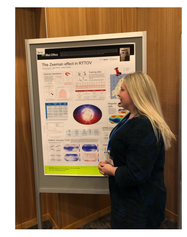
A diverse atmospheric dataset of 100,000 profiles from CAMS short range forecasts

Emma Turner (emma.turner@ecmwf.int), Angela Benedetti, Oakley Young*

Research Department, ECMWF, Reading, United Kingdom, *Met Office, Exeter, United Kingdom



Selection Algorithm

Most (90%) of the profiles are selected randomly to preserve the statistical qualities of the original model. 10% (both steps) of the remaining profiles are selected via the algorithm shown below, such that inter-profile differences are maximised following Eresmaa and McNally (2014).

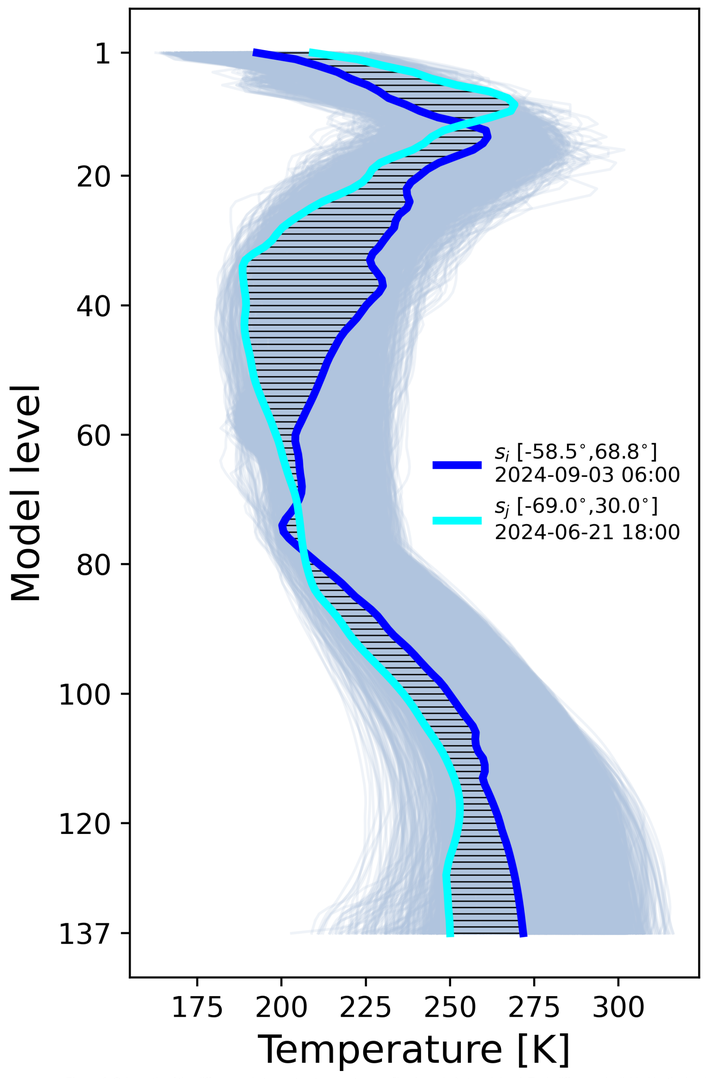


Figure 1. Schematic of the differences between two temperature profiles (s_i and s_j) at each level. Each difference is divided by the level standard deviation, the quantity is then squared and summed over every level to give D, a measure of the inter-profile difference.

Step 1

For 12 months (March 2024 – February 2025):
  For 3 days (1st, 10th, 20th):
    For 4 forecast steps (36, 42, 48, 56):
      1350 of 348,528 profiles selected at random
      150 of 347,178 profiles selected via a tunable threshold

Number of bins K; Number of levels L; Variable in bin k from profile i on level l; Profile i; Standard deviation of variable k on level l

$$D(s_i, s_j) = \sum_{k=1}^{K} \sum_{l=1}^{L} \left(\frac{\theta_{ik}(l) - \theta_{jk}(l)}{\sigma_k(l)} \right)^2$$

$D(s_i, s_j) > threshold \ \forall \ s_j \in S_o$ ← Set of output profiles

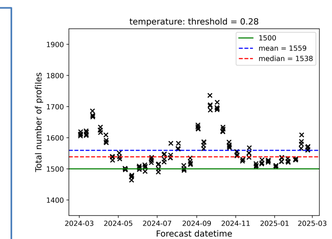
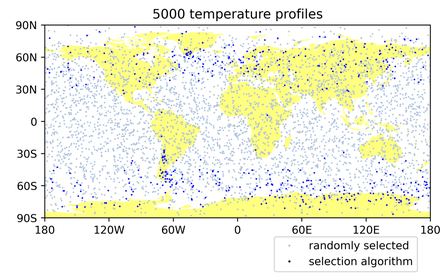


Figure 2. Number of profiles selected at each time step using a threshold of 0.28 in step 1 of the selection algorithm

Step 2

4500 of ~216,000 profiles selected at random
500 of ~211,500 profiles selected by tuning a threshold



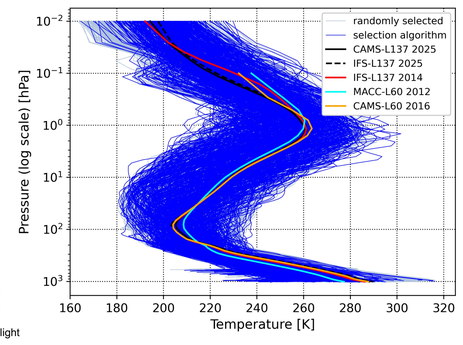


Figure 3. (above) Location of all profiles in the temperature subset of the CAMS-L137 2025 dataset. The 500 profiles chosen via the selection algorithm are shown with dark blue pixels and the remaining 4500 chosen by random are indicated with light grey blue pixels.
Figure 4. (right) All temperature profiles converted from model to pressure levels. Blue and light blue profiles as above. Medians are shown for the respective current and previous datasets.

CAMS and the IFS

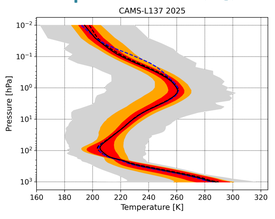
	CAMS AC	IFS
Horizontal resolution	T511/N256 (~40 km) 348,528 pixels	T1279/O1280 (~9 km) 6,599,680 pixels
Vertical resolution	137 levels (~80 km) 1013.25 - 0.01 hPa	
Forecasts	0000/1200 to 5 days	0000/1200 to 15 days 0600/1800 to 6 days
Cycle	Cy49r1 (12th November 2024) Cy48r1 (27th June 2023)	

20 variables

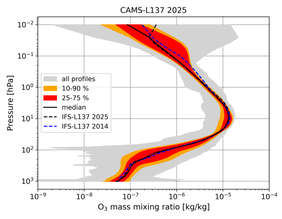
Long name	Short name	bins
Temperature	temp	1
Specific humidity	hum	1
Ozone	oz	1
Cloud condensate	clw, ciw	2
Precipitation	rain, snow	2
Carbon dioxide	co2	1
Methane	ch4	1
Nitrous oxide	n2o	1
Carbon monoxide	co	1
Nitrogen dioxide	no2	1
Sulphur dioxide	so2	1
Formaldehyde	ch2o	1
Sea salt aerosol	salt1, salt2, salt3	3
Desert dust aerosol	dust1, dust2, dust3	3
Sulphate aerosol	sulphate	1
Organic matter	hphil_om, hphob_om	2
Black carbon	hphil_bc, hphob_bc	2
Ammonium	ammonium	1
Nitrate aerosol	nitrate1, nitrate2	2
Secondary organic matter	bio_om, anthr_om	2

NWP variables

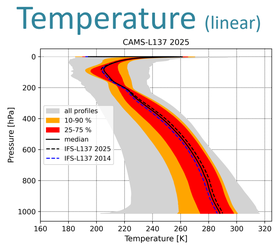
Temperature (log)



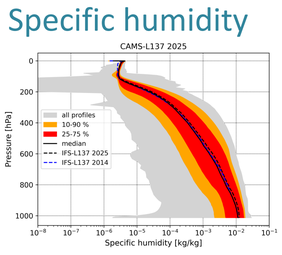
Ozone



Temperature (linear)

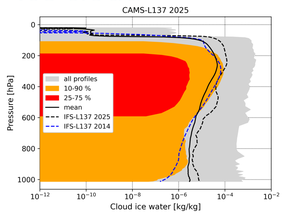


Specific humidity

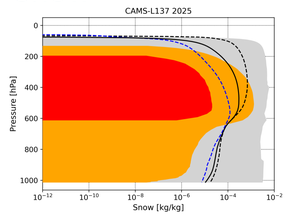


Hydrometeors

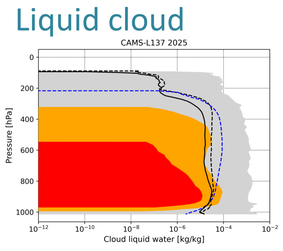
Ice cloud



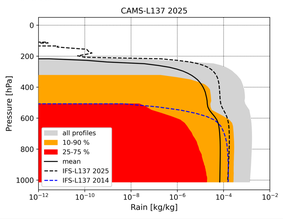
Snow



Liquid cloud

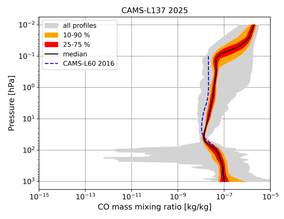


Rain

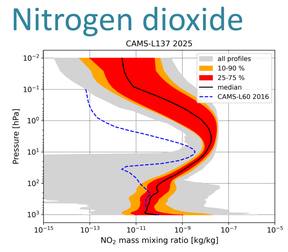


Atmospheric gases

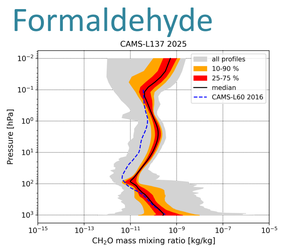
Carbon monoxide



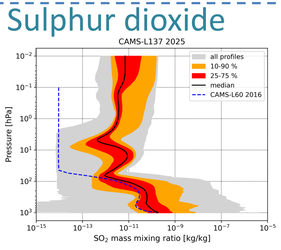
Nitrogen dioxide



Formaldehyde



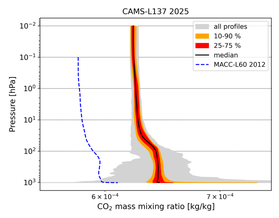
Sulphur dioxide



Greenhouse gases

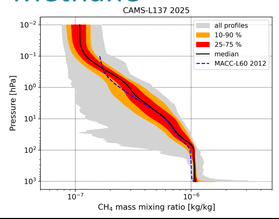
CAMS GHG model is used for CO_2 and CH_4

Carbon dioxide

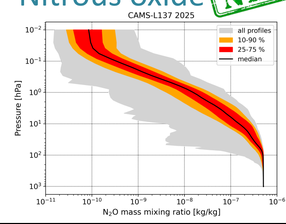


	CAMS GHG
Horizontal resolution	T1279/O1280 (~9 km) 6,599,680 pixels
Vertical resolution	137 levels (~80 km) 1013.25 - 0.01 hPa
Forecasts	0000 UTC to 5 days
Cycle	Cy49r1 (12th Nov 24)

Methane



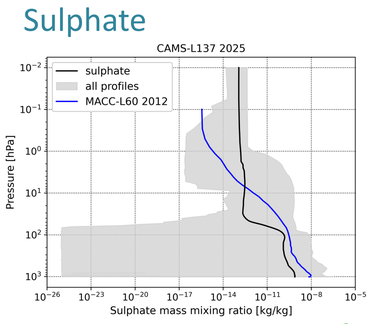
Nitrous oxide NEW



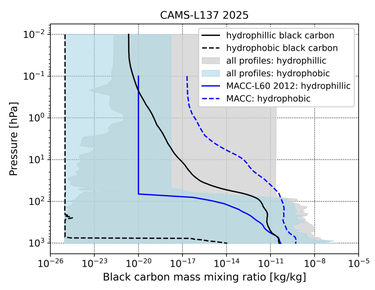
Aerosols

New aerosol optical properties from CAMS have also been added to RTTOV for release in version 14.1. All aerosols are treated as Mie spheres apart from dust.

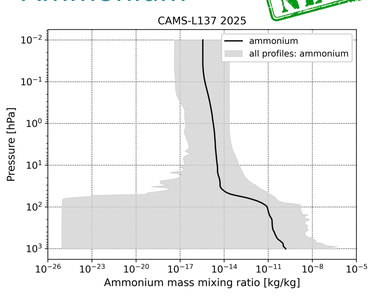
Sulphate



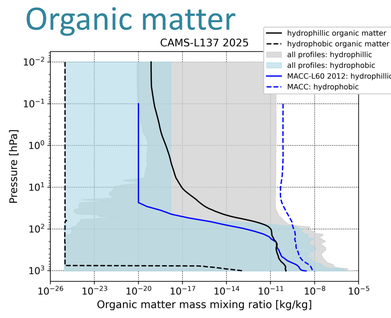
Black carbon



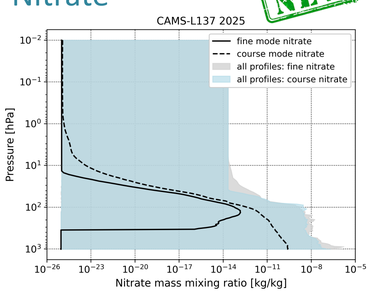
Ammonium NEW



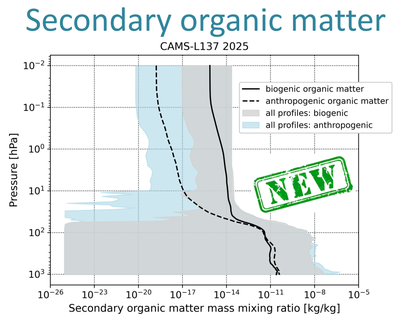
Organic matter



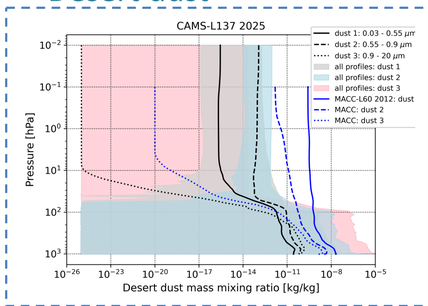
Nitrate NEW



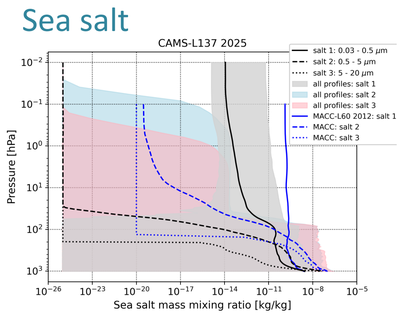
Secondary organic matter



Desert dust

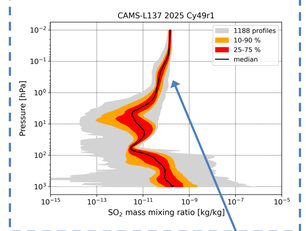


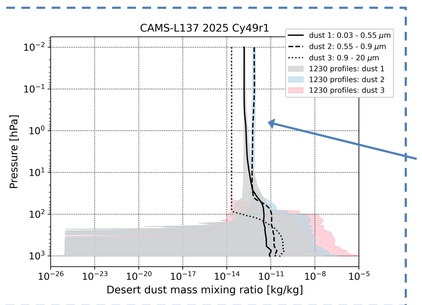
Sea salt



Cy49r1

Cy49r1 represents a major upgrade in CAMS for aerosols (and SO_2). There is an increase in AOD globally.





The latter ¼ of the new dataset is from Cy49r1. Many concentrations increase and plumes are narrower at upper levels.

Key Points

- The new dataset will be available at: https://nwp-saf.eumetsat.int/site/software/atmospheric-profile-data/
- Atmospheric gas and aerosol profiles deviate considerably from previous datasets
- N_2O, ammonium, nitrate and biogenic and anthropogenic organic matter are new.
- Hydrometeor variables are affected by the lower resolution of CAMS with respect to the IFS.
- Cy49r1 comprises ~¼ of the profiles and has a significant effect on aerosols and SO_2, particularly at upper levels.

References

Eresmaa, R., and A. McNally, (2016): Diverse profile datasets based on the CAMS atmospheric composition forecasting system. NWP SAF Report, 11 p.

Eresmaa, R., and A. McNally, (2014): Diverse profile datasets from the ECMWF 137-level short-range forecasts. NWP SAF Report No. NWPSAF-EC-TR-017, 12 p.

Eresmaa, R., A. Benedetti, and A. McNally, (2012): Diverse profile database of aerosol and trace gas concentrations from the Monitoring Atmospheric Composition and Climate short-range forecasts. NWP SAF Report No. NWPSAF-EC-TR-015, 12 p.