AeroCom aircraft comparison experiment

Duncan Watson-Parris, University of Oxford ([duncan.watson-parris@physics.ox.ac.uk](mailto:duncan.watson-parris@physics.ox.ac.uk))

*Aircraft participants*

Sarah Doherty (University of Washington), Jens Redemann (NASA Ames), Shuka Schwarz (NOAA), Mian Chin (NASA Goddard), Christina Williamson (NOAA), Paola Formenti (CNRS), Rob Wood (University of Washington), Andi Anreae (MPI-C, Mainz), Markus Hermann (IAGOS CARIBIC), Tom Lachlan-Cope (BAS), Ken Carslaw (University of Leeds)

*Model participants*

Dirk Olivie (MetNo), Bjørn H. Samset (CICERO), Gunnar Myhre (CICERO), Haochi Che (University of Oxford), Laurent Labbouz (Toulouse), Jialei Zhu (University of Michigan), Joyce Penner (University of Michigan), Huisheng Bian (NASA Goddard), Andrew Gettleman (NCAR), Philip Stier (University of Oxford)

*We propose an experiment in which any participating AeroCom model can provide a common set of diagnostics on a consolidated set of aircraft sampling points in order to facilitate the large number of anticipated analyses leveraging recent aircraft data efforts.*

Recent dedicated aircraft measurement campaigns and data collection efforts have delivered a large amount of in-situ aerosol measurements of great value to AeroCom modellers. The Global Aerosol Synthesis and Science Project (GASSP) dataset[1] brings 1000s of separate aircraft measurement flights across 10s of campaigns into a single consistent database. Combining this with data from recent campaigns such as CLARIFY, ORACLES[2], ATom[3] and ACE-ENA provides a unique opportunity to evaluate AeroCom model aerosol distributions across a wide range of regions and meteorological conditions. Each campaign includes different measurements of aerosol properties such as size distributions and speciation, and each focuses on different regions or phenomena; however, they all provide valuable model constraints and all require similar sampling considerations.

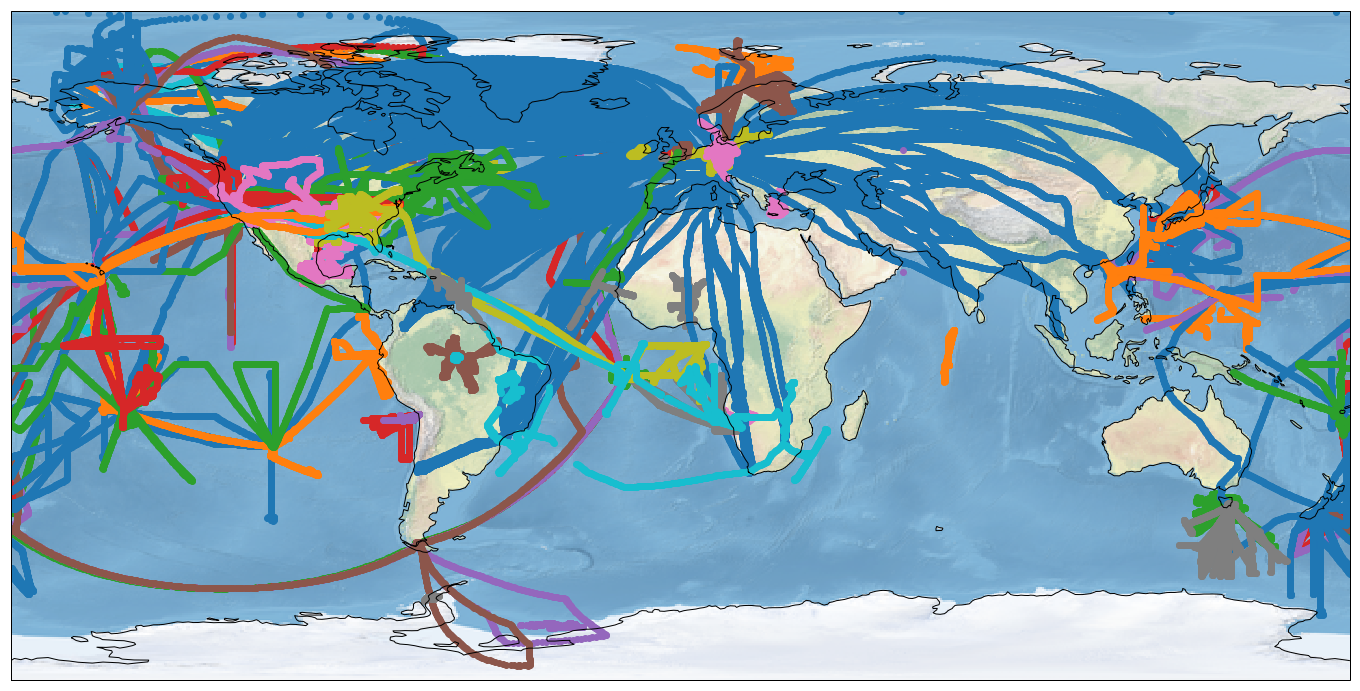


Figure 1: All aircraft flight data across a number of different campaigns and databases

Building on the Phase II experiments this effort will support the interpolation of consolidated flight track points from high-temporal resolution model output to minimise the large sampling biases that would otherwise be present[4]. An indication of the biases that can be introduced otherwise is given in Figure 2, which shows Taylor diagrams for a single model field sampled at the same points but from different temporal resolution output. Using monthly mean model output can result in correlations of less than 0.7 and biases larger than 15% compared to on-line interpolation.

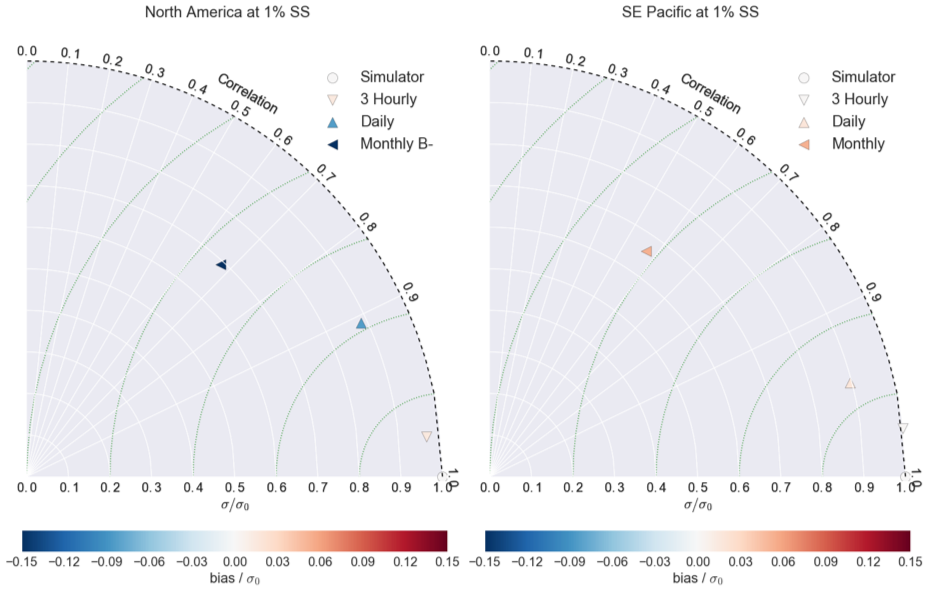


Figure 2: Taylor diagrams showing the sampling error and reduction in correlation and variability introduced by comparing instantaneous point observations with different temporal aggregations of model output. All datasets are linearly interpolated from the same underlying 4-D model field (CCN at 1% supersaturation) and compared with an online interpolation. (a) shows points taken over North America, (b) shows points over the South East Pacific.

For this experiment the flight track points will be provided in a single CF-conformant NetCDF format with time, latitude, longitude, altitude and pressure coordinates. A post-processing script will also be provided allowing interpolation from high-temporal resolution output (at least 3 hourly) using the CIS tool[5] to output in the same CF-compliant NetCDF format as the sample data, and then deletion of the full output fields. Vertical interpolation will automatically be performed by height or pressure as required.

Some models have implemented a ‘flight-track simulator’ to allow on-line interpolation of these spatially sparse measurement points, thus avoiding significant output storage requirements. These implementations can be used where they allow the interpolation of around 106 points, however conversion between the data formats supplied and received must be done by the modelling group. Strict controls on the data format will be enforced to ensure that each models output can be used easily by all of the aircraft groups.

Due to the (potentially) large data processing required we have split the experiment in to two Tiers. The first Tier is mandatory for inclusion in the main analysis, while the second Tier provides additional investigation opportunities but are not mandatory. Both setups are identical to the Phase III control experiment, the only differences are the period. Modelling groups are free to choose the most appropriates biomass burning emissions inventories to use. Ideally these should be the same as those used for the CRTL experiment.

## Tier 1 – 2017 only

This one-year experiment will support the main analyses. Evaluations against all campaigns will be done assuming the inter-annual variability is small in remote regions (previously shown in ECHAM-HAM and to be further tested in Tier 2 experiments). This particular year, however, will allow direct comparison with the CLARIFY, ORACLES-2017 and ATom-3 campaigns where the biomass South-East Atlantic burning plume shows large inter-annual variability.

## Tier 2 – Hindcast (optional)

A full hindcast is requested to run for 2008 through to 2018. This will allow for comparisons in the period with the most amount of flight data (2008), during the control run (in 2010) and against data in the most recent available campaigns (ATom-4, ACE-ENA and SOCRATES in 2018). It will explore the inter-annual variability of remote aerosol and the assumptions made in the Tier 1 experiment.

## Tier 2 – Pre-Industrial (optional)

We also (optionally) request a pre-industrial run (1850) to investigate natural aerosol and how representative remote campaigns such as ATom and SOCRATES are of ‘pristine’ conditions.

The interpolated model fields requested are listed in the accompanying spreadsheet. Files should adhere to the normal AEROCOM rules: one NetCDF file per variable, CF-compliant. Both the variable and the coordinate fields should share a single dimension (preferably named ‘obs’)[6]. This format will automatically be adhered to by the provided script.

**Note** that 3 months isn't long enough to spin up the accumulation mode in the Upper Troposphere in ECHAM-HAM. Please use a 6-month spin-up - unless you're sure that your model doesn't require it.

# Linked studies and analyses

The requested model setup and diagnostics will inform a number of different analyses proposed by both modellers and aircraft groups. Examples of these different efforts and how each aircraft dataset contributes to them are demonstrated in Figure 3.

Some analyses build on the baseline experiment with their own sensitivity experiments or specialist diagnostics. Please refer to each extension analysis for further details.



Figure 3: Schematic of different experiments and campaigns included in this experiment. This is not an exclusive list and many more analyses are anticipated and encouraged

# Timeline

~~Summer 2019 – Finalised the experiment protocol (Done)~~

~~Autumn 2019 – Received initial model submissions~~

~~Winter 2019 – Early analysis performed~~

(Spring – Updated protocol)

**Summer 2020 – Final deadline for model submission**

Autumn 2020 -> Publication

# Submission requirements

To be used in these analyses each model submission must include as a *minimum* the Priority ‘1’ fields requested in the attached spreadsheet. Priority ‘2’ fields could nonetheless greatly enhance some the analyses and are requested if available.

Each model must also strictly follow the NetCDF-CF conventions for file structure and metadata – particularly with regard to the units and vertical coordinates. File naming should follow the standard convention:

aerocom3\_<ModelName>\_<ExperimentName>\_<VariableName>\_<VerticalCoordi nateType>\_<Period>\_<Frequency>.nc,

But note that the vertical coordinate type should be ‘Aircraft’ for interpolated spatio-temporal points. e.g:

aerocom3\_ECHAM6.3HAM2.3\_ GASSP\_CCN0040\_Aircraft\_2008\_3hourly.nc

Submissions not adhering to these requirements will be not be accepted. Storage requirements for the diagnostics are minimal. All the requested (and optional) diagnostics output at every point requires less than 1Gb output.

# References

[1] Reddington, C.L., et al. (2017).  THE GLOBAL AEROSOL SYNTHESIS AND SCIENCE PROJECT (GASSP): Measurements and modelling to reduce uncertainty. *Bull. Amer. Meteor. Soc.,* **0**, <https://doi.org/10.1175/BAMS-D-15-00317.1>

[2] <https://espo.nasa.gov/ORACLES>

[3] Wofsy, S.C., et al., *ATom: Merged Atmospheric Chemistry, Trace Gases, and Aerosols*. 2018, ORNL Distributed Active Archive Center.

[4] Schutgens, N. A. J., Partridge, D. G., & Stier, P. (2016). The importance of temporal collocation for the evaluation of aerosol models with observations. *Atmos. Chem. Phys.*, *16*(2), 1065–1079. http://doi.org/10.5194/acp-16-1065-2016

[5] Watson-Parris, D., et al. (2016). Community Intercomparison Suite (CIS) v1.3.2: A tool for intercomparing models and observations. *Geosci. Model Dev., 9,* 3093-3110, https://doi.org/10.5194/gmd-9-3093-2016, 2016.

[6] http://cfconventions.org/Data/cf-conventions/cf-conventions-1.7/cf-conventions.html#point-data