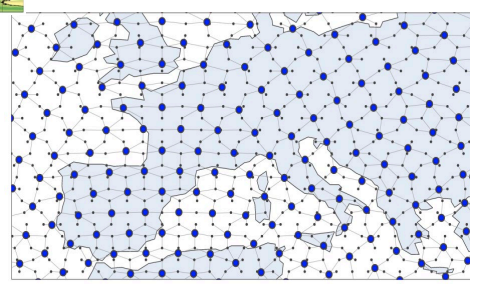
# From AIFS to AIFS-COMPO

Input
Atmospheric state+AC variables:
X(t), X(t-3h)

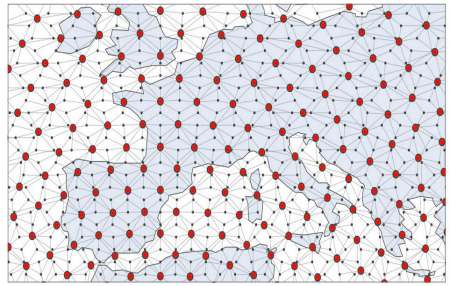


AIFS-COMPO

encoder → processor → decoder

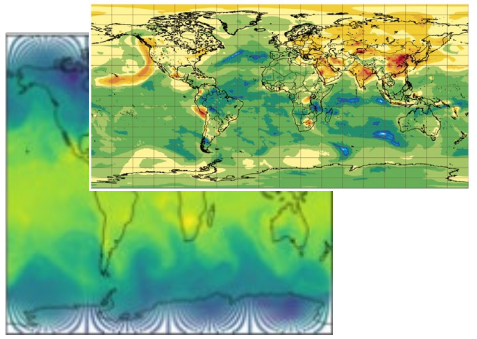


Graph attention



Graph attention

Output
Prediction of
Atmospheric state+AC variables :
X(t+3h)



**Variables**:

- Atmospheric composition variables
  - AODs, PMs, Reactive gases
  - Mixing ratio at pressure levels
- Upper-air variables at 13 pressure levels (t,v,u,w,z)
- Surface variables (temperature, winds, pressure, radiation)
- Static geographical features, location and time information as input forcing

**Training Scheme:**

1. Train on EAC4 (yrs 2003-2023)
2. Fine-tune on operational analysis/forecast (yrs 2018-2023)
3. Rollout finetune on lead times up to 72h with operational data

Implemented in anemoi

Credit: Paula Harder

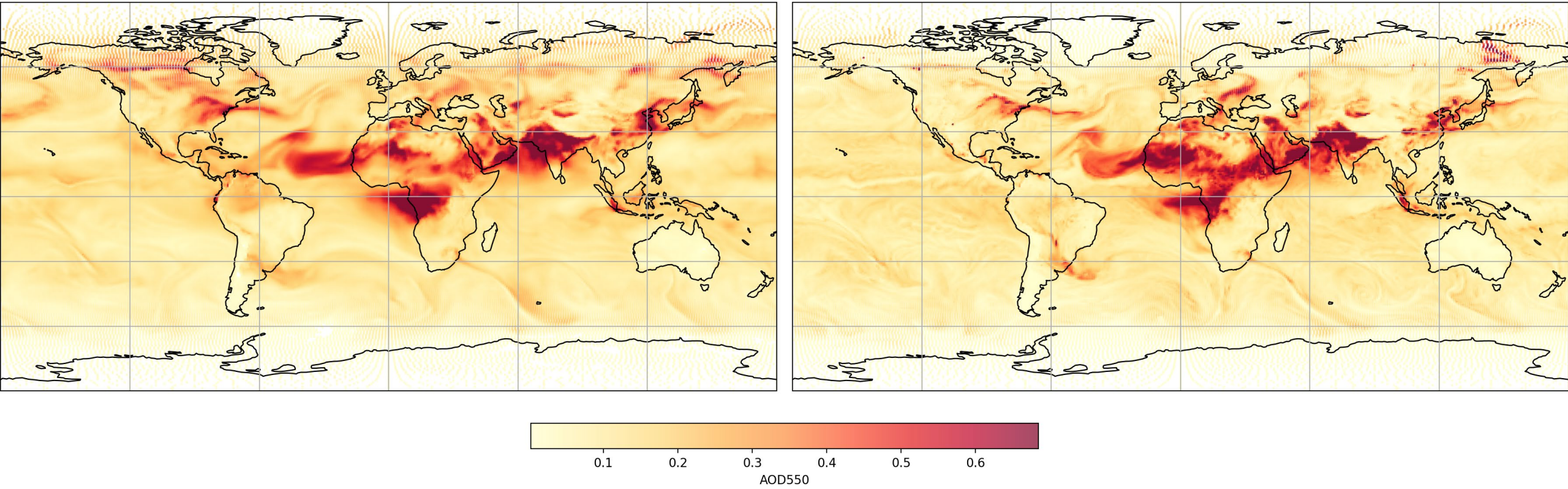
# Results: Example AOD Day 3 forecast

## AIFS-COMPO

AIFS-COMPO AOD550
Init: 2024-06-20 00 UTC, Lead: +72 h

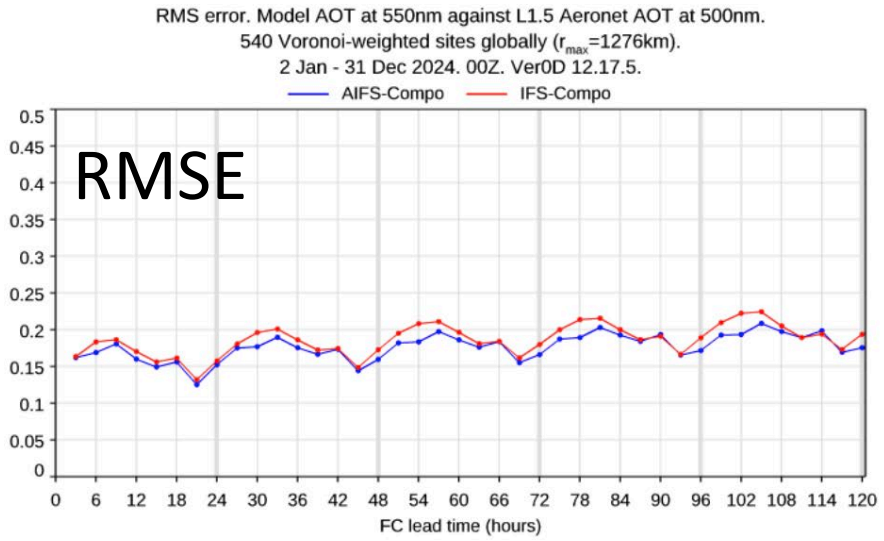
## IFS-COMPO

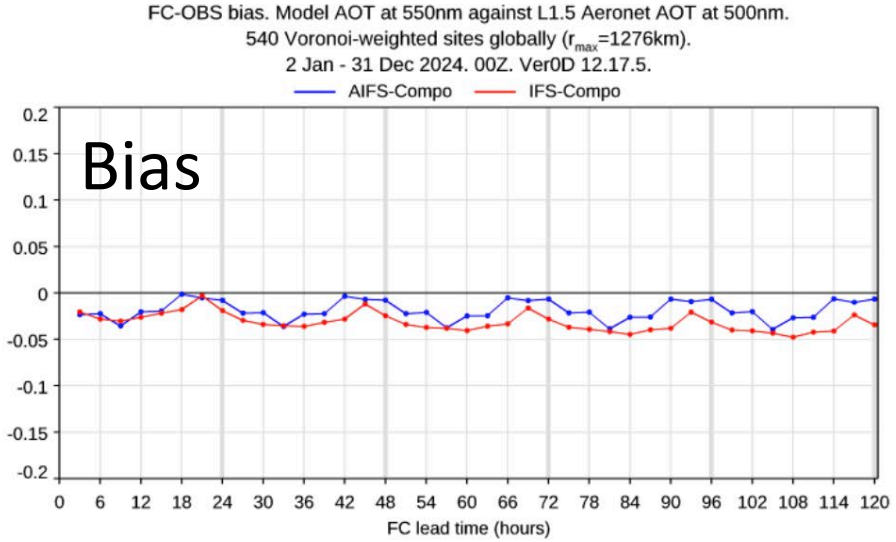
IFS-COMPO AOD550
Init: 2024-06-20 00 UTC, Lead: +72 h

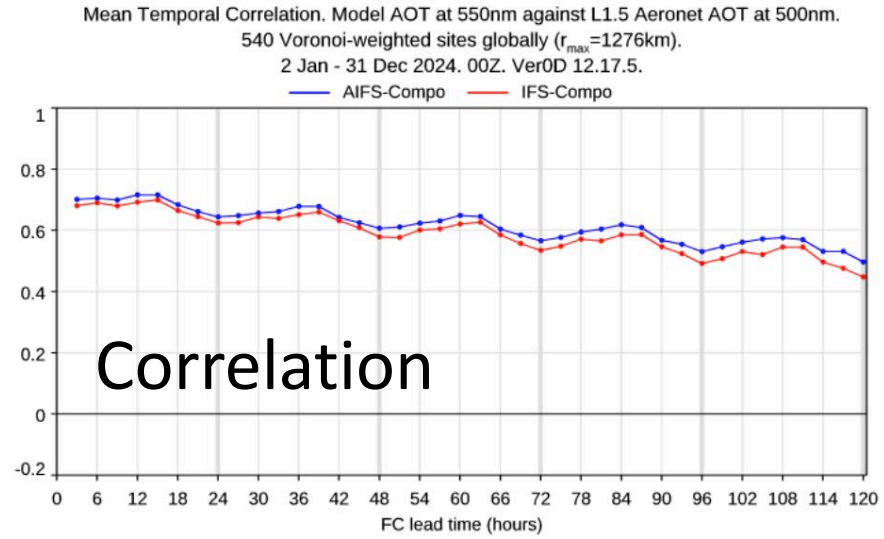


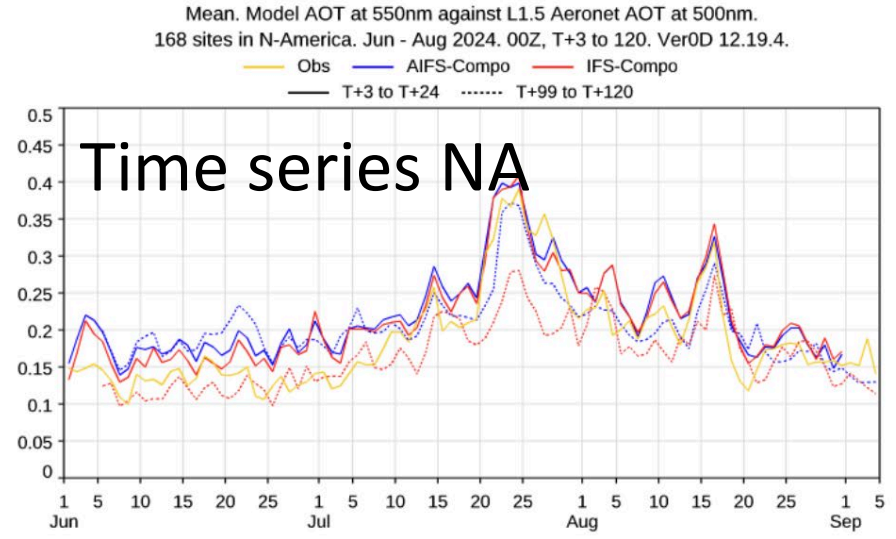
Credit: Paula Harder

# Results: Global AOD metrics



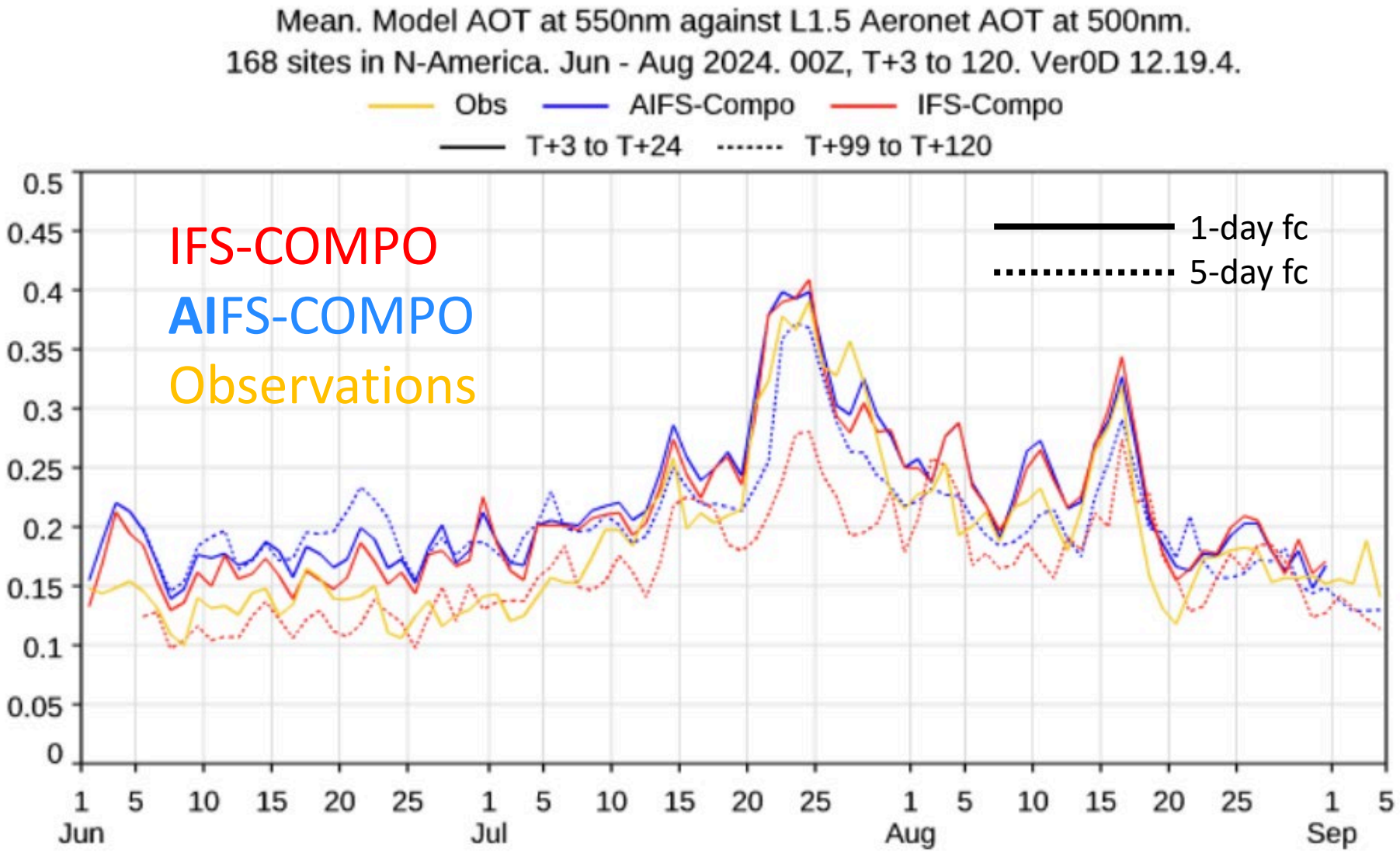

IFS-COMPO

AIFS-COMPO

AIFS-COMPO improves RMSE, bias, and correlation compared to IFS-COMPO wrt to observations

Credit: Paula Harder

# Results: AOD North America Wildfire Season 2024




AIFS-COMPO is able to predict extremes, even 5 days ahead

Credit: Paula Harder

# Results: AOD – Different training stages

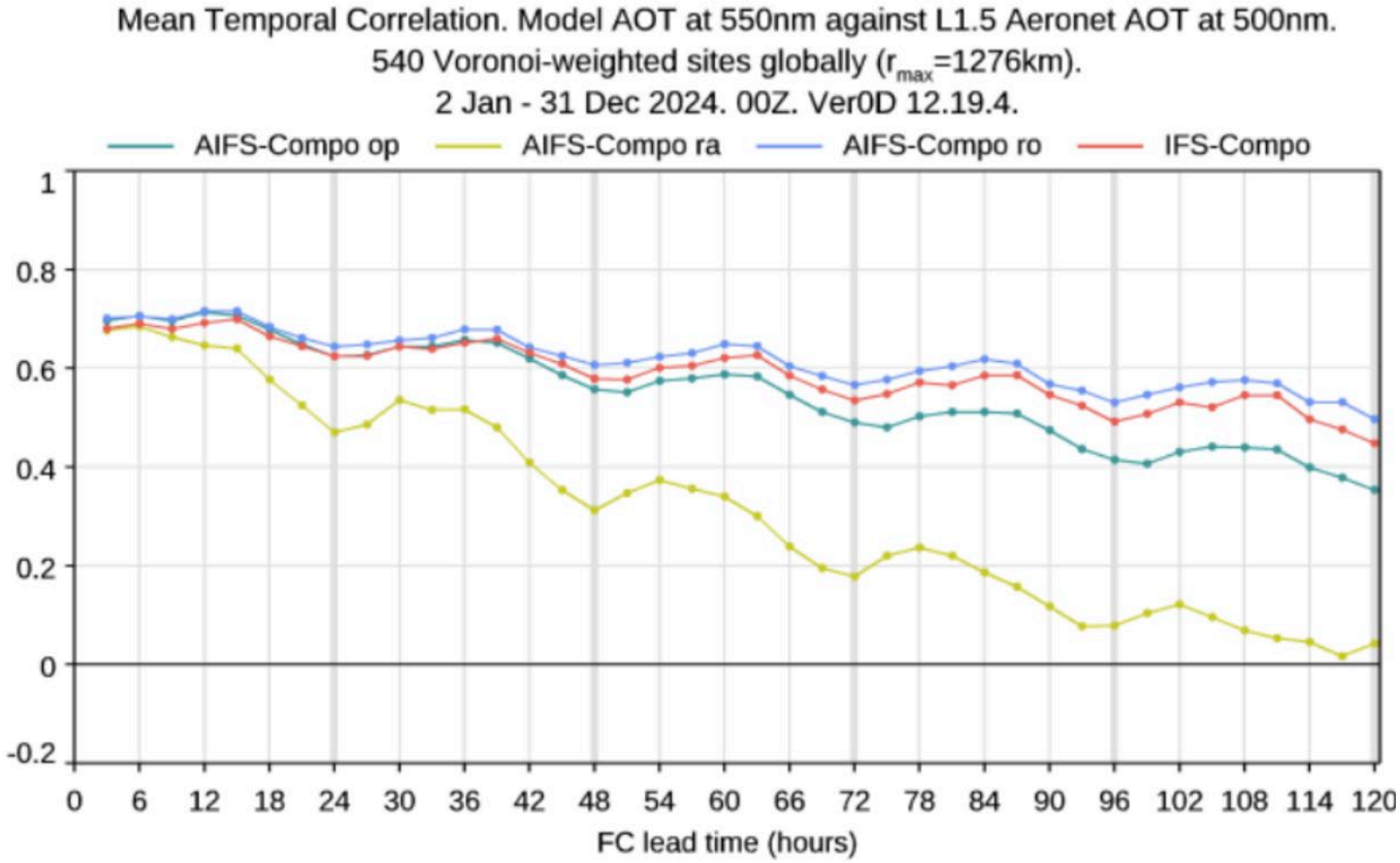
**Training Scheme:**
1. Train on EAC4
2. Fine-tune on operational analysis/forecast
3. Rollout finetune on lead times up to 72h

IFS-COMPO

AIFS-COMPO final

AIFS-COMPO oper

AIFS-COMPO EAC4 only




For AIFS-COMPO finetuining on operational data improves the performance significantly

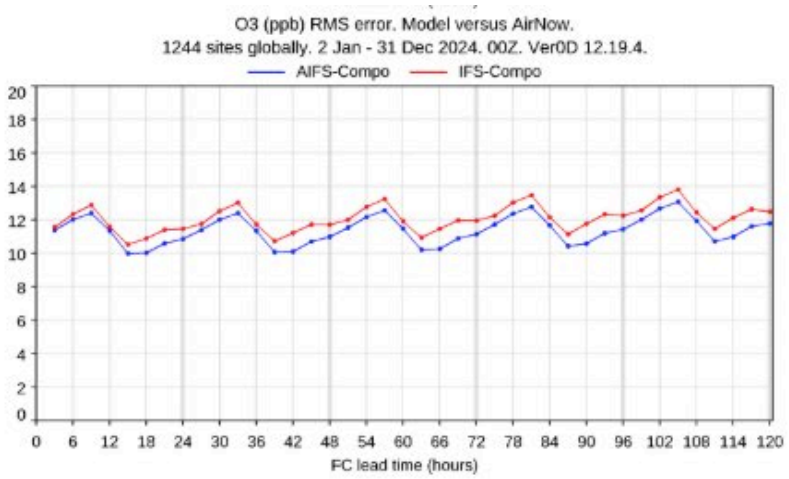
Credit: Paula Harder

# Results: Surface Ozone RMSE

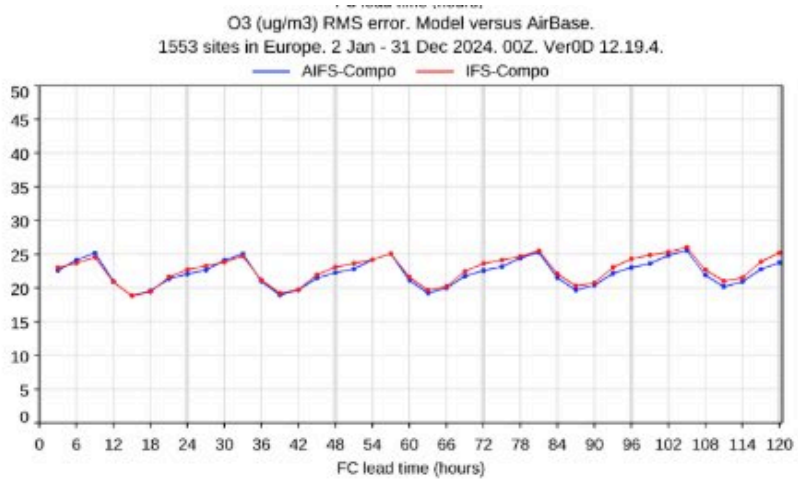
IFS-Compo

AIFS-Compo

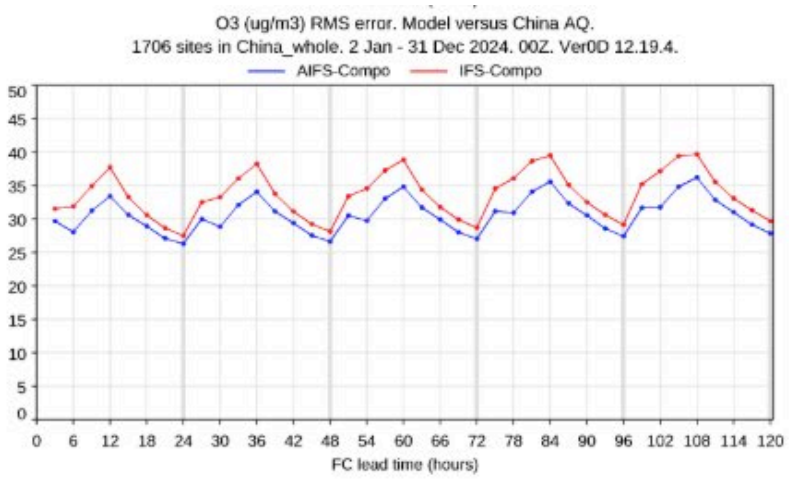
North America




Europe



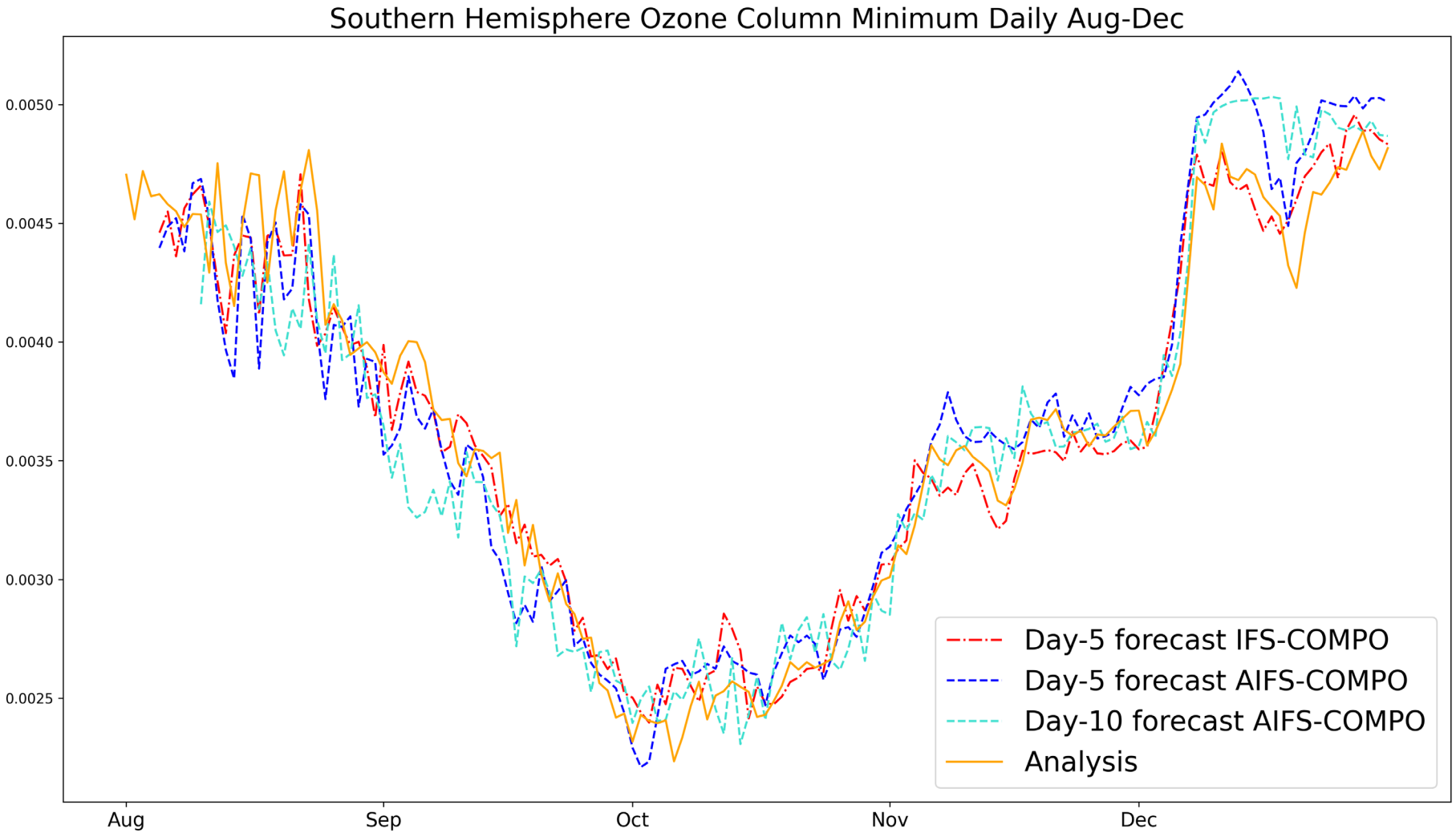

China



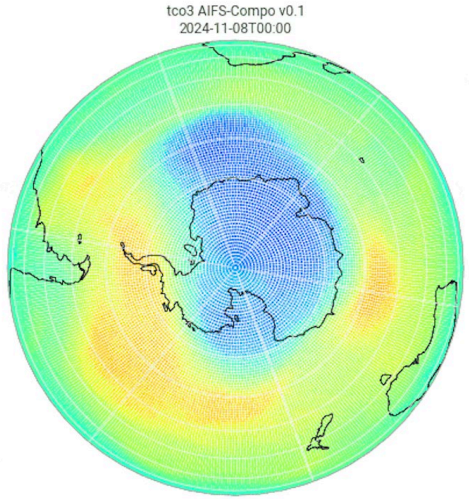

- Depending on the area AIFS-COMPO either matches or outperforms IFS-COMPO
- Largest improvement in China

Credit: Paula Harder

# Results: Ozone hole prediction 2024

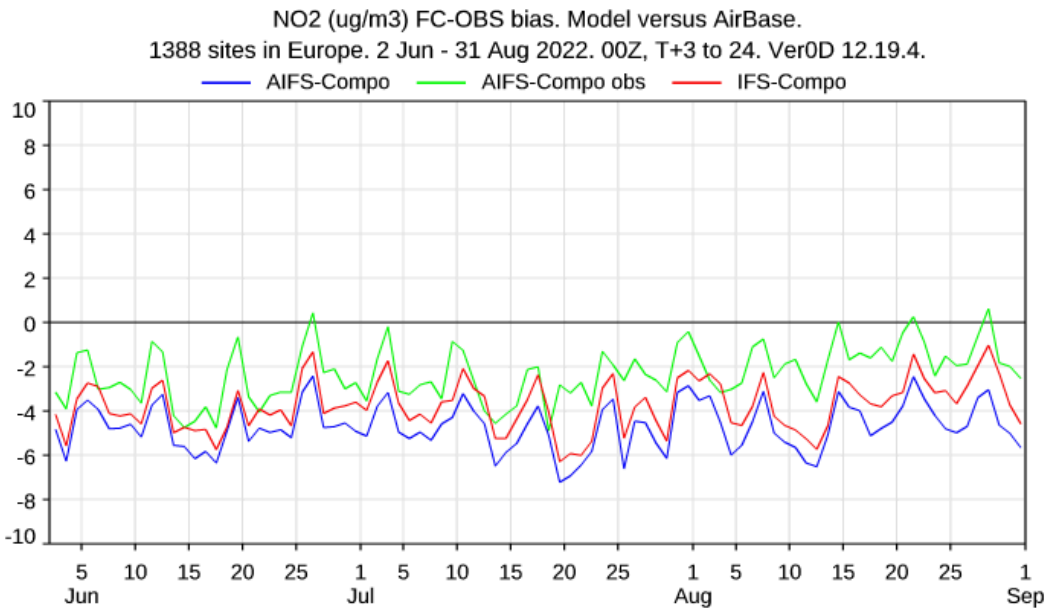



AIFS-Compo can predict the ozone hole and even extend to a 10 day forecast reliably

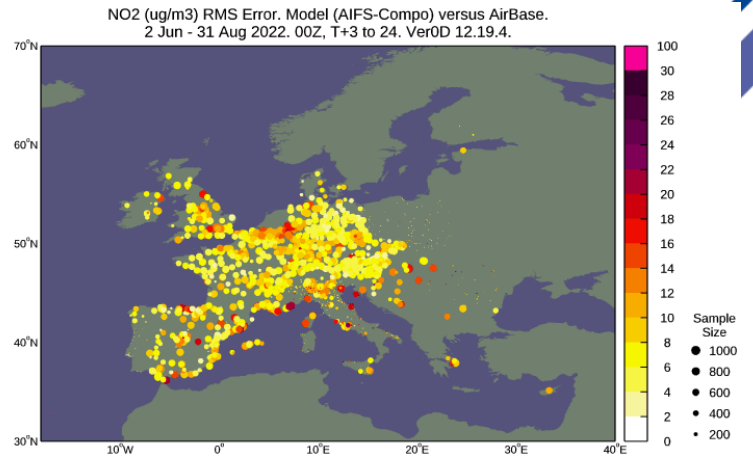


# Learning from Observations

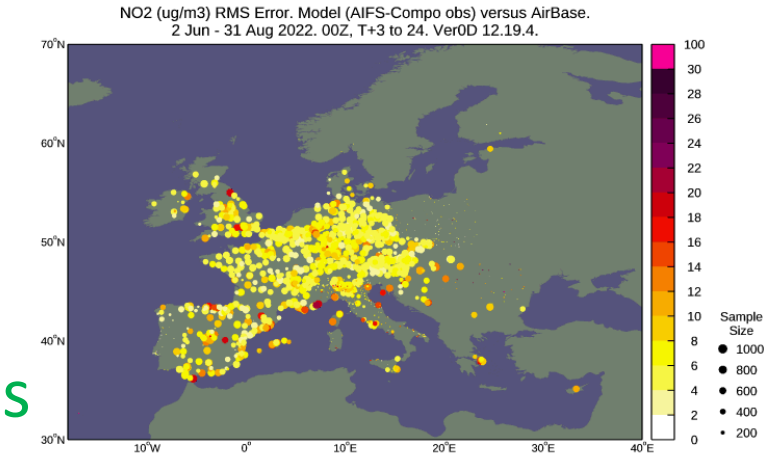
- First experiment with gridded NO2 observation over Europe
- Including observations for a final rollout finetuning
- RMSE and bias of AIFS-COMPO+obs reduced compared to both IFS-COMPO and AIFS-COMPO



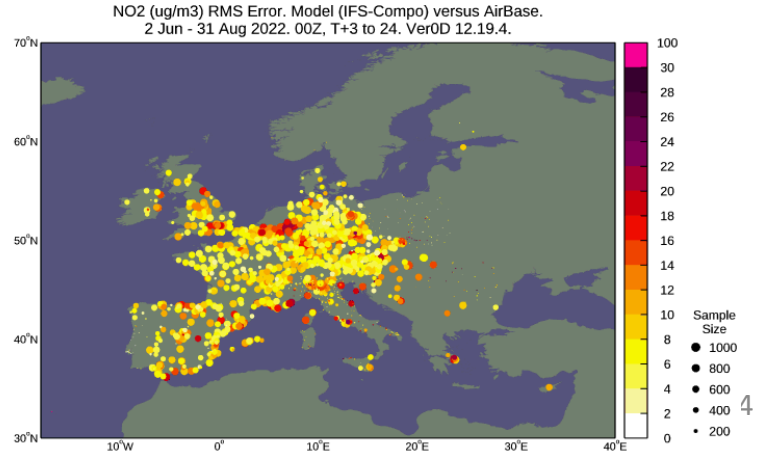

AIFS-COMPO




AIFS-COMPO+obs




IFS-COMPO

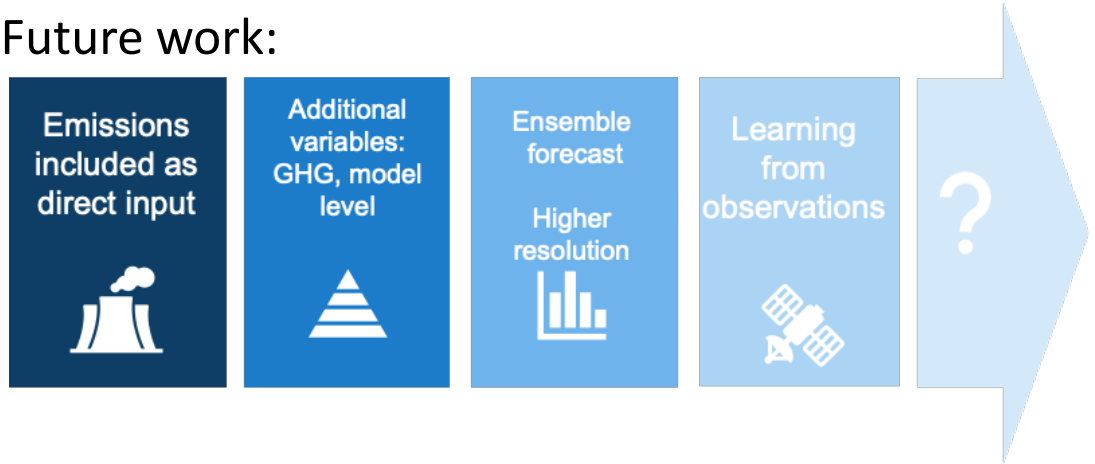



Credit: Paula Harder

# Summary AIFS

- AIFS-COMPO provides a realistic integrated NWP+AC forecast for 5 days leadtimes at a fraction of the computational costs
- AIFS-COMPO matches or outperforms IFS-COMPO when verified against observation for AOD, PM and surface concentrations of ozone, NO2, SO2 and CO
- Including three different training stages, in particular operational finetuning increases the skill of the model significantly
- The accuracy and computational efficiency of AIFS-COMPO can be used to e.g extend the ozone hole prediction to longer forecast horizons of e.g 10 days

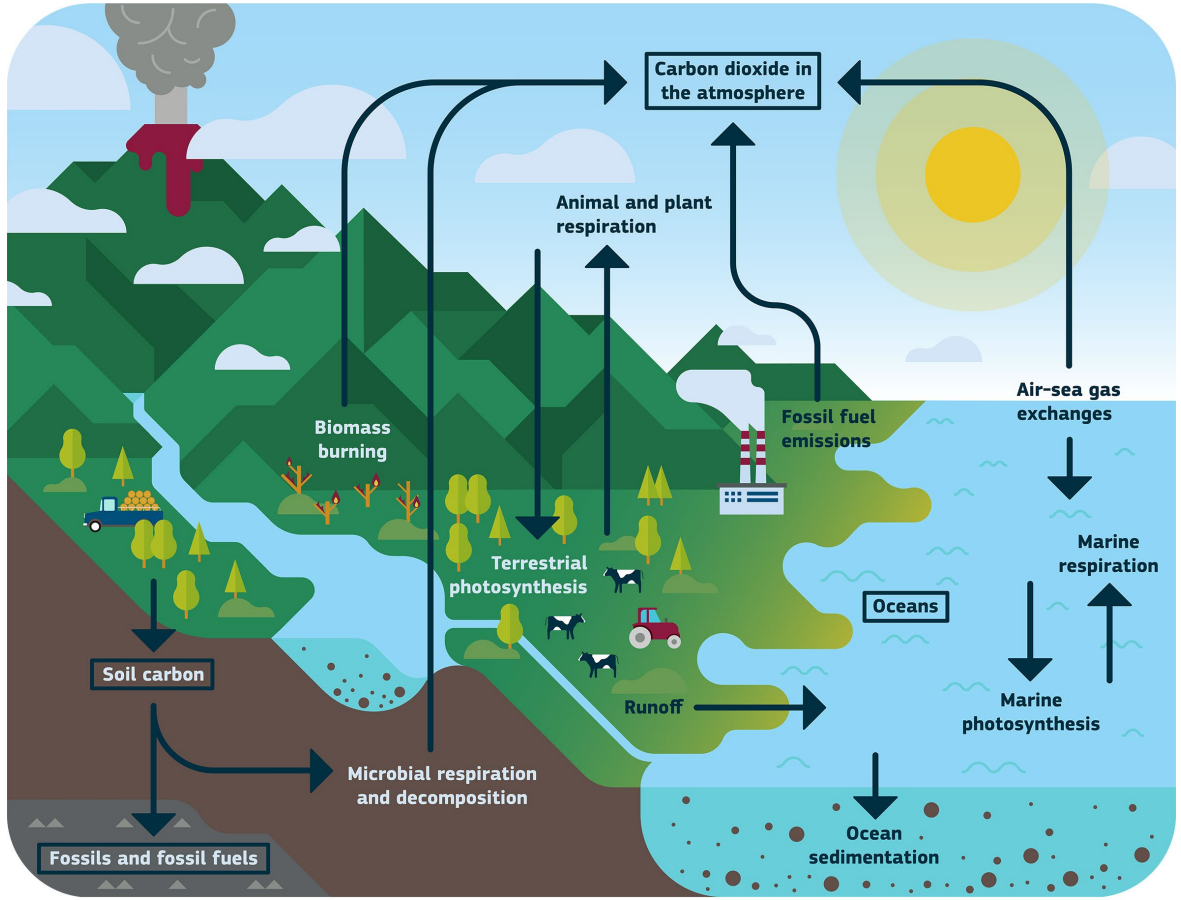
Future work:




Preprint Harder et al. (2026):

# 3. Emission inversion

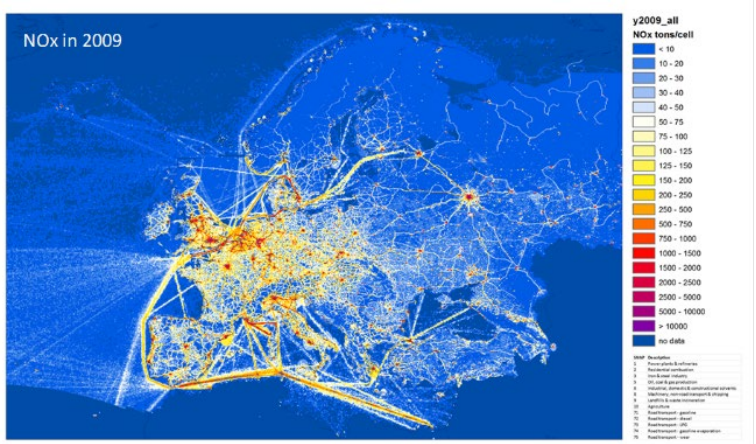
Thanks to Auke Visser and Flora Kluge for their contributions



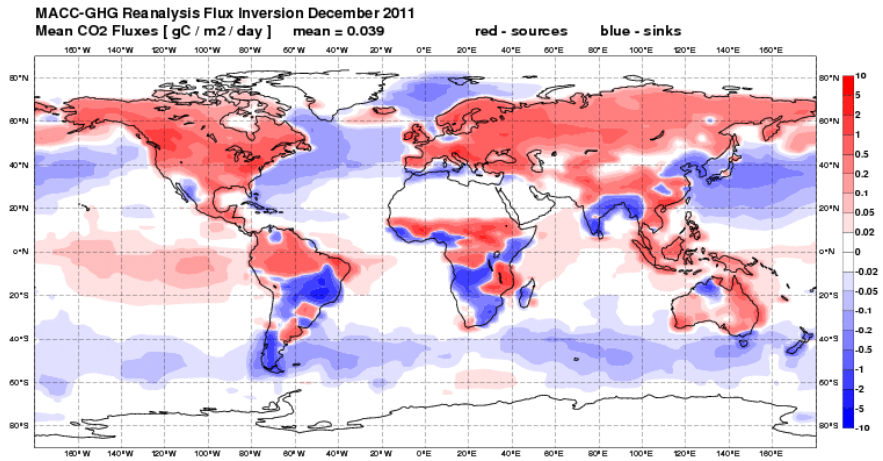

- NWP 4D-Var is mostly defined as an initial value problem. Only initial conditions are changed and model error is relatively small.
- AC modelling depends on initial state and surface fluxes.
- Large part of chemical system not sensitive to initial conditions because of chemical equilibrium, but dependent on other parameters (e.g. emissions, deposition, reaction rates, …) which all might have errors.

# Examples of emissions

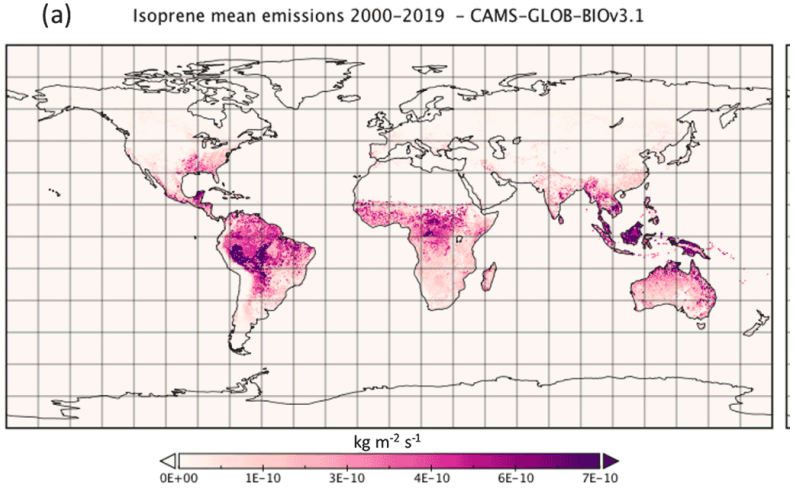
TNO European anthropogenic NOx emissions



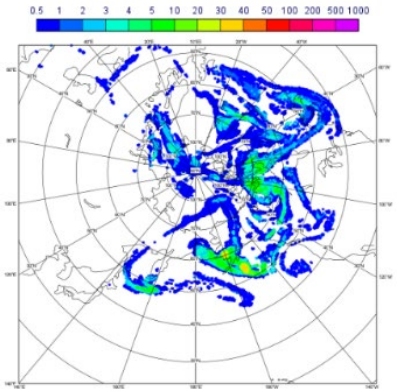

$CO_2$ fluxes




CAMS_GLOB_BIO isoprene emissions

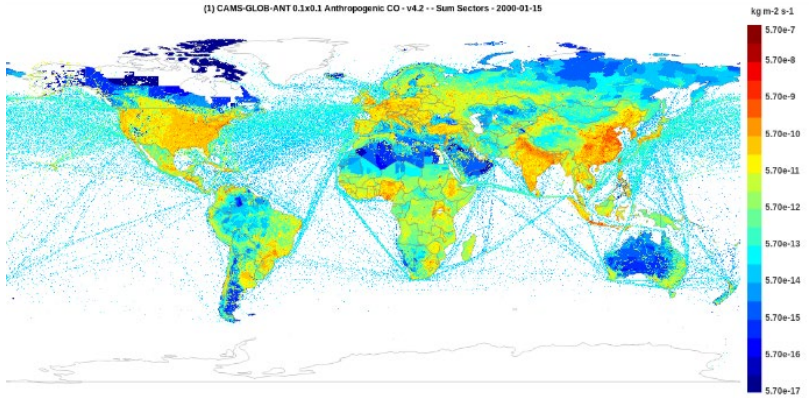


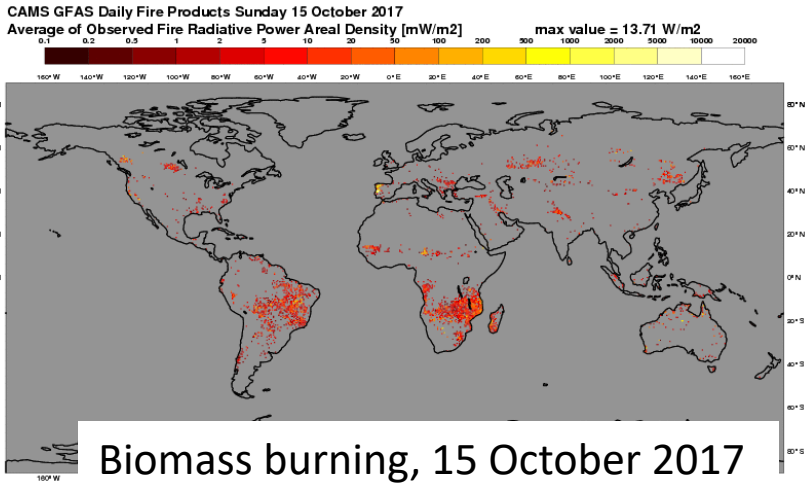

Volcanic $SO_2$





CAMS_GLOB anthropogenic CO emissions



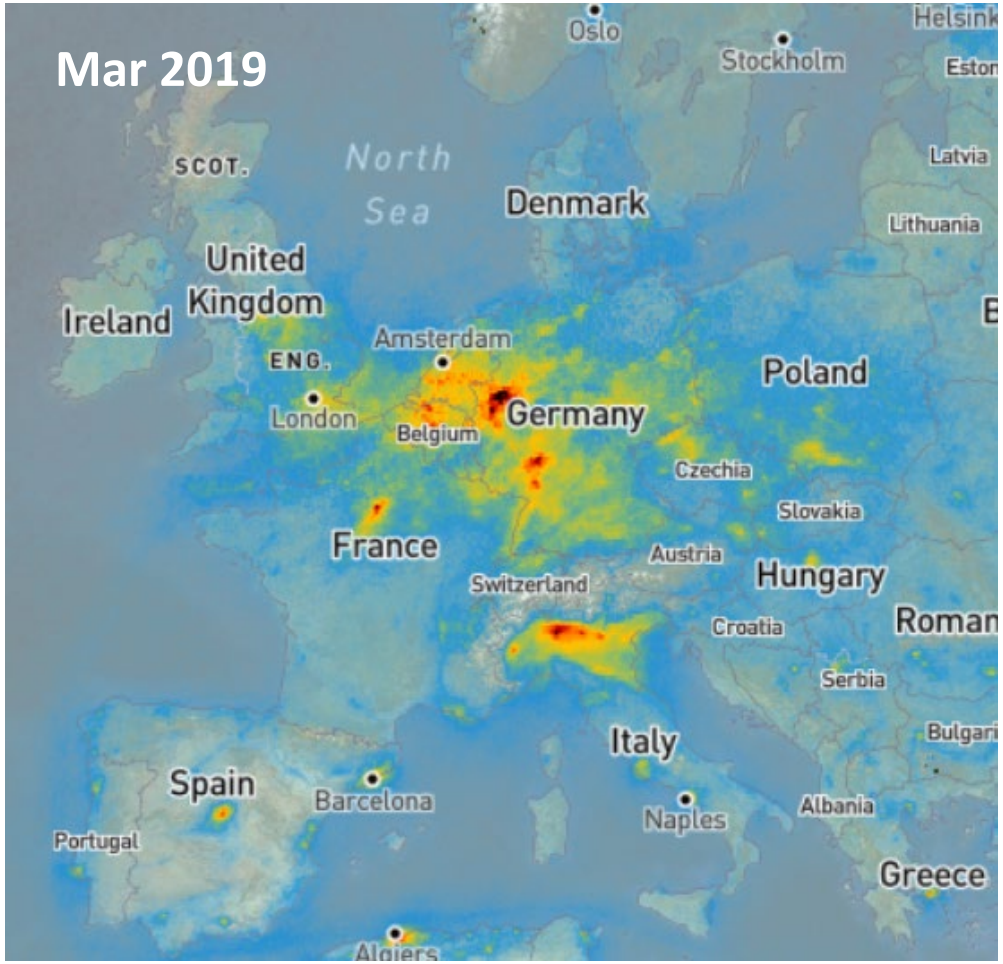



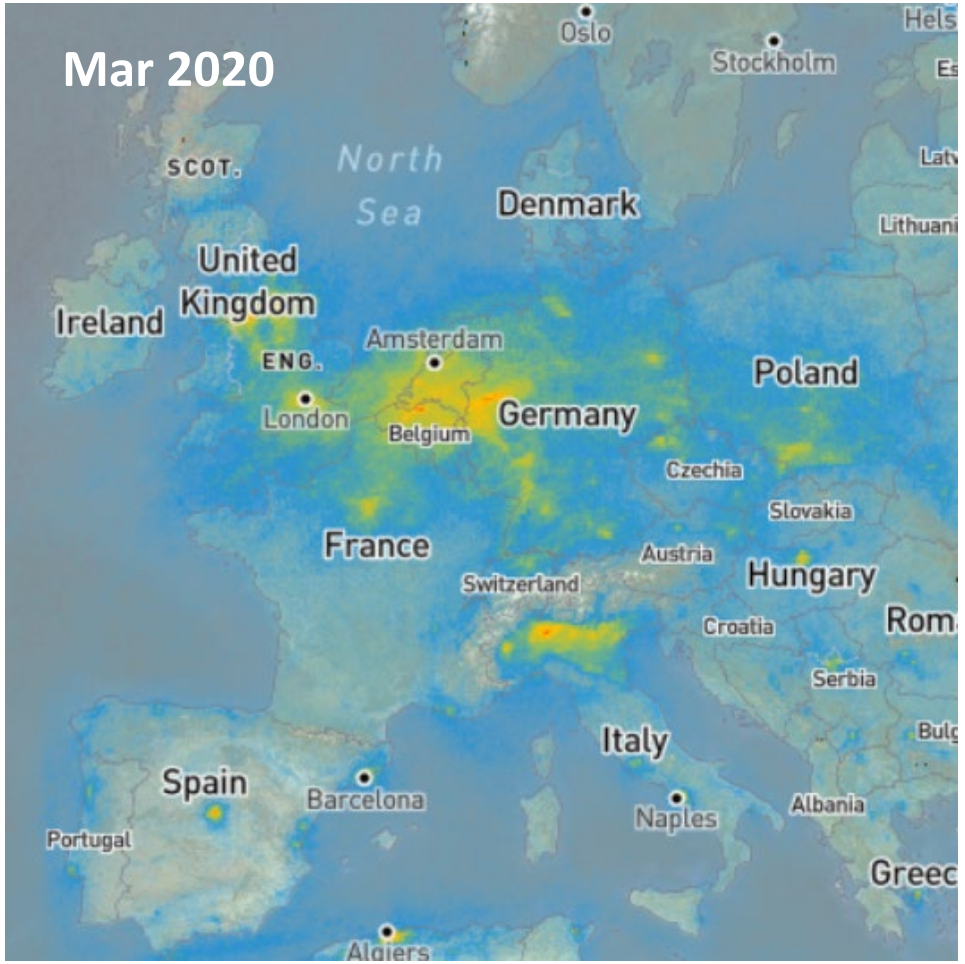

Biomass burning, 15 October 2017

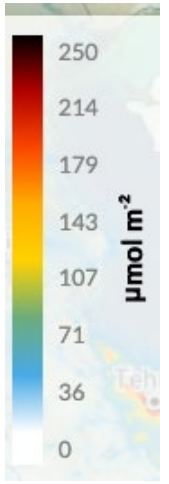
# Emissions

- Emissions are one of the major uncertainties in composition modeling (can not be measured directly)
- The compilation of emissions inventories is a labour-intensive task based on a wide variety of socio-economic and land use data
- Some emissions can be “modeled” based on wind (dust and sea salt aerosol) or temperature (biogenic emissions)
- Some emissions can be observed indirectly from satellites instruments (Fire radiative power, burnt area, volcanic plumes)
- Trends are applied to inventories from previous years to produce future emission datasets
- “Inverse” methods can be used to correct prior emission estimates using observations of concentrations and models

# TROPOMI NO2 columns March 2019 & 2020




https://viewer.terrascope.be

# 4D-Var cost function

$$J(x) = \underbrace{(x - x_b)^T B^{-1}(x - x_b)}_{J_b \text{ Background term}} + \underbrace{\sum_{i=0}^{n}(y_i - H_i[x_i])^T R_i^{-1}(y_i - H_i[x_i])}_{J_o \text{ Observation term}}$$

Cost function

x: control vector
$x_b$: model background (short forecast)
B: Background error covariance matrix
y: Observations
H[x]: Model equivalent of observations
R: Observation error covariance matrix

- **Strong constraint 4D-Var** assumes perfect model over assimilation period
- Weak constraint 4D-Var includes a model error term

**Analysis**: x that minimizes cost function

# Joint State/Emissions 4D-var inversion system

$J_b$: background constraint for x

$J_p$: constraint for emission scaling factors

$$J(x,p) = (x - x_b)^T B^{-1}(x - x_b) + (p - p_b)^T B_p{}^{-1}(p - p_b) + \sum_{i=0}^{n}(y_i - H_i[x_i, p])^T R_i^{-1}(y_i - H_i[x_i, p])$$

State control vector

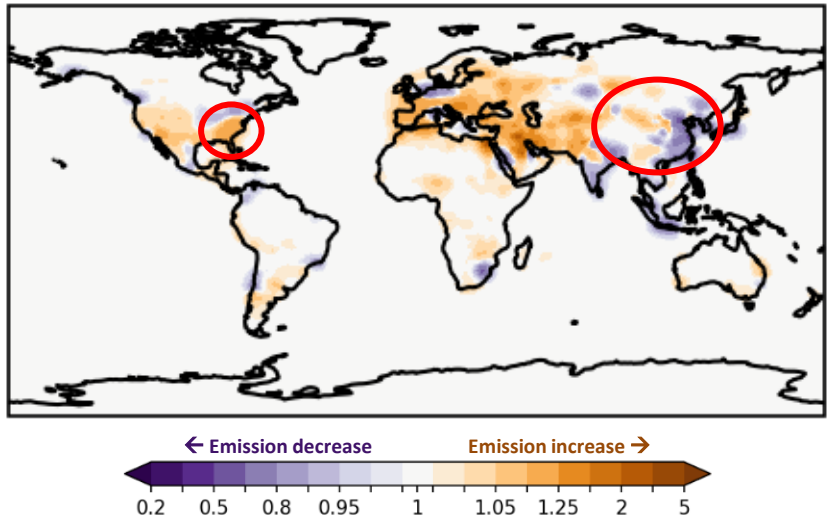
Parameter (e.g. scaling factors)

$J_o$: observation constraint

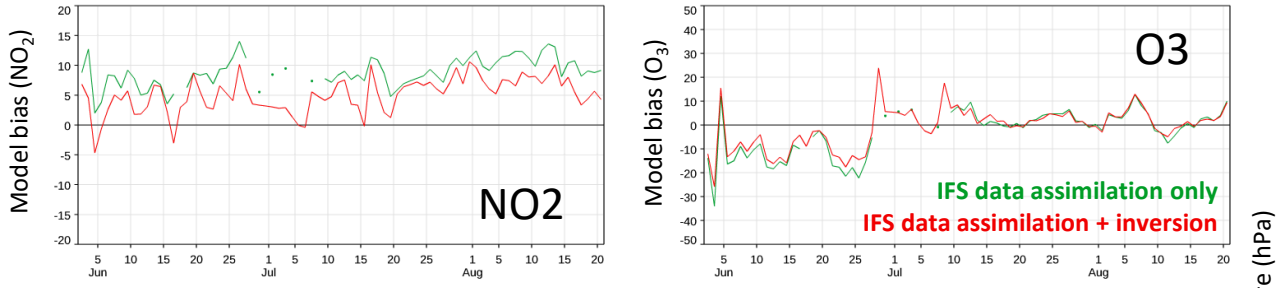
- IFS inversion system **jointly optimizes meteorology, concentrations and emissions** within each 12-hour assimilation window
- **Optimization of 2D scaling factor (SF) fields** applied to emission inventories or online-simulated emissions (anthropogenic, soils, biomass burning)
- IFS inversion currently being tested as **a candidate for operational implementation in IFS-COMPO CY51R1** (target species: $NO_x$, CO)

# IFS emission inversions: impact on $NO_2$ and $O_3$

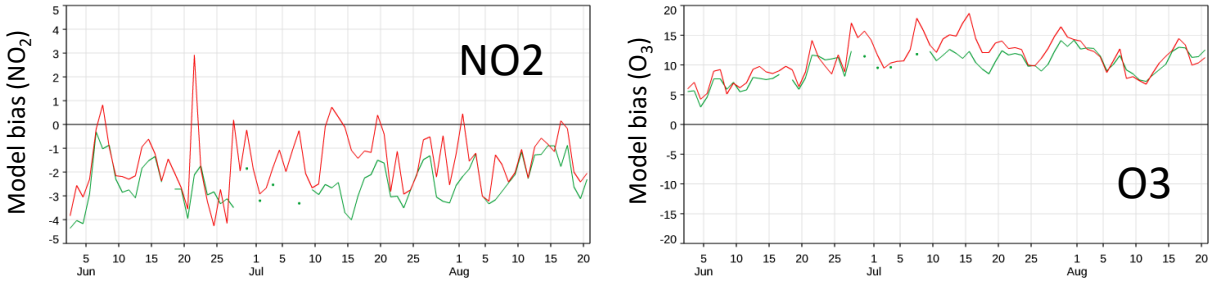
**Anthropogenic $NO_x$ emission posterior scaling factor**
(June-August 2025)




**China: reduced surface and boundary-layer $NO_2$ bias, small positive impact on $O_3$**




**Southeast US: $NO_2$ improvement, small $O_3$ degradation**
Role of biogenic VOCs emissions?



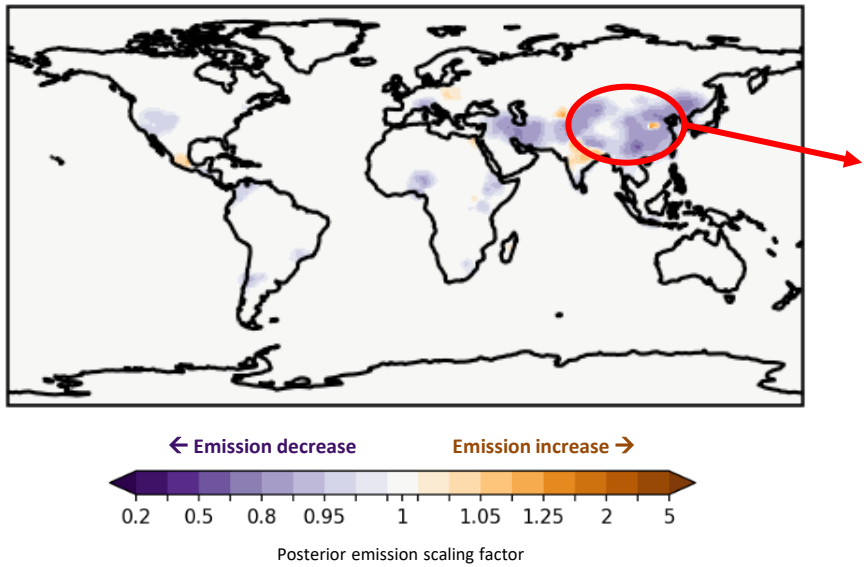

→ Optimized $NO_x$ emissions evaluated against independent surface and profile observations

→ The optimized $NO_x$ emissions show improvements in high-emission regions (esp. China) and highlight areas for model and inversion improvement

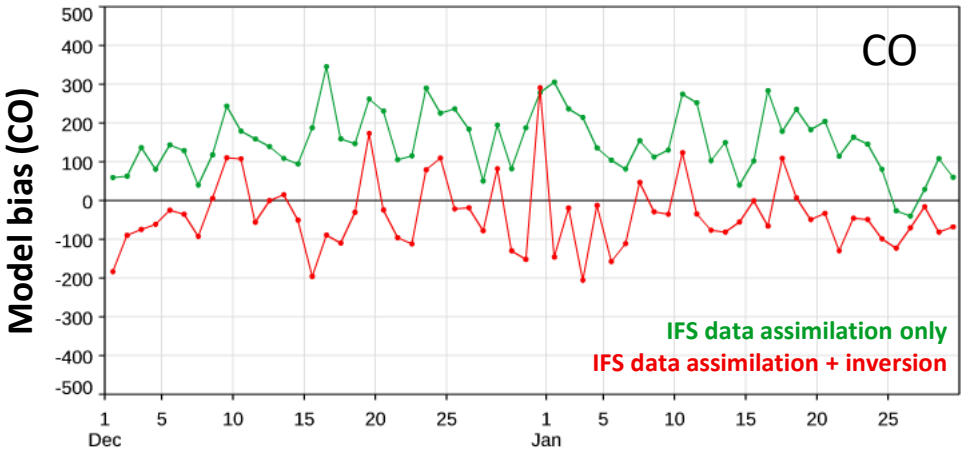
Credit: Auke Visser

# IFS emission inversions: impact on CO

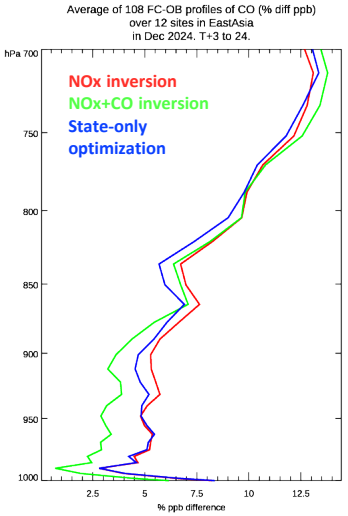
**Anthropogenic CO emission scaling factor**
(1-27 Dec 2024)




**IFS CO bias against independent surface (China AQ, below) and profile (IAGOS, right) observations**




Comparison with IAGOS data

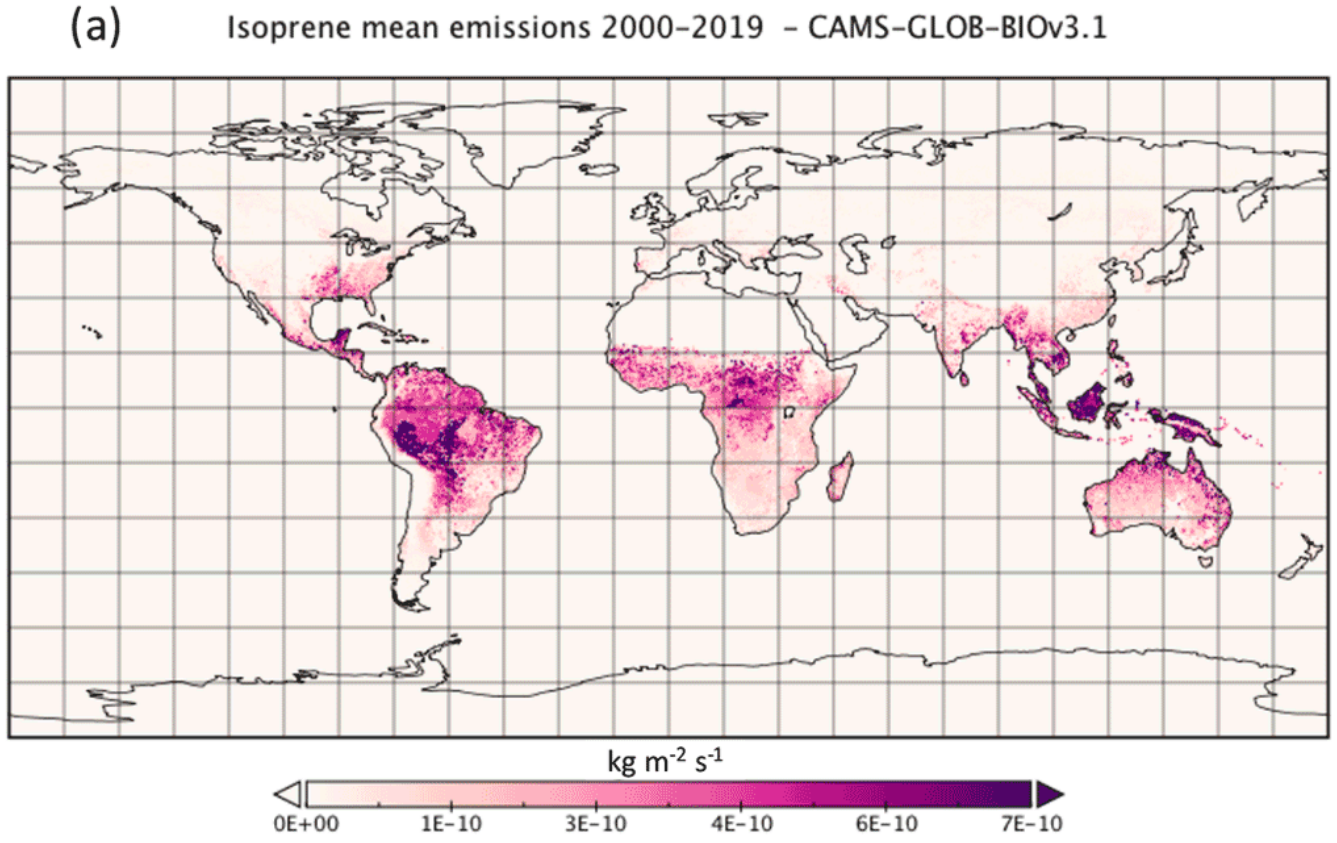



→ Widespread reduction of CO emissions leads to an improved match of IFS-COMPO with in-situ surface and profile observations in China

Credit: Auke Visser

Credit: Flora Kluge

# Biogenic compounds in the atmosphere



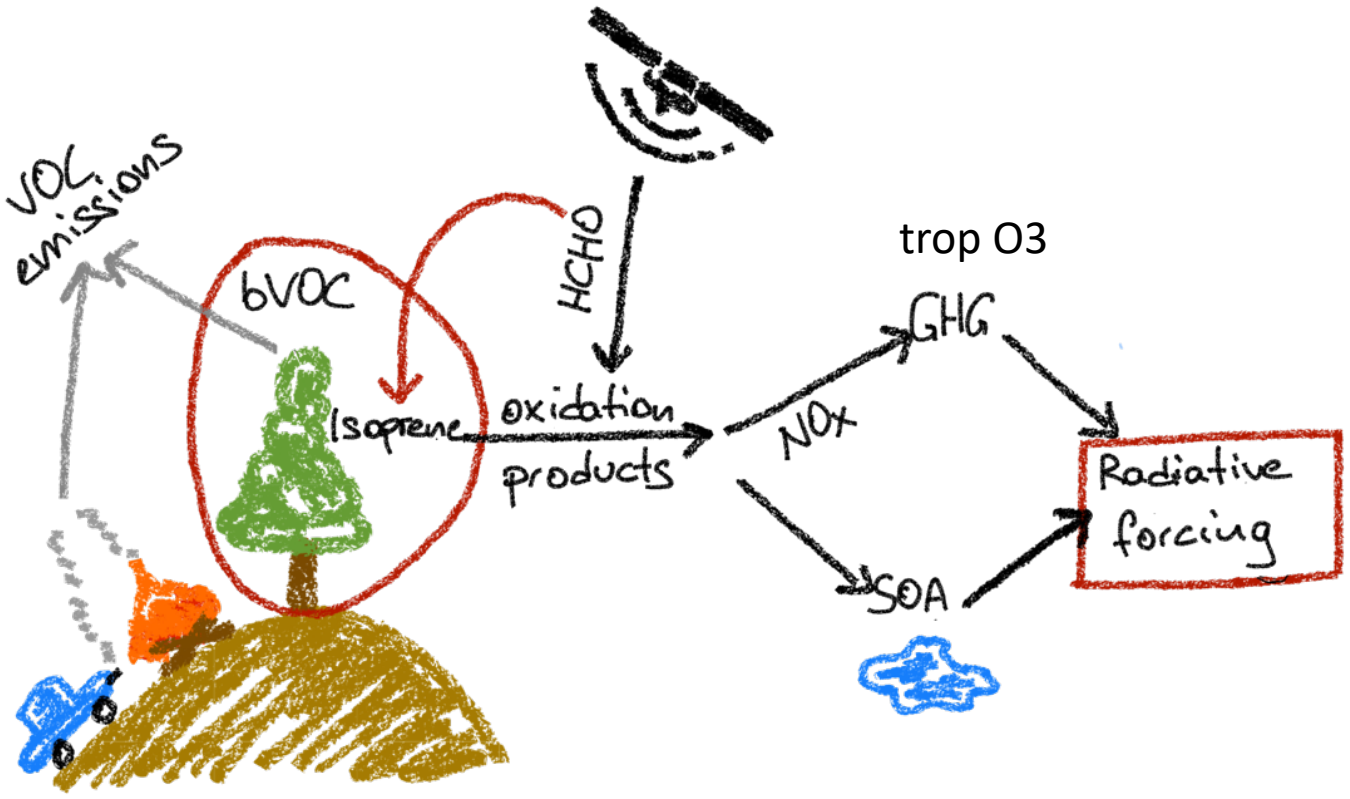

- Biogenic organic compounds are highly reactive and key drivers of atmospheric chemistry
- Pre-cursors for tropospheric O3 & contribute to formation of SOAs
- Direct impact on radiative forcing via scattering and absorption of sun light
- Indirect impact on radiative forcing: GHG concentrations & lifetime, OH, aerosols (cloud formation)
- Long-term increasing impact expected with increasing global surface temperatures (climate change)

- Total overall impact on chemistry and climate forcing largely uncertain
- Many involved feedback mechanisms not fully understood
- Emissions are highly variable in space and time

➢ **Better constrain biogenic emissions in IFS-COMPO by using satellite observations of bVOC products (HCHO)**

Credit: Flora Kluge

# Biogenic compounds in the atmosphere




- Biogenic organic compounds are highly reactive and key drivers of atmospheric chemistry
- Pre-cursors for tropospheric O3 & contribute to formation of SOAs
- Direct impact on radiative forcing via scattering and absorption of sun light
- Indirect impact on radiative forcing: GHG concentrations & lifetime, OH, aerosols (cloud formation)
- Long-term increasing impact expected with increasing global surface temperatures (climate change)

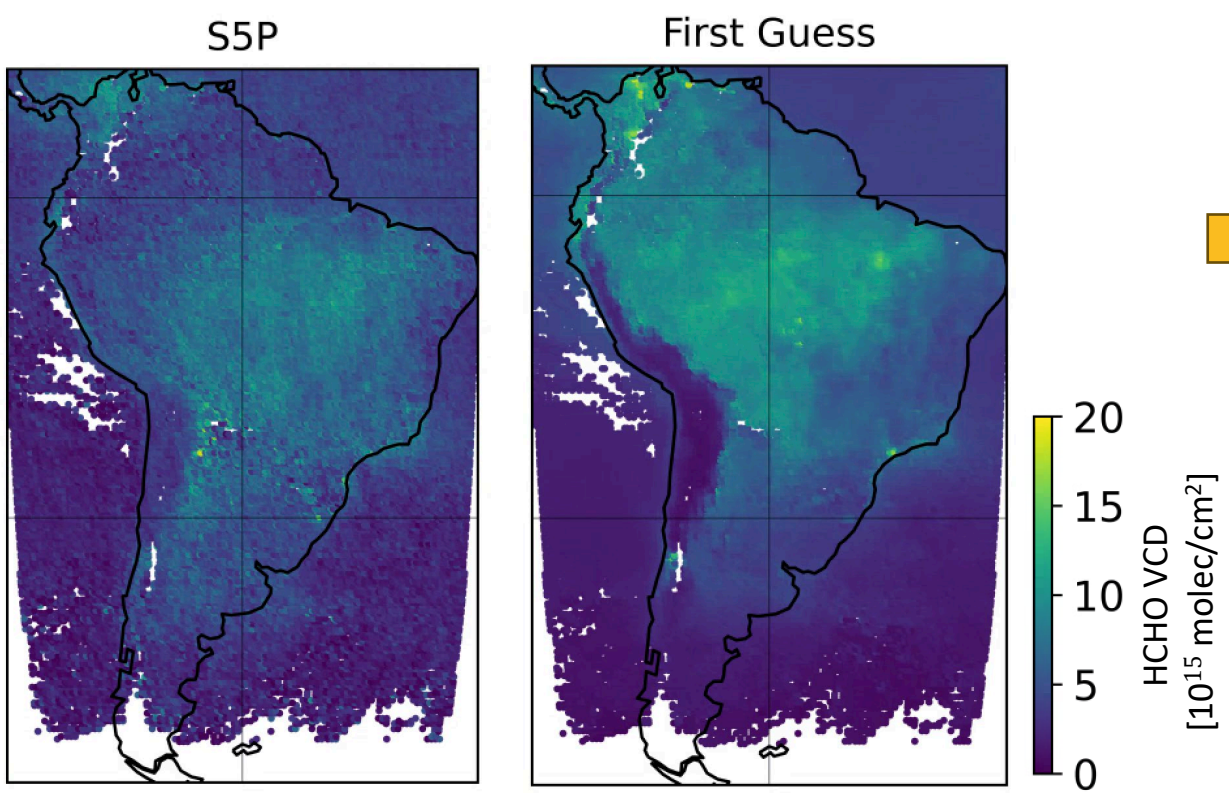
- Total overall impact on chemistry and climate forcing largely uncertain
- Many involved feedback mechanisms not fully understood
- Emissions are highly variable in space and time

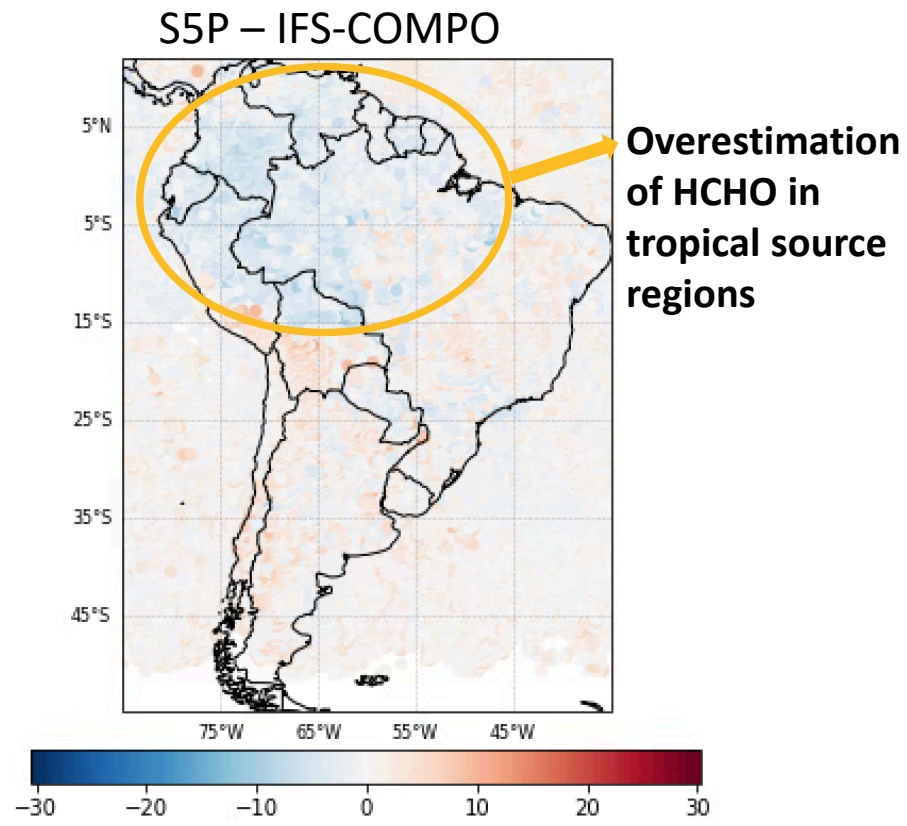
➢ **Better constrain biogenic emissions in IFS-COMPO by using satellite observations of bVOC products (HCHO)**

# S5P HCHO observation-based optimization of biogenic surface emissions in IFS-COMPO

Credit: Flora Kluge

South America:
Comparison to TROPOMI HCHO shows overestimation of biogenic compounds over Amazonian rain forest and underestimation in higher latitudes (20250602-20250609)

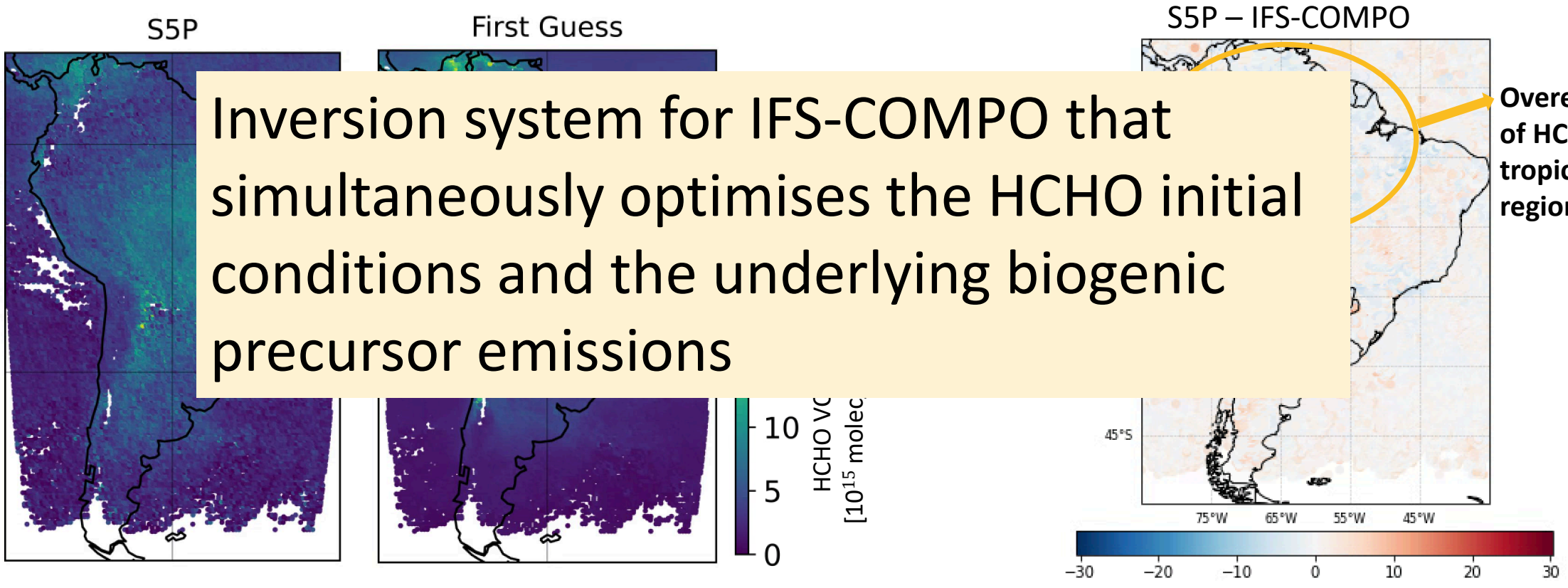


# S5P HCHO observation-based optimization of biogenic surface emissions in IFS-COMPO

Credit: Flora Kluge

South America:
Comparison to TROPOMI HCHO shows overestimation of biogenic compounds over Amazonian rain forest and underestimation in higher latitudes (20250602-20250609)



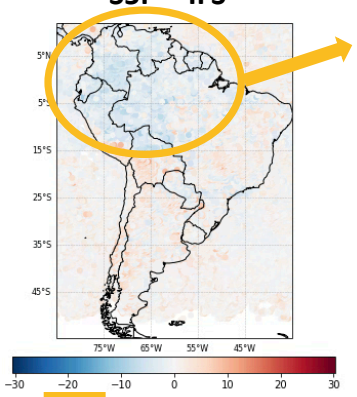

Inversion system for IFS-COMPO that simultaneously optimises the HCHO initial conditions and the underlying biogenic precursor emissions

**Overestimation of HCHO in tropical source regions**

Credit: Flora Kluge

# S5P HCHO observation-based optimization of biogenic surface emissions in IFS-COMPO

**Starting point: HCHO S5P – IFS**

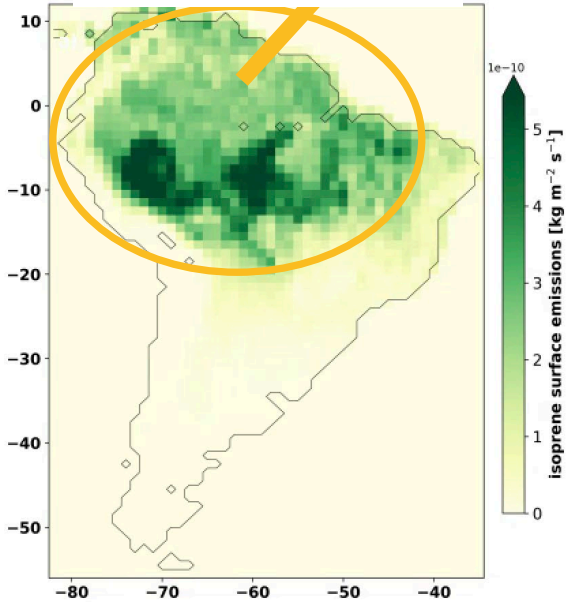


**Overestimation**

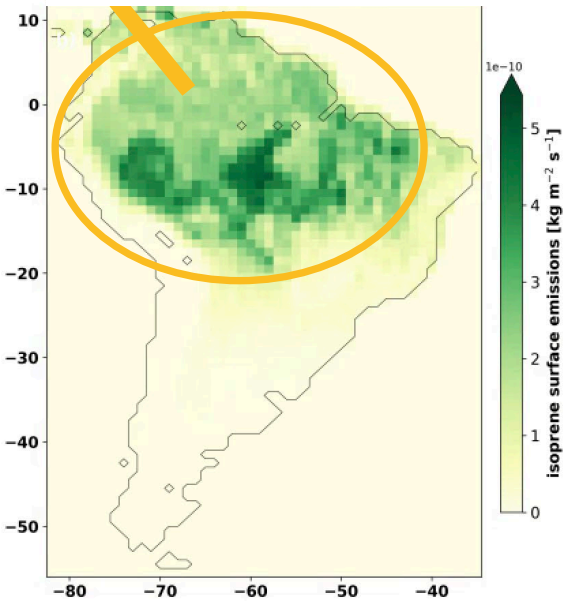
**-17% median emission reduction**

**IFS-COMPO BVOC assimilation and inversion system optimizes HCHO, isoprene, and isoprene surface emissions based on S5P, TEMPO, GEMS HCHO observations**

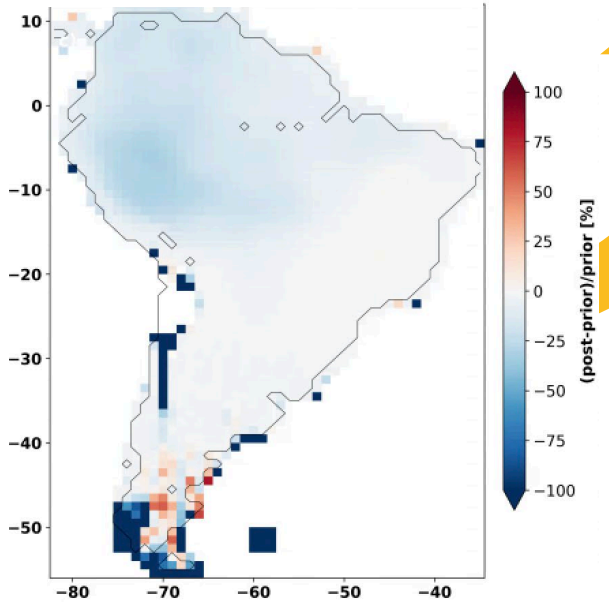
**isoprene prior emissions**



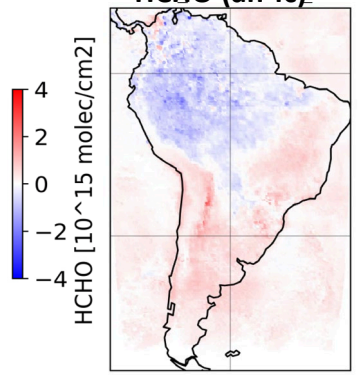

**isoprene optimised emissions**




**Change in isoprene emissions**




**Resulting change in HCHO (an-fc)**



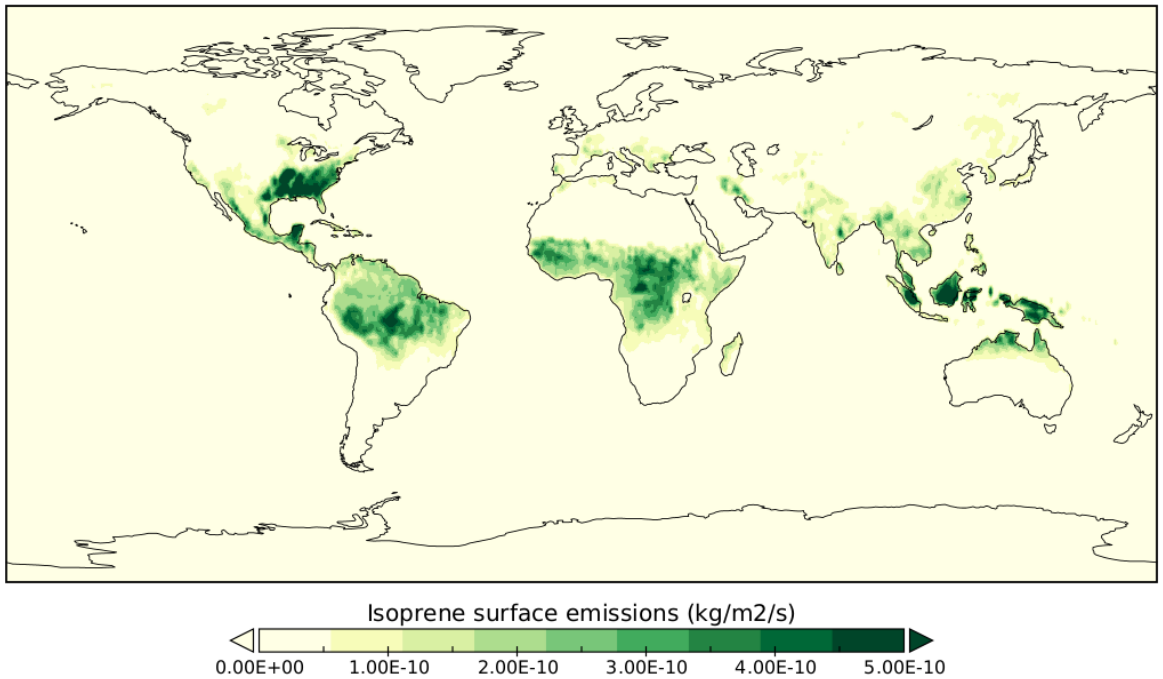

**Reduced observation bias by decreased HCHO in tropics and increased HCHO in higher latitudes**

Credit: Flora Kluge

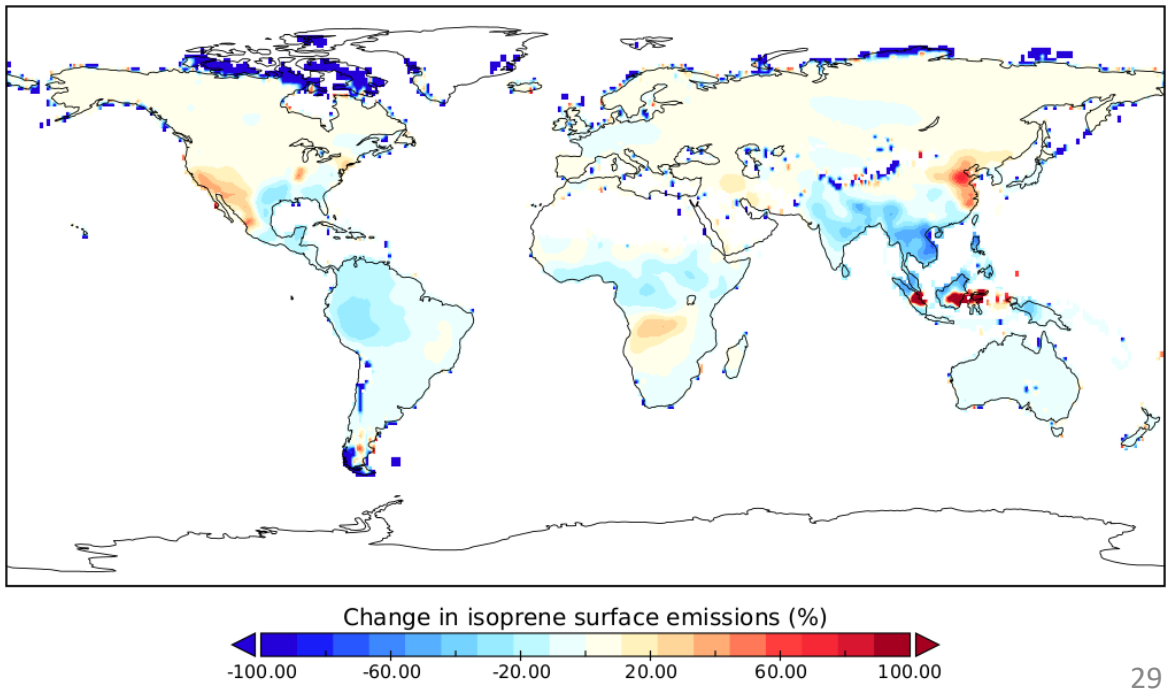
# Observation based optimization of biogenic surface emissions

## Optimized isoprene emissions using satellite observations of formaldehyde from the TROPOMI, GEMS, and TEMPO instruments

**Optimised isoprene surface emissions**
**1.6.2025 – 9.6.2025**




**Impact of the optimization:**
**Relative change of the prior emissions**
**1.6.2025 – 9.6.2025**



# Summary inversion

- Joint optimisation of emissions and atmospheric composition fields is now operationally feasible within the IFS-compo 4D-Var framework
- Emission inversion of reactive gases gives promising first results
  - e.g. improved NO2 forecasts over China
  - reduction of positive CO biases
- Biogenic emission inversion system was developed in HE CAMEO project and uses HCHO observations to constrain isoprene emissions
  - Reduction of -17% medium reduction of isoprene emissions over S-America
- Potential candidate for implementation in CY51R1 but more evaluation needed

Thanks to Cathy Li for her contributions

# 4. Impact of AC on NWP forecast scores

- ECMWF's NWP system currently uses a monthly aerosol climatology, see *https://www.ecmwf.int/en/newsletter/179/earth-system-science/new-time-varying-tropospheric-aerosol-climatology-ifs*
- CAMS uses prognostic aerosols interactively in the radiation scheme
- What is the benefit of using prognostic aerosol in the IFS radiation scheme compared to using the aerosol climatology?



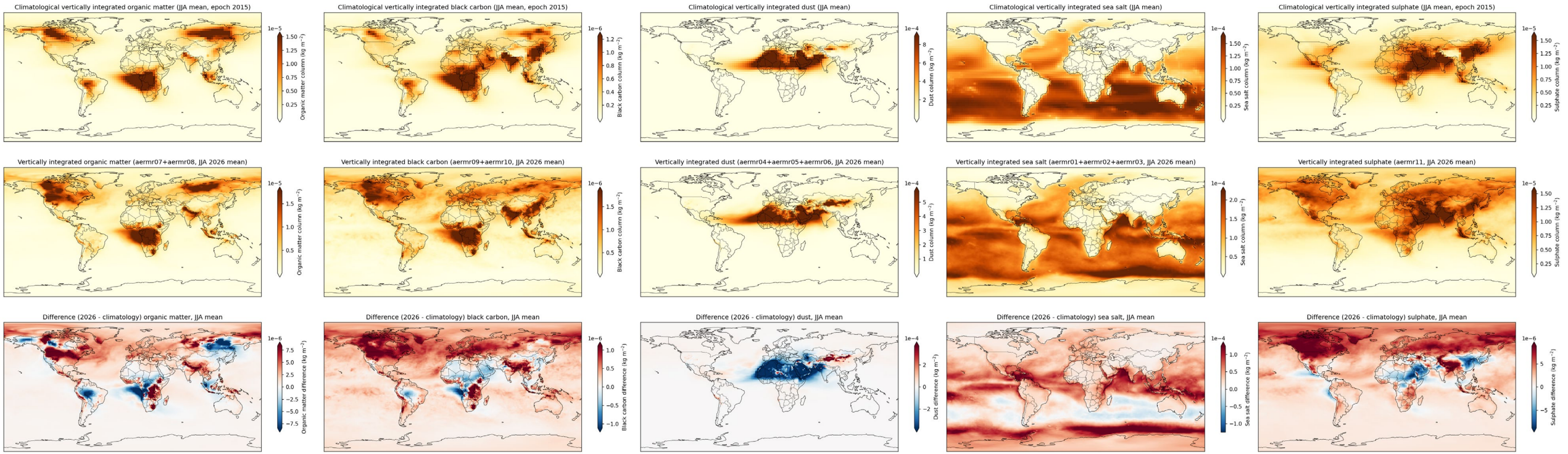
Pete Laing/BC Wildfire Service



Pavliha/iStockphoto

# Mean differences between climatological and prognostic aerosols (total column mass) JJA 2026

OM | BC | Dust | Sea Salt | SO4

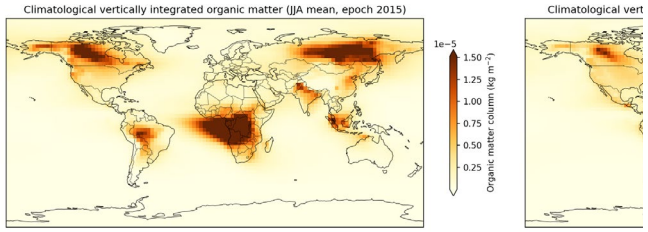


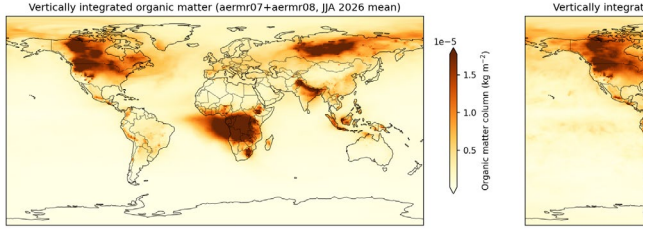

- Decadally-varying aerosol climatology based on CY47R3 CAMS full chemistry free-running simulations, see *https://www.ecmwf.int/en/newsletter/179/earth-system-science/new-time-varying-tropospheric-aerosol-climatology-ifs*
- Considerable differences between 2026 and climatology (2015 epoch)
- Strong biomass burning signals in OM and BC: Increase over N-America and Europe, reduction over S-America and parts of Africa

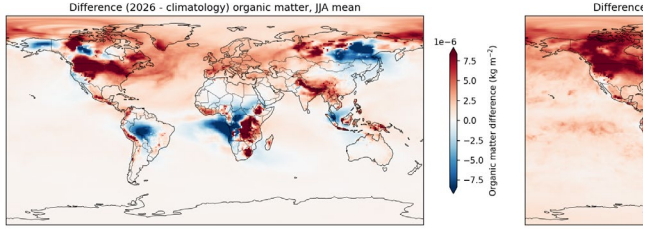
PROGRAMME OF THE EUROPEAN UNION

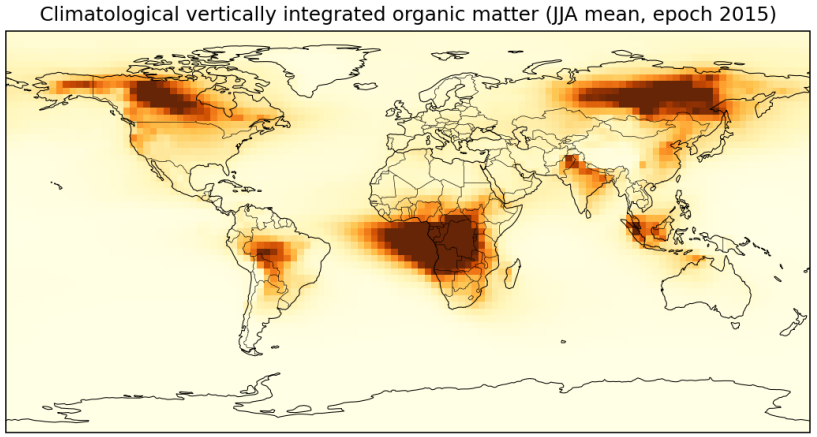
# Mean differen … gical and … prognostic ae … ass) JJA 2026

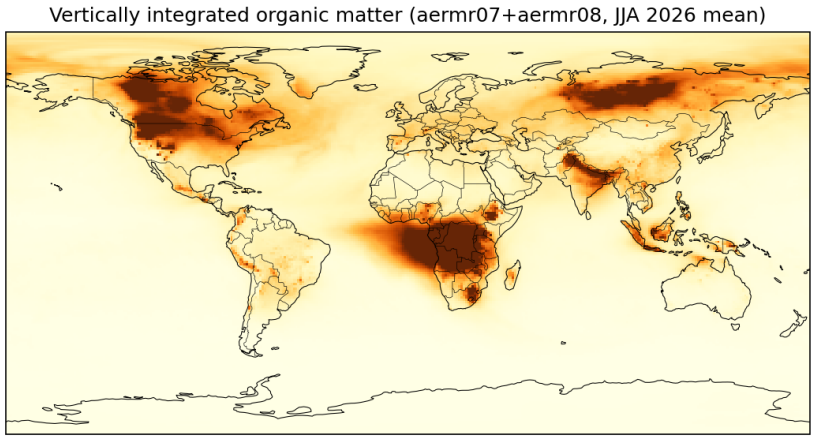
OM

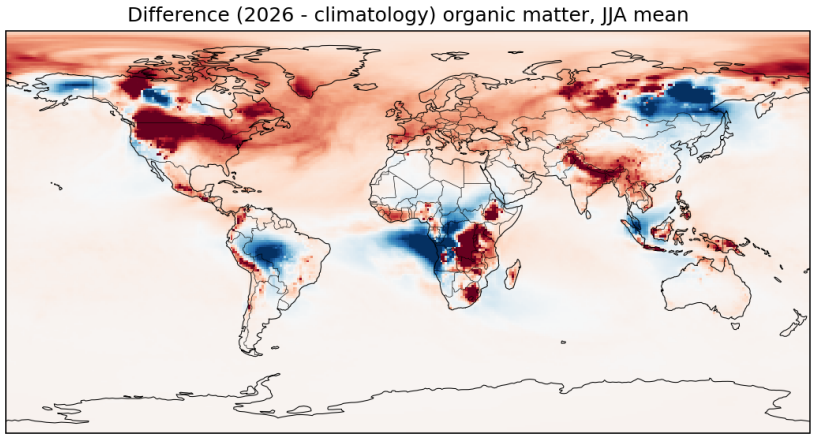


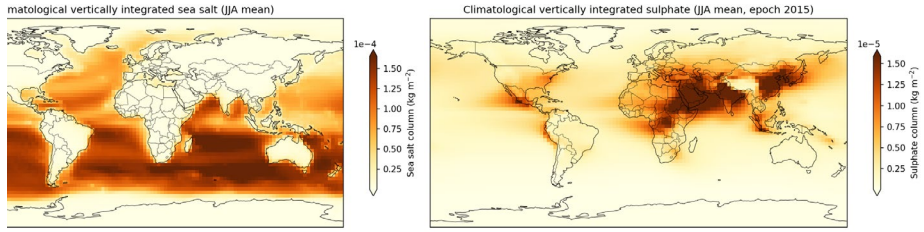



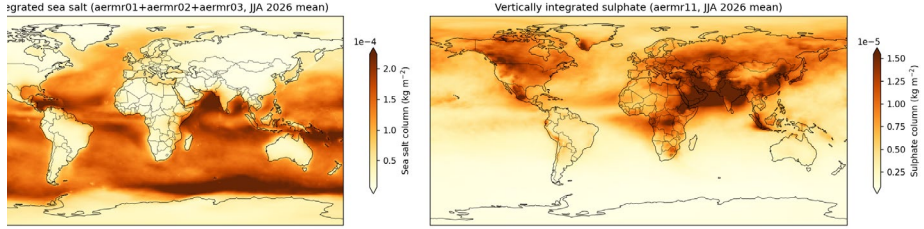



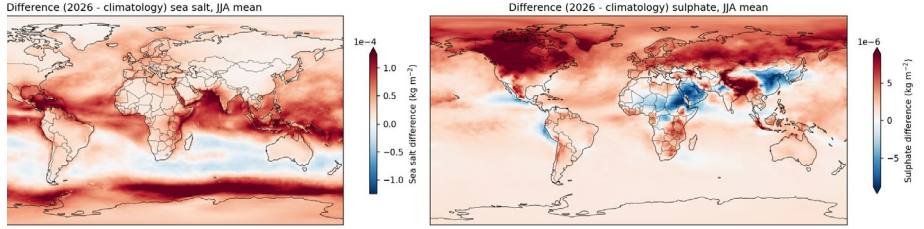




Sea Salt




SO4



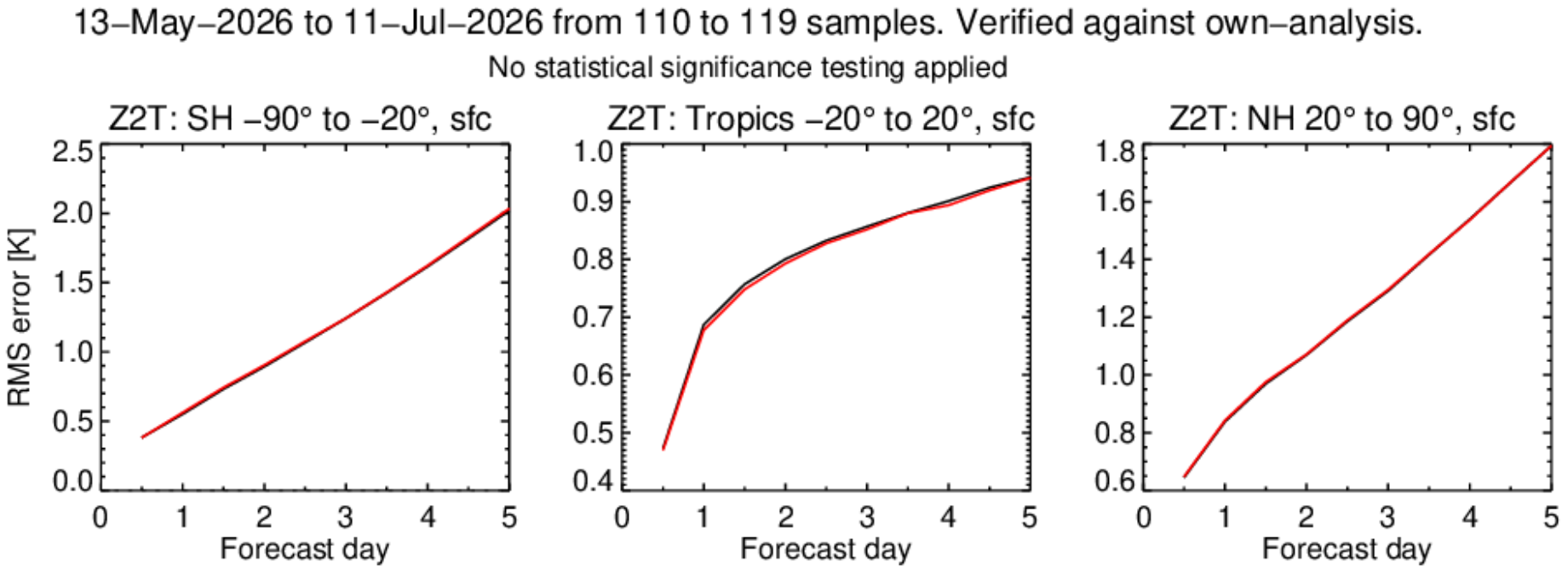







- Decadally-varying aerosol clir … g simulations, see *https://www.ecmwf.int/en/new… ospheric-aerosol-climatology-ifs*
- Considerable differences betv …
- Strong biomass burning signa … pe, reduction over S-America and parts of Africa

# Impact on 2mT forecast scores (CY50R1)



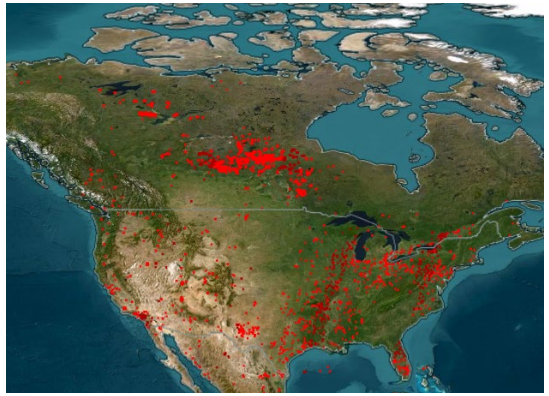

Prognostic AER
CLIM AER

Using prognostic aerosol makes no noticeable difference for “average” NWP scores, e.g. RMS 2mT with lead time

Credit: Cathy Li

# North American wildfires August 2025

- Impact of using prognostic aerosols in radiation scheme for large fires
- Reduction in 2mT under aerosol plume

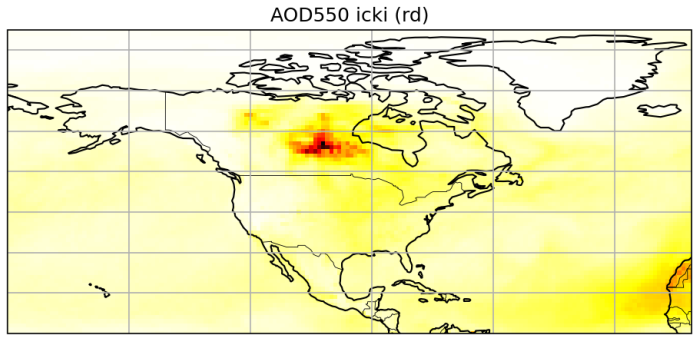


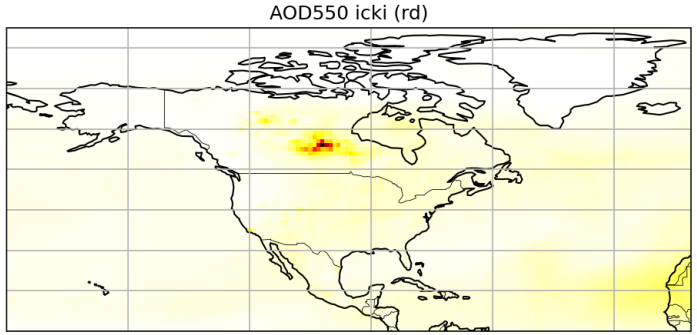
https://firms.modaps.eosdis.nasa.gov/map/

1-day fc

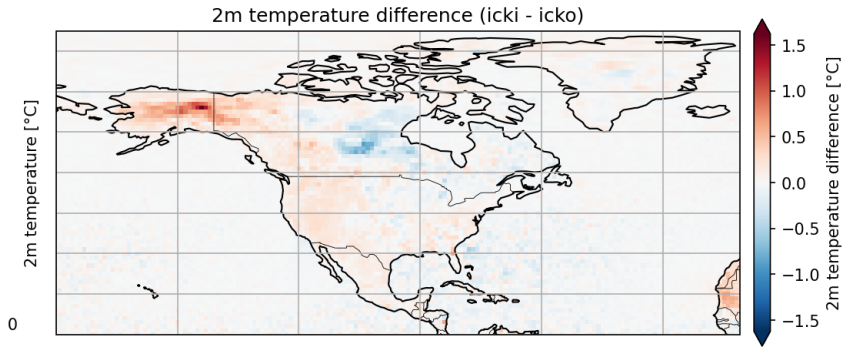
5-day fc

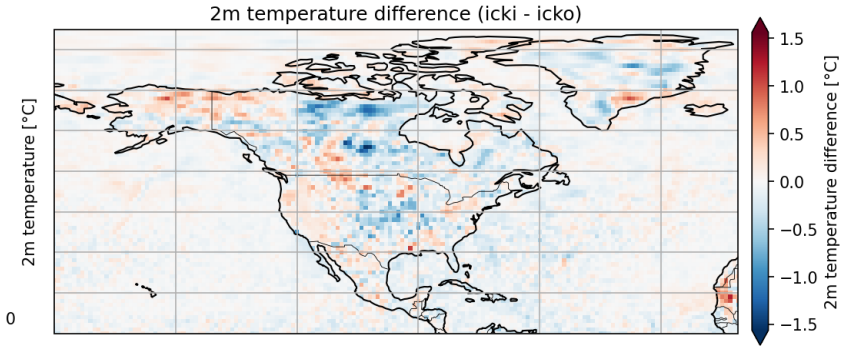
AOD550 1-12 Aug 2025




Difference 2mT: AERprog – AERclim

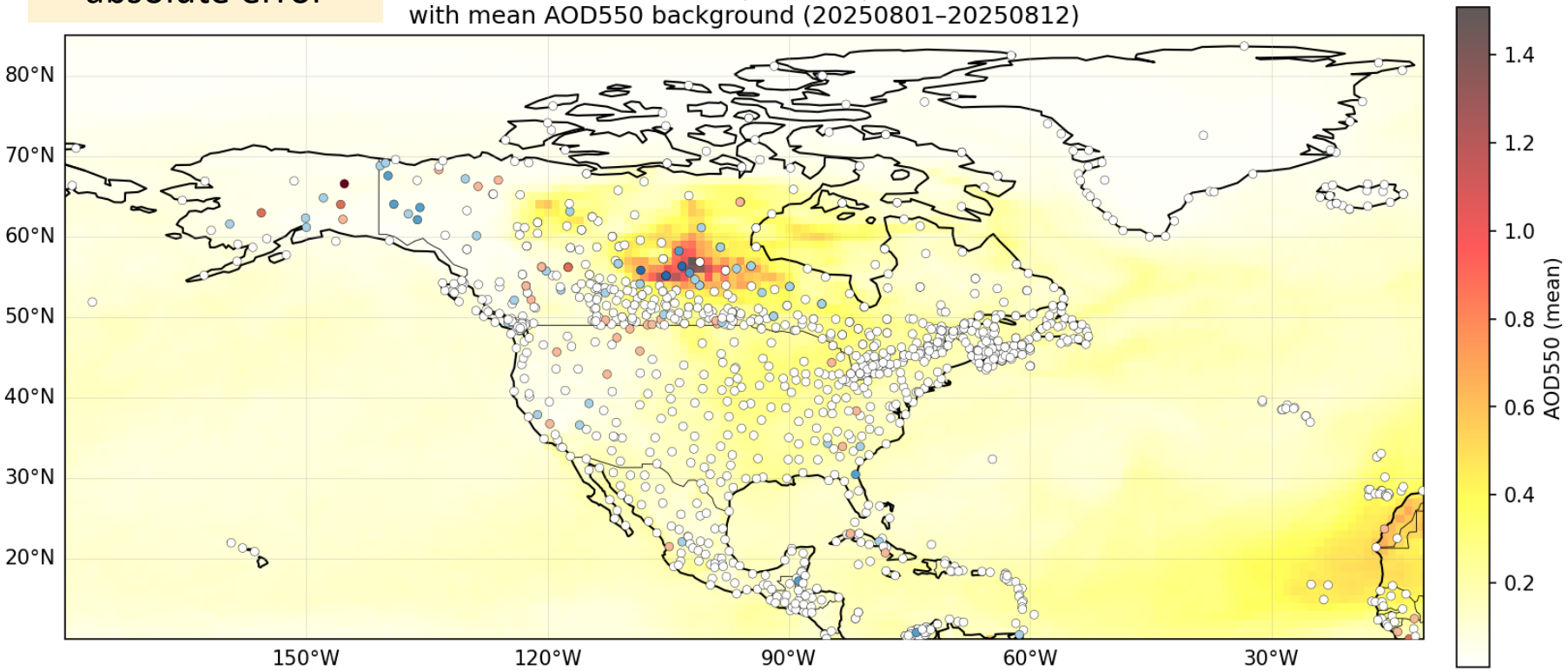


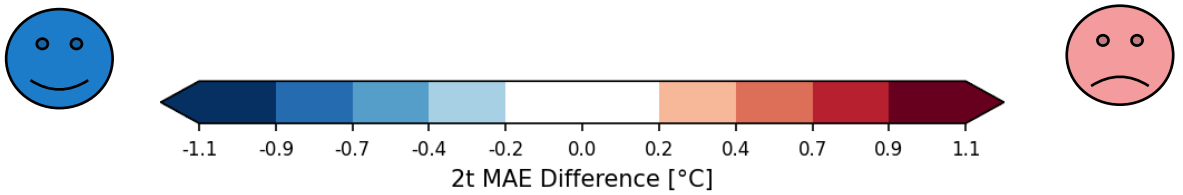



Credit: Cathy Li

# Comparison with 2mT synop observations: 1-day fc

Diff mean absolute error



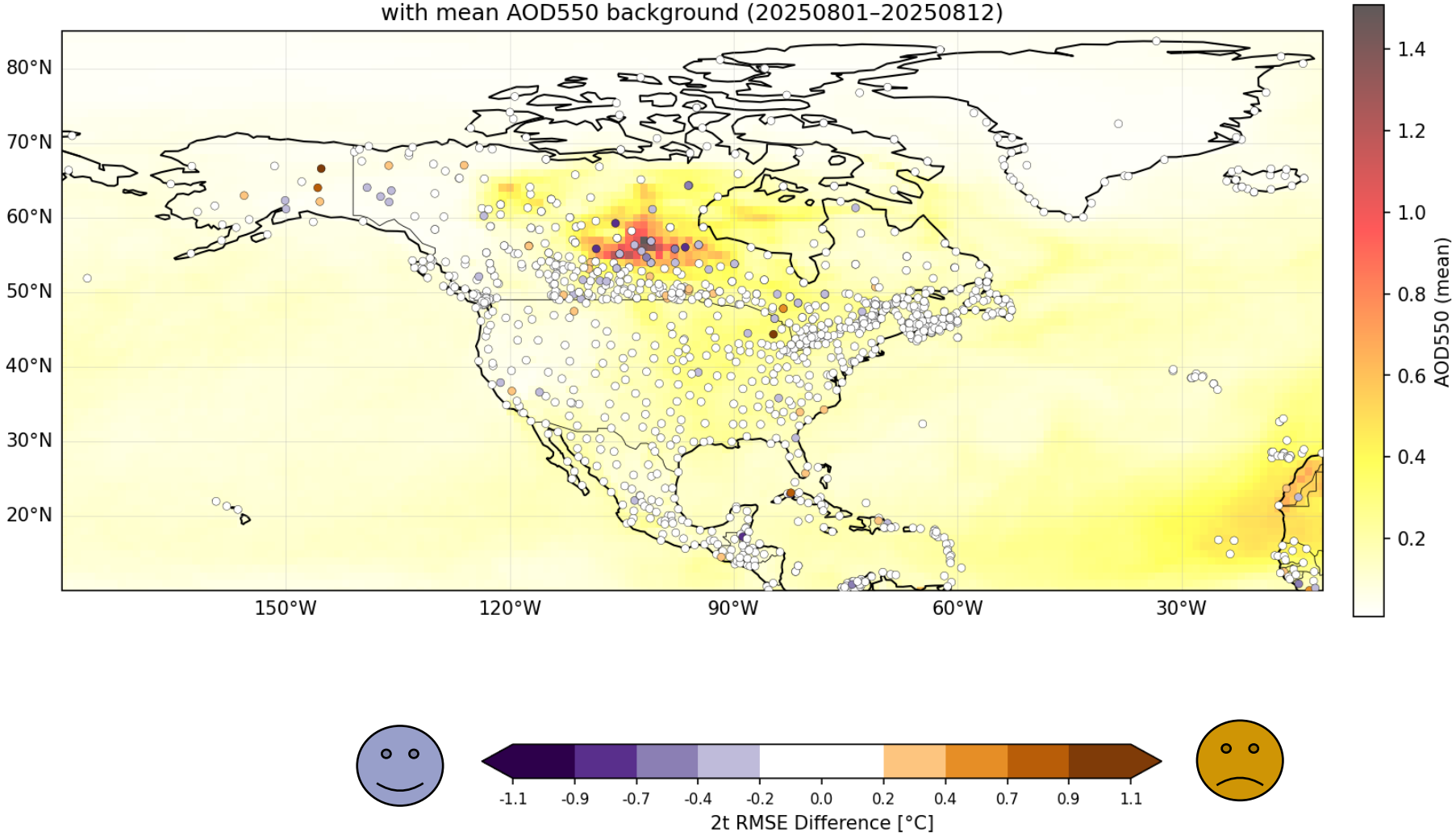




- Significant reduction of MAE and RMSE with interactive aerosols at lead time of +24h.

Credit: Cathy Li

# Comparison with 2mT synop observations: 1-day fc

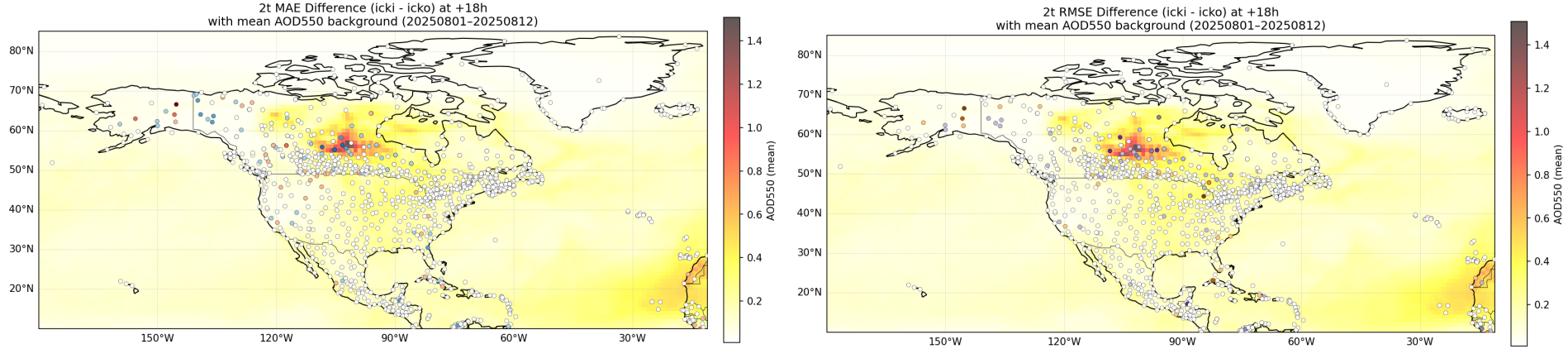
Diff RMSE




- Significant reduction of MAE and RMSE with interactive aerosols at lead time of +24h.

Credit: Cathy Li

# Comparison with 2mT synop observations



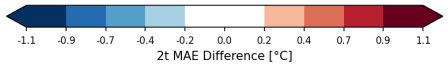


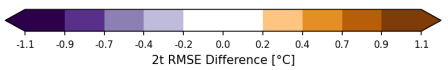


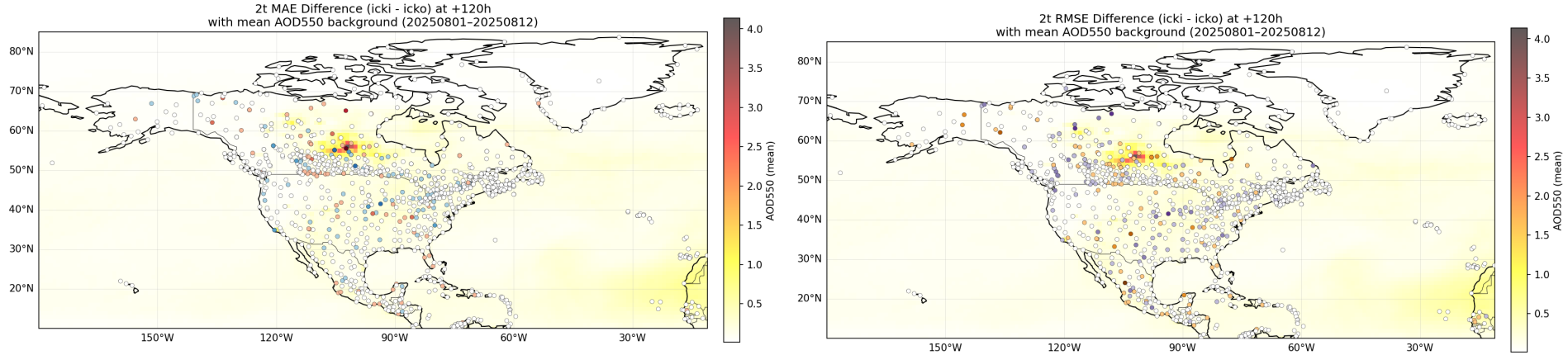
Mean absolute error

1-day fc

Mean RMSE



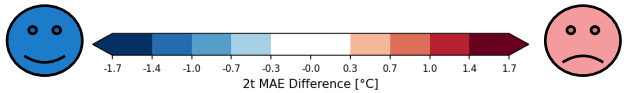



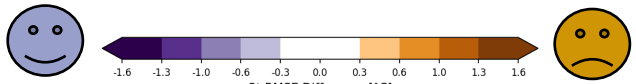







5-day fc

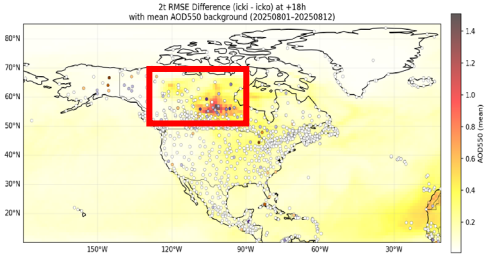


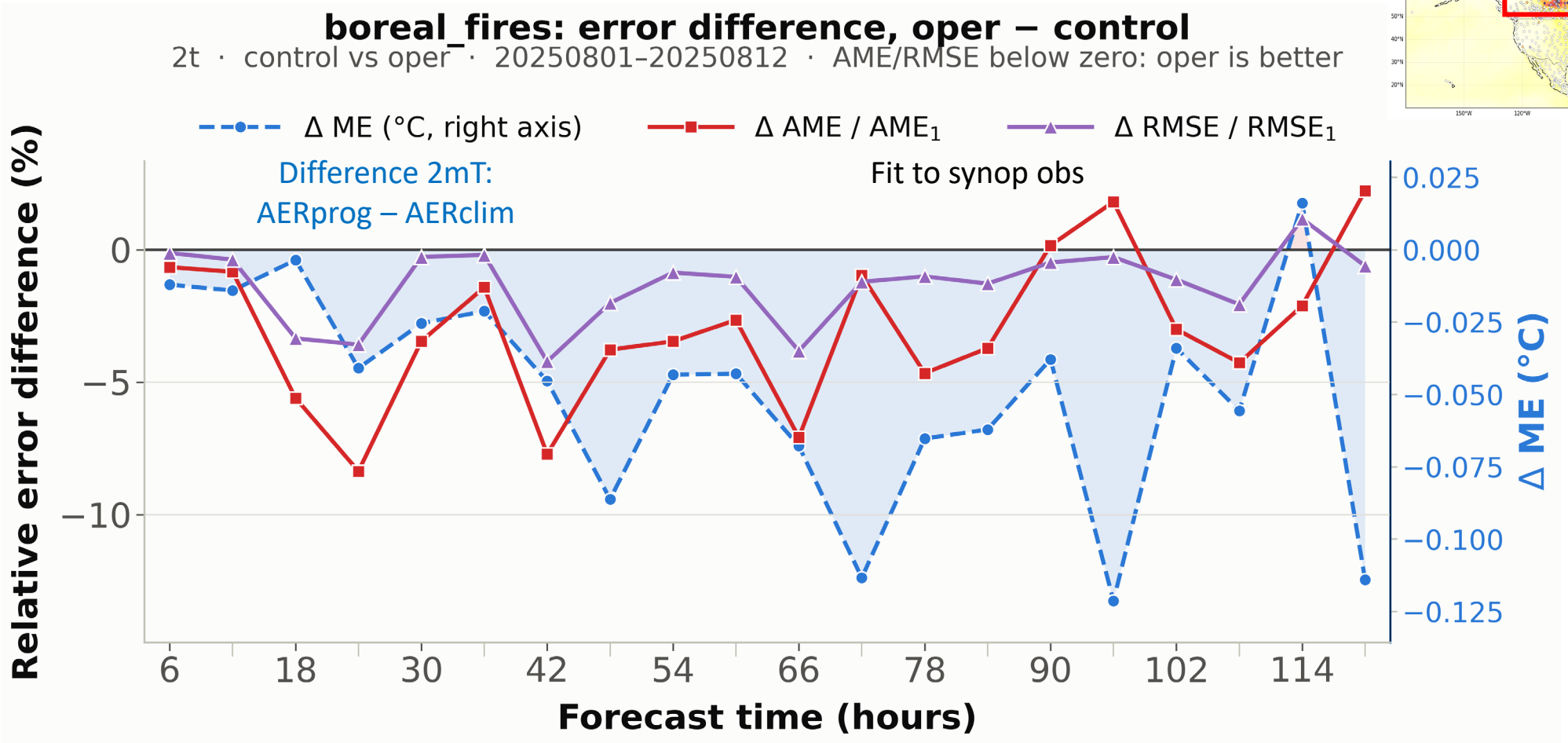




- Significant reduction of MAE and RMSE with interactive aerosols
- especially at lead time of +24h.
- At the lead time of +120h, the improvement is more significant south of the plume.

Credit: Cathy Li

# Comparison with 2mT synop observations

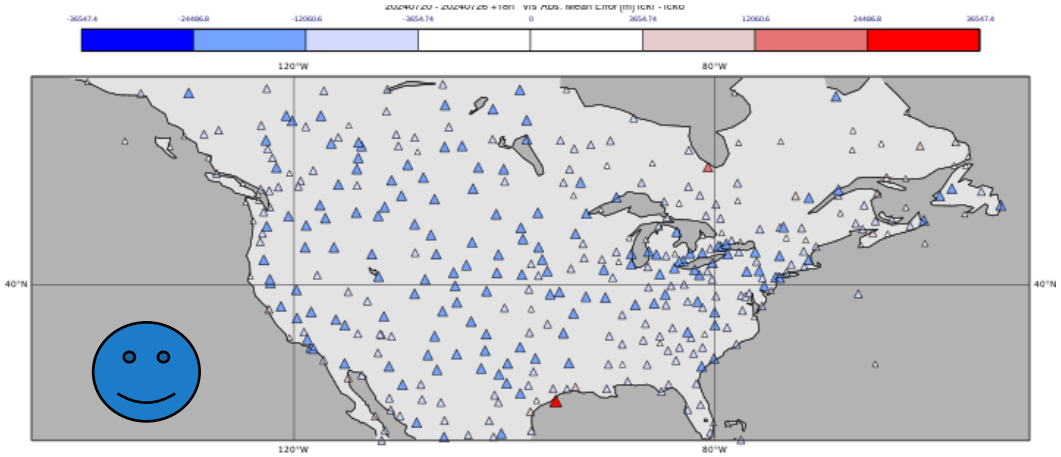


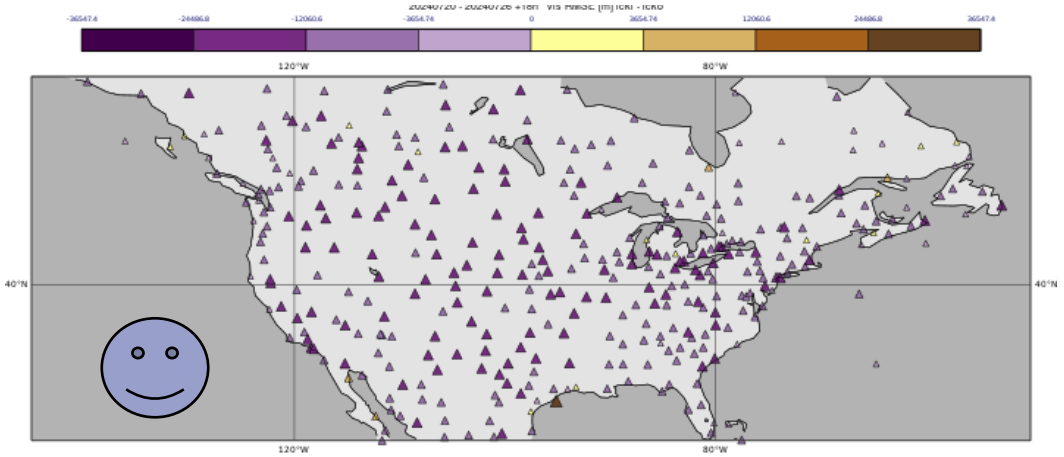
Credit: Cathy Li

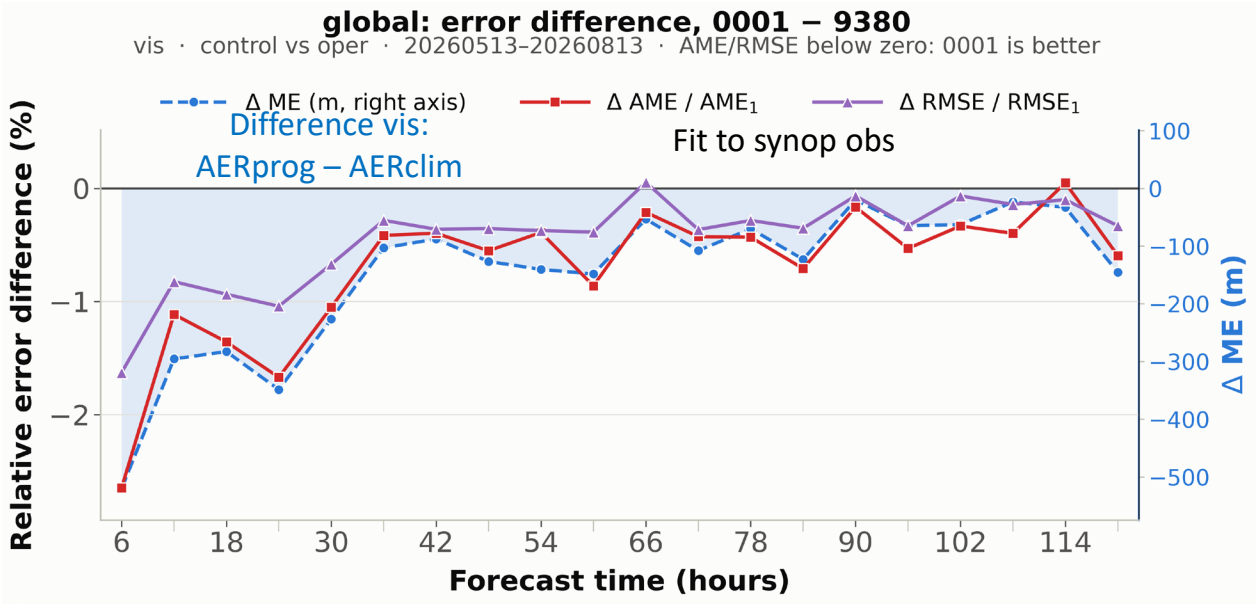
# Visibility forecast (20-26 July 2024)

Absolute Mean Bias (icki-icko) of vis (20-26 July 2024)



RMSE (icki-icko) of vis (20-26 July 2024)






- Global mean visibility lower in AERprog
- Improved fit to visibility obs seen globally and across all lead times

# Summary interactive aerosols

- Prognostic aerosols can lead to cooling of up to several K in case of large AOD anomalies
- Using prognostic aerosols improves the forecasts of both 2m temperature and visibility
- For 2m temperature, the impact from the aerosol settings is mostly local, i.e. noticeable in the cases/areas where AOD > 1
- No overall improvement in standard global or hemispheric NWP scores for 2mT
- For visibility, the improvement is global (regardless of the events).

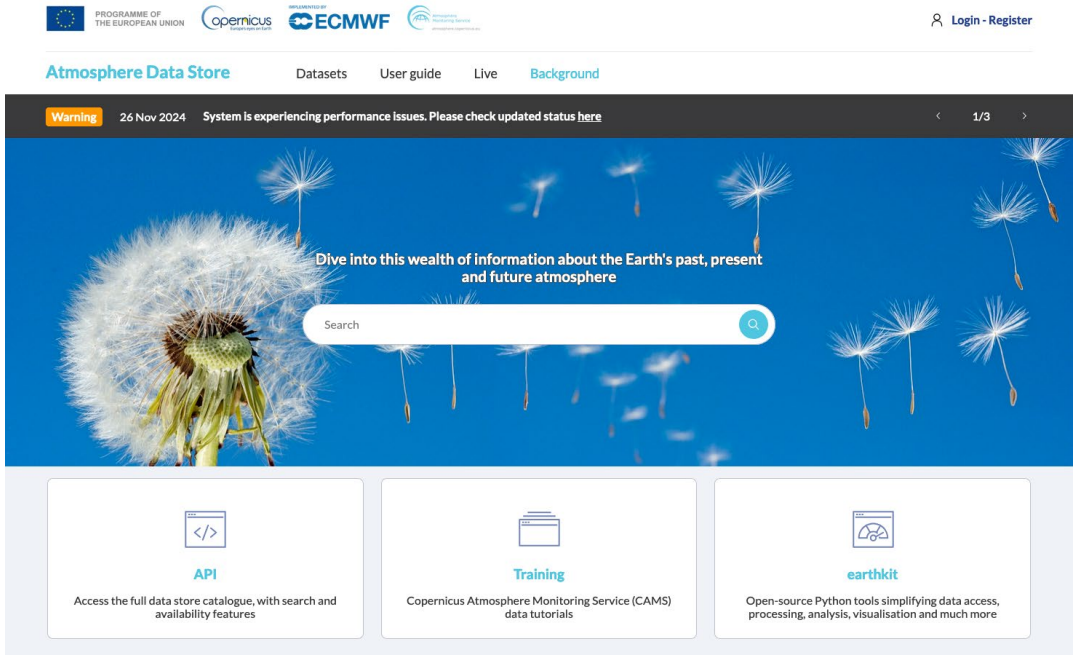
- Prognostic aerosol in IFS for NWP purposes is ongoing research topic at ECMWF
- Potential to improve medium-range and monthly seasonal forecast recognised

# 5. Summary

- CAMS provides a wide range of atmospheric composition products
  - Useful in their own right and freely available
- Exciting potential to improve NWP forecasts with atmospheric composition information
  - Use of prognostic aersols interactively in the radiation scheme (used by CAMS)
- AIFS-compo shows great results
- Emission inversion can improve atmospheric composition analysis and forecast fields

# The Atmosphere Data Store (ADS)

All CAMS data are freely available: **https://ads.atmosphere.copernicus.eu/**




Copernicus EU

@copernicusecmwf

Copernicus ECMWF

www.copernicus.eu
atmosphere.copernicus.eu

@copernicusecmwf.bsky.social

**Documentation of atmospheric composition in IFS:**
https://www.ecmwf.int/en/elibrary/81738-ifs-documentation-cy50r1-part-8-atmospheric-composition

**Validation reports and validation servers:**
https://atmosphere.copernicus.eu/quality-assurance

PROGRAMME OF THE EUROPEAN UNION

Copernicus Europe's eyes on Earth

Atmosphere Monitoring Service atmosphere.copernicus.eu

IMPLEMENTED BY ECMWF

#EUSpace



# Thank you


**© European Union 2023**

Unless otherwise noted the reuse of this presentation is authorised under the CC BY 4.0 license. For any use or reproduction of elements that are not owned by the EU, permission may need to be sought directly from the respective right holders.

Slide xx: element concerned, source: e.g. Fotolia.com; Slide xx: element concerned, source: e.g. iStock.com