# Data Assimilation: Integrating satellite data into the CAMS global system

# Antje Inness (ECMWF)

Fourth Joint School on Atmospheric Composition 28 Sept – 6 October 2022













# Outline

Atmosphere Monitoring

- 1. Data assimilation methodology for atmospheric composition
- 2. Emission inversion
- 3. Potential issues when assimilating satellite data
- 4. Reanalysis
- 5. Summary

With thanks to the ECMWF CAMS team, particularly Melanie Ades and Nicolas Bousserez.





# **1. Data Assimilation Methodology**



2









# Variational Data assimilation

Atmosphere Monitoring

- NWP definition: Combining data and model in an 'optimal' way to produce the best possible initial conditions for a numerical forecast
- Optimal in a statistical sense: minimize error and/or maximize probability of the analysis being correct
- CAMS uses ECMWF's 4-dimensional variational data assimilation system or 4D-Var
- For AQ other parameters than
   IC might be of interest





# Cost function

Atmosphere Monitoring Data assimilation for atmospheric composition is in principle no different from NWP data assimilation data assimilation

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

x: control vector

x<sub>b</sub>: 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 constrained 4D-Var includes a model error term

Europe's eyes on Earth

THE EUROPEAN UNIO

# Data assimilation methodology

Atmosphere Monitoring

47

Data assimilation for atmospheric composition is in principle no different from NWP data assimilation

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

$$\frac{NWP}{\substack{\text{vorticity}\\ \text{divergence}\\ \text{surface pressure (logarithm)}\\ \text{specific humidity}}$$

$$Atmospheric Composition$$

$$\frac{Ozone}{arbon monxide}\\ nitrogen dioxide\\ formaldehyde\\ \text{sulphur dioxide}\\ \text{carbon dioxide}\\ methane\\ aerosol mixing ratio$$

$$W = \sum_{i=0}^n (y_i - H_i[x_i])^T R_i^{-1}(y_i - H_i[x_i]) + \sum_{i=0}^n (y_i - H_$$







# Incremental 4D-Var



# **2. Emission Inversion**



2









# Initial condition vs boundary problem



- 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

**Atmosphere** 

Monitoring

 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



# Short-lived memory of NO2 assimilation



# OMI NO2 analysis increment [%]





Lonaitude

-15

Differences between Assim and CTRL (JF 2008)

Difference between 12h forecasts from ASSIM and CTRL (JF 2008)

20

25

[10<sup>15</sup> molec/cm<sup>2</sup>]

30

100

10



- Large differences between analyses of ASSIM and CTRL
- Impact is lost during subsequent 12h forecast
- Constraining emissions (in addition of IC) would give a better initial state and persistence of forecast improvements throughout the DA window
   Inness et al. (2015, ACP)

VF

# Examples of emissions

### TNO European anthropogenic NOx emissions





CO2 fluxes

# CAMS\_GLOB biogenic CO emissions



Volcanic SO2

0.5 1 2 3 4 5 10 20 30 40 50 100 200 500 100





### 

# Biomass burning, 15 October 2017



# Emissions

Atmosphere Monitoring

- 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





# Joint state/emissions 4D-var inversion system

**Atmosphere** Monitoring



J<sub>o</sub>: observation constraint

- ٠ Joint optimisation of emissions and initial conditions
- Optimized emissions e.g. CO2, CH4, CO & NO2 ٠
- TL/AD of simplified chemistry: link between NO emissions and NO2 observations ٠
- 2D scaling factors p applied to emission fields ٠
- Prior error definition:
  - Global constant or 2D map of standard error
  - Spatial correlation length scale (via  $B_{p}$ )
  - NO/CO2 emission error correlation in  $B_p \rightarrow NO2$  obs can contrain CO2 emissions



Credit: Nicolas Bousserez

# Impact of Covid lockdown on US anthropogenic emissions

 Differences between posterior emissions May 2020 – May 2019 show impact of covid lockdown

Atmosphere

- Based on CAMS operational emissions in the prior and a fixed prior uncertainty of 40%.
- 10-20% reduction consistent with previous studies (e.g., *Keller et al.* (2020); Liu et al. (2020))
- Provided uncertainties in NO/CO<sub>2</sub> emission ratios are accounted for, topdown NO<sub>2</sub> estimates could help quantify CO<sub>2</sub> emissions variability

NO<sub>x</sub> emission changes (%) between May 2020 and May 2019



# **3. Potential issues when assimilating satellite data**











# Biases and Bias correction

Atmosphere Monitoring



- 4D-Var assumes that errors are unbiased and gaussian
- Biases in the input [y h(x)] can arise from:
  - errors in the actual observations
  - errors in the model background
  - errors in the observation operator
- Often no true reference in the real world
- Need an adequate bias model to correct for this
  - Correct observation biases before use
  - Adaptive bias correction system (better for complex and changing observation system)



Europe's eyes on Earth

IMPLEMENTED BY

**EEFCMWF** 





# Examples of variational bias correction

Atmosphere Monitoring

Variational bias correction:

- Spins up
- Adjusts to changes in model or data
- Removes biases





# Observation operators: Mapping model and obs Atmosphere nMonitoring N (n = U[M(n)]) $T D^{-1}(n = U[M(n)])$

$$\sum_{i=0}^{n} (y_i - H_i[M_i(x_0)])^T R_i^{-1} (y_i - H_i[M_i(x_0)])$$

Observation operator **H** maps model state at beginning of the assimilation window (t=0) to the observation time and location

# Direct assimilation of radiance observations:

The observation operator must incorporate an additional step to compute radiances from the model state variables (radiative transfer model, e.g. RTTOV)

CAMS hopes to explore this in the future

# Assimilation of retrievals:

Good characterization of retrieval is crucial:

Europe's eyes on Earth

- Averaging kernels
- A priori
- Error estimates
- Quality flags





# Assimilating retrievals: Column retrieval

Atmosphere Monitoring

# We can make use of the averaging kernel A in the observation operator

$$\hat{\mathbf{y}} = \mathbf{y}_{a} + \mathbf{A}(\mathbf{y} - \mathbf{y}_{a}) + \varepsilon$$

Retrieved value: true state y smoothed by the averaging kernel A;  $y_a$ : a-priori,  $\epsilon$ : retrieval error

$$d = y^{\hat{}} - H(\mathbf{x}_m) = \mathbf{y}_a + \mathbf{A}(\mathbf{y} - \mathbf{y}_a) + \varepsilon - H(\mathbf{x}_m)$$

Without averaging kernels in observation operator (e.g. simple vertical integral)

$$d = \hat{y} - \hat{H}(\mathbf{x}_m) = \mathbf{y}_a + \mathbf{A}(\mathbf{y} - \mathbf{y}_a) + \varepsilon - (\mathbf{y}_a + \mathbf{A}(H(\mathbf{x}_m) - \mathbf{y}_a))$$

$$= \mathbf{A}(\mathbf{y} - H(\mathbf{x}_m)) + \varepsilon$$

With averaging kernels in observation operator

- We remove the influence of the a-priori profile if we use the averaging kernel to sample the model profile according to the assumptions made in the retrieval.
- The model data is smoothed by the averaging kernel to produce a profile that is directly comparable to the product derived from the instrument radiances
- We still need to know y<sub>a</sub> and A in the observation operator calculations

Examples of total column CO averaging kernels

W. Australia, 11/01/2000, Daytime

100



MOPITT AKs Deeter et

Deeter et al. (2003, JGR)

W. Australia, 11/01/2000, Nighttime

- night day Pressure (mb) Pressure (mb) 100 1000 -0.2 0.2 -0.2 0.2 0.0 0.4 0.6 0.0 0.4 0.6 Mean CO Averaging Kernels Mean CO Averaging Kernels
- TIR retrievals give information about mid troposphere
- Diurnal variations of Tsurf affect TIR retrieval over land
- CO near surface more detectable during day, AKs shift downwards
- Diurnal variability of AKs largest over e.g. deserts, smallest over sea
- Clear TROPOMI data have some sensitivity to lower troposphere and PBL

THE EUROPEAN UNION

Europe's eyes on Earth



# Impact of TROPOMI CO retrievals in DA





# Increment from one TC ozone retrieval



# Increment from one TC ozone retrieval



Formulation of the B-matrix is very important for AC

# An extreme example: Ozone 7 October 2004

# Atmosph **GEMS** reanalysis CAMS reanalysis Monitor [DU]

- Similar TCO3 analysis from (old) GEMS reanalysis and CAMS reanalysis
- Huge differences between corresponding O3 profiles
- No profile data (MIPAS, MLS) were assimilated in GEMSRA in Oct 2004 and model had a large O3 bias leading to very bad vertical O3 analysis profiles
- Shows importance of using limb sounding data for O3 analysis





# Aerosol analysis

Atmosphere Monitoring

- CAMS aerosol model has 14 aerosol bins:
  - 3 size bins each for sea-salt and desert dust
  - 2 bins (hydrophilic and hydrophobic) each for organic matter and black carbon
  - 1 bin for sulphate
  - 2 bins (fine and coarse) for nitrate
  - 1 bin for ammonium
- Assimilated observations are AOD at 550 nm from MODIS (Aqua and Terra) over land and ocean & PMAp (Metop-BC) over ocean
- Assimilation tests with VIIRS and SLSTR AOD
- Control variable is formulated in terms of the total aerosol mixing ratio.
- Analysis increments are repartitioned into the species according to their fractional contribution to the total aerosol mixing ratio.
- The repartitioning of the total aerosol mixing ratio increment into the different bins is difficult





# Dust storm February 2021



NASA Worldview – MODIS Aqua and Terra AOD 550nm observations for 20210222

The CAMS forecast does a good job of forecasting the AOD plume from Africa over Northern Europe Credit: Melanie Ades

## CAMS Total AOD at 550nm 12hr forecast valid at 20210222 12hr

Aerosol forecasts - Sunday 21 Feb 2021, 00 UTC VT Sunday 21 Feb 2021, 12 UTC Step 12 © ECMWF 2021



# Dust test case February 2021

.... A



- AOD increments are attributed to the ٠ different species according to their proportion in the nonlinear forecast.
- If there is no dust in the forecast in a ٠ specific location then the increment will be given to whatever species are there in this case Sulphate



AOD incr at 550nm

MWF

AOD at 550nm

# Dust test case February 2021



LMD IASI 10um obs 20210222 12hr



# SO2 assimilation

Atmosphere Monitoring

- Volcanic eruptions put ash and SO2 into the atmosphere
- Altitudes of volcanic plumes and emission strengths vary
- CAMS uses SO2 outgassing emissions but has no emission data set providing information about volcanic eruptions in NRT
- How can CAMS provide SO2 forecasts?



# Current use of SO2 data in CAMS NRT system

Atmosphere Monitoring

 CAMS assimilates GOME-2BC and TROPOMI total column SO2 TCSO2 retrievals making use of the volcanic flags provided by data providers (AC-SAF, ESA; algorithm from DLR)



Europe's eyes on Earth



THE EUROPEAN UNION



# Current use of SO2 data in CAMS NRT system

Atmosphere Monitoring

- CAMS assimilates GOME-2BC and TROPOMI total column SO2 TCSO2 retrievals making use of the volcanic flags provided by data providers (AC-SAF, ESA; algorithm from DLR)
- We need to make assumptions about the plume height if this is not known in NRT
- Default: SO2 is placed in mid-troposphere at model level 98 (~ 550 hPa, 5 km) by using a prescribed bg-error stdv profile
- This can be modified if injection height is known
- Currently: Globally constant injection height
- 'Baseline configuration: BLexp'



# Current use of SO2 data in CAMS NRT system

Atmosphere Monitoring

- CAMS assimilates GOME-2BC and TROPOMI total column SO2 TCSO2 retrievals making use of the volcanic flags provided by data providers (AC-SAF, ESA; algorithm from DLR)
- We need to make assumptions about the plume height if this is not known in NRT
- Default: SO2 is placed in mid-troposphere at model level 98 (~ 550 hPa, 5 km) by using a prescribed bg-error stdv profile
- This can be modified if injection height is known
- Currently: Globally constant injection height
- 'Baseline configuration: BLexp'

SO2 background error standard deviation



# Use of SO2 Layer Height data by CAMS

Atmosphere Monitoring

- DLR have developed algorithm to provide information about the plume height in NRT from TROPOMI (Hedelt et al., 2019, doi.org/10.5194/amt-12-5503-2019)
- Full-Physics Inverse Learning Machine (FP\_ILM) algorithm
- SO2 LH project one of ESA's S5P Innovation projects
- Data useful for SO2 > 20 DU
- CAMS is testing the use of these data: 'LHexp'
- Inness et al. (2022): https://doi.org/10.5194/gmd-15-971-2022





SO2 Layer height project



PROGRAMME OF THE EUROPEAN UNION



# Case Study: Raikoke eruption June 2019

Atmosphere Monitoring

- Raikoke volcano erupted on 22 June 2019
- SO2 and ash injected up to 13-17 km -> into stratosphere
- SO2 plumes could be observed from space for a long time in NH
- Compare BLexp and LHexp



# Height of SO2 plume from S5P LH data

Atmosphere Monitoring



# Comparison of CAMS plume height with IASI



Period: 22 -29 June 2019

CAMS SO2 analysis shows improved agreement with IASI LATMOS/ULB SO2 altitude data if TROPOMI SO2 LH data are used

**Biases against IASI:** 

BL exp: -5.1 ± 2.1 km LH exp: 0.4 ± 2.2 km

Using the LH data leads to improved SO2 analyses and SO2 forecasts

Europe's eyes on Earth



Inness et al. (2022): doi.org/10.5194/gmd-15-971-2022

# Hunga Tonga eruption January 2022

### **Atmosphere**

Miles

- Monitoring Eruption of Hunga Tonga-Hunga Ha'apai volcano in January 2022 caused atmospheric shock waves, sonic booms, and tsunami waves
  - It injected SO2, ash and water vapour into the stratosphere



Hunga Tonga eruption



Credits: NASA's Goddard Space Flight Center/Mary Pat Hrybyk-Keith



# CALIPSO 19 January 2022

### source Simon Carn on twitter:

80.58

https://twitter.com/simoncarn/status/1484527916960595973?s=20)

86.17

Europe's eyes on Earth

# Beginning of eruption 14 January 2022, Oz



### Animation of CAMS SO2 analysis **Atmosphere** CAMS TCSO2 CAMS sulphate aerosol CAMS Total Column Sulphur Dioxide: 20220113, 00z CAMS Sulphate Aerosol Optical Depth: 20220113, 00z 10°N 10°N 0° - 01 10°S 10°S 20°S 20°S 30°S 30°S 40°S 40⁰S 50°S 50°S 120°E 180 120°E 0.00 [DU] 0.00 5,00 10,00 0,50 1.00 Credit: Mark Parrington

- Assimilation brings the volcanic SO2 signal into the analysis
- SO2 is converted to sulphate aerosols
- Initial plume is advected south-eastwards because SO2 is placed too low, i.e. in the mid-troposphere
- On subsequent days the plume is moved further westwards by adjusting initial conditions (while still at the wrong altitude)







Europe's eyes on Earth

# 4. Reanalysis



20









# Near-real time model versus reanalysis

Atmosphere Monitoring



CAMS Reanalysis NRT CAMS analysis **NRT global CAMS system** (daily analyses and 5-day forecasts):

- Evolves with time: Usually 1-2 model updates per year
- Horizontal and vertical resolution can change
- Observation usage changes
- Emission data sets might change

# **Reanalysis** (retrospective):

- Consistent long term dataset produced with one model version (from 2003 onwards)
- Consistent emissions
- Consistent, reprocessed observations
- Can be used for trend analysis



PROGRAMME OF THE EUROPEAN UNION

Europe's eyes on Earth



# CAMS global reanalysis 2003 – 2021 (updated every 6 months)



# CAMS global reanalysis (CAMSRA, eac4)

- 2003 –2021, with new years being added
- Aerosols, chemical pollutants, CO<sub>2</sub> & CH<sub>4</sub>
- 80 km spatial resolution
- Inness et al. (2019): <u>https://doi.org/10.5194/acp-19-3515-2019</u>
- Wagner et al. (2021): <u>https://doi.org/10.1525/elementa.2020.00171</u>
- <u>atmosphere.copernicus.eu/eqa-reports-global-services</u>
- Available from ADS <u>https://atmosphere.copernicus.eu/data</u>



### Reanalysis

Using a combination of observations and computer models to recreate historical climate conditions.

DATA DESCRIPTION		
Data type	Gridded	
Horizontal coverage	Global	
Horizontal resolution	0.75°x0.75°	
/ertical coverage	Surface, total column, model levels and pressure levels.	
/ertical resolution	60 model levels. Pressure levels: 1000, 950, 925, 900, 850	
Femporal coverage	2003 to 2020	
Temporal resolution	3-hourly	
File format	GRIB (optional conversion to netCDF)	
/ersions	Only one version	
Update frequency	Twice a year with 4-6 month delay	

- CB05 tropospheric chemistry
- Cariolle-Déqué scheme for stratospheric ozone
- Interactive prognostic O3 and AER

# Use EAC4 to study CO trends

CO burdens

**Atmosphere** 





Fig. 2.62. Column-averaged CO (xCO, in ppb) at the Park Falls TCCON station. Monthly mean observations are shown by the black dots, and corresponding monthly mean xCO columns calculated using the TCCON-averaging kernels are shown by the blue triangles. The continuous blue line is the monthly xCO from the CAMS reanalysis.

Flemming et al. (2020), BAMS State of Climate 2019

TCCON data from: https://tccondata.org/

(ab) Carbon Monoxide



PROGRAMME OF THE EUROPEAN UNION



Europe's eyes on Earth



-20 -15 -10 -5 0 5 10 15 20 Anomalies from 2003-21 (%) Flemming and Inness (2022), BAMS State of Climate 2021

TCCO Anomaly 2021

# Antarctic October TCO3 trends

1979-2016

# 1979-2020



Linear trends calculated for periods 1979-1996, 1997-2020

CCI merged data set and Multi Sensor Reanalysis from: <u>cds.climate.copernicus.eu</u>

THE EUROPEAN UNION

Europe's eyes on Earth

Good agreement

1979-2016

**Atmosphere** 

 Trend depends on turnaround year and length of dataset

# Antarctic ozone hole 2019, 2020 & 2021

In addition to long-term O3 recovery there is a lot of interannual variability



**Atmosphere** 





- 2019 and 2020 both had exceptional Antarctic ozone holes
- 2019 small and short-lived because of unusual stratospheric warming
- 2020 deep, big & long-lived due to very cold stratosphere and stable polar vortex
- 2021 very similar to 2020

PROGRAMME OF

(1979-2002 from ERA5; 2003-2020 from CAMSRA; 2021 CAMS NRT)



yes on Earth

# O3 hole 2022



[DU]

] (provided by CAMS, the Copernicus Atmosphere Monitoring

• It will be interesting to see how it evolves

# Summary

Atmosphere Monitoring

- 4D-Var data assimilation methodology, including the importance of emission inversion
- Examples of challenges or limitations when assimilating atmospheric composition observations
  - Bias correction
  - Observation operators and application of averaging kernels
  - Issues when assimilating total column observations
  - Aerosol assimilation
  - SO2 assimilation
- Reanalysis and examples of data usage, e.g. trends, Antarctic Ozone hole
- CAMS provides atmospheric composition data at global and European regional scale
- CAMS data freely available from ADS https://atmosphere.copernicus.eu/data



Europe's eyes on Earth





# The Atmosphere Data Store (ADS)

### Atmosphere Monitoring All CAMS data are freely available

# https://atmosphere.copernicus.eu/data

Home Search Datasets FAQ	ernicus CECMWF Atmosphere Monitoring Service Vour reedback helps us to improve	register
Atmosphere Data	Store	Home
10		Sear
	Welcome to the Atmosphere Data Store	cams
lysis monthly mean of total column carbon monoxide	It is freely available and functions as a one-stop shop to explore Atmosphere data. Register for free to obtain access to the ADS and its Toolbox.	Sort by Rele
A CONTRACTOR	We are constantly improving the services and adding new datasets. For more information, please consult the catalogue, our FAQ $\omega$ or the CAMS forum $\omega$ .	Title Type
	Enter search term(s) Al 💙 Search	> Varia
V Vo Sy		> Para
4 10 4 10 10 10 10 10 10 10 10 10 10 10 10 10		> Prod
		> Tem
	Atmosphere Data Store API Arress the CAMS Envirm Arress the CAMS website	

	icus
Search Datasets FAQ @	
rch results	
s reanalysis Q	All
y evancy	Showing 1-7 of 7 results for cams reanalysis *
2	CAMS global reanalysis (EAC4) monthly averaged fields CAMS global reanalysis (EAC4) monthly averaged fields
iable domain	Section CAMS global reanalysis (FAC4)
ameter family	CAMS global reanalysis (EAC4)
atial coverage	About CAMS
duct type	Copernicus Atmosphere Monitoring Service The Copernicus Atmosphere Monitoring Service (CAMS
nporal coverage	Second State CAMS solar radiation time-series
	CAMS solar radiation time-series
	CAMS European air quality forecasts
	CAMS European air quality forecasts

# http://atmosphere.copernicus.eu

# @CopernicusECMWF







