National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

VISION: towards seamless integration of Model, Satellite, and In-Situ Observation data

Sadie Bartholomew (NCAS-CMS & UoR), on behalf of the VISION team

Dept. Meteorology Lunchtime Seminar, 2025-04-29

UNIVERSITY OF CAMBRIDGE | Yusuf Hamied Department of Chemistry

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

FAAM

UNIVERSITY OF LEEDS

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

- Context: VISION project

UNIVERSITY OF CAMBRIDGE | Yusuf Hamied Department of Chemistry

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

FAAM

UNIVERSITY OF LEEDS

Virtual **I**ntegration of **S**atellite & **I**n-Situ **O**bservation **N**etworks

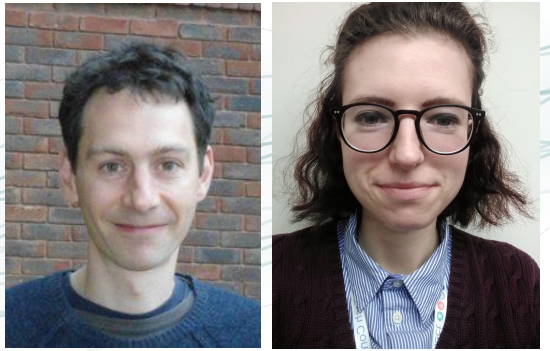
- **Pre-2024**: project initiated with work by Maria Russo (NCAS + Cambridge) and team to produce a tool '`ISO_simulator`' (i.e. Toolkit Version 1) which reads modelled variables and limited (mainly flight) observational data files and outputs model data interpolated in space and time to match the observations, as detailed in Russo. et al, 2025: https://gmd.copernicus.org/articles/18/181/2025/gmd-18-181-2025.html
- **Jan 2024 - March 2025**: funding grant from NERC **TWINE** (**TWIN**ning capability for the natural **E**nvironment) programme 'Digital Twins for Environmental Science' for cross-NCAS/NCEO team of atmospheric modellers, software engineers, and satellite experts to deliver new, extended toolkit and a digital twin of FAAM. 15 months funding, £622K. Lead: Luke Abraham.
- **Going forwards**: maintenance of tool by NCAS-CMS, with science guidance from rest of team. Potential routes for new funding to explore new applications of the toolkit and develop it further.

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Introducing the VISION team

NCAS Cambridge

NCAS Leeds



NCAS Reading

NCEO RAL STFC

FAAM (project partner)

Introducing the VISION team

NCAS Cambridge

NCAS Leeds

NCAS-CMS @UoR

David Hassell Sadie Bartholomew

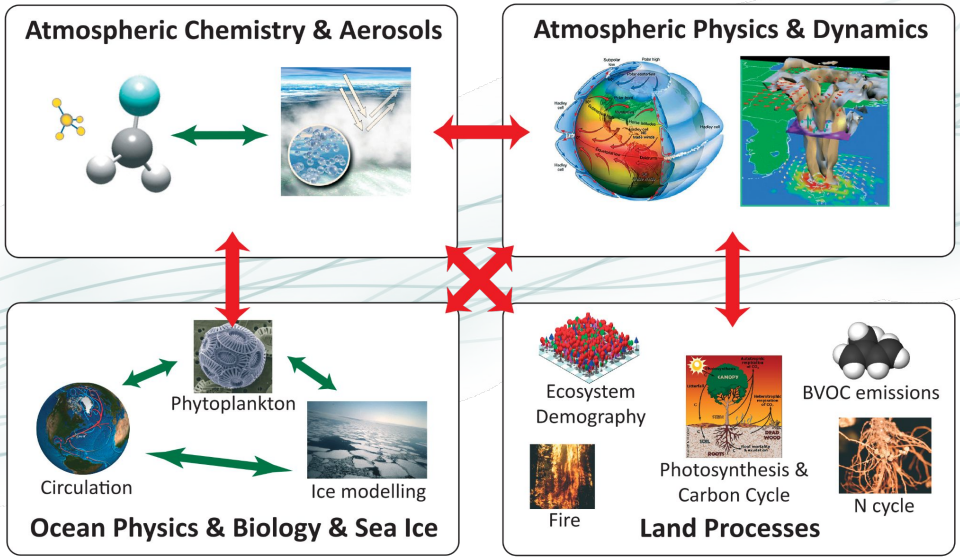
NCEO RAL STFC

FAAM (project partner)

VISION Project Aims

1. *Reduce NERC carbon emissions* by optimising impact and outcomes from the FAAM Airborne Laboratory.

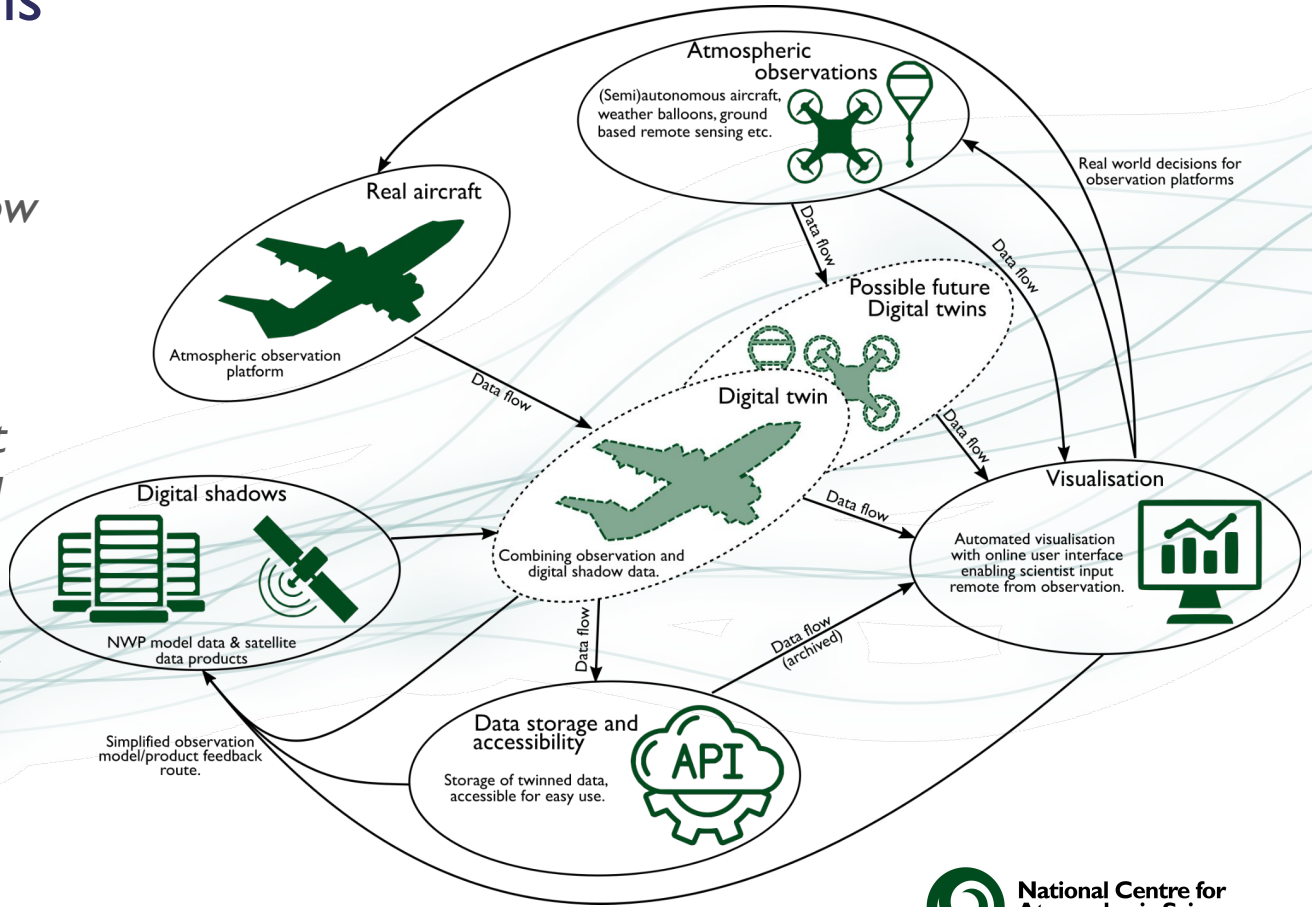
2. Improve policy decision making for climate mitigation by *providing tools for integrating right-time observational data into Earth System Models.*



National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

VISION Project Aims

*"This project will **deliver a toolkit and novel visualisations** that will allow for better integration of models and observations to enhance our confidence in future climate projections. It will **also develop a digital twin** to improve the operational flights of the FAAM Airborne Laboratory atmospheric research aircraft."**



* From https://www.ukri.org/news/digital-twin-projects-to-transform-environmental-science/

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

- Context: VISION Toolkit

UNIVERSITY OF CAMBRIDGE | Yusuf Hamied Department of Chemistry

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

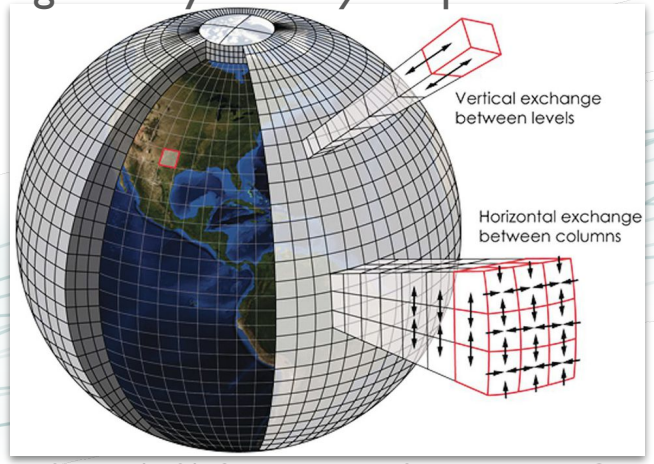
FAAM

UNIVERSITY OF LEEDS

Simulations and measurements as sources of earth system info.

(Earth System) Models (ESMs)

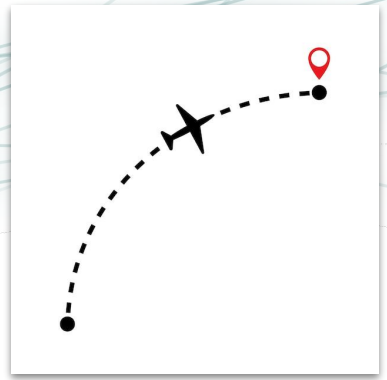
- spatially *gridded* output, with cells typically ~100 km wide
- regular and much *lower temporal resolution*, commonly hourly
- generally *orderly* output

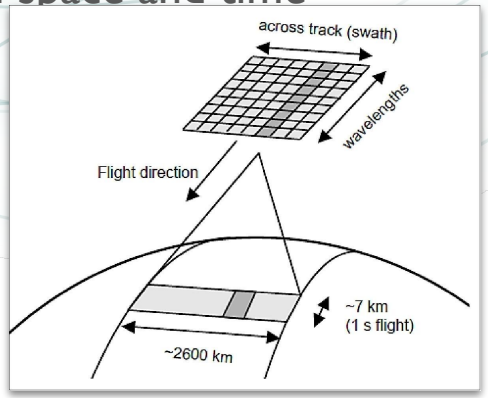


vs.

Observations

- spatially *discrete points* as a set forming trajectories or more complicated geometries (e.g. satellite swaths), much *less structured*
- *higher temporal resolution* (possibly irregular) e.g. can be sampled every second
- *variable coverage* in space and time





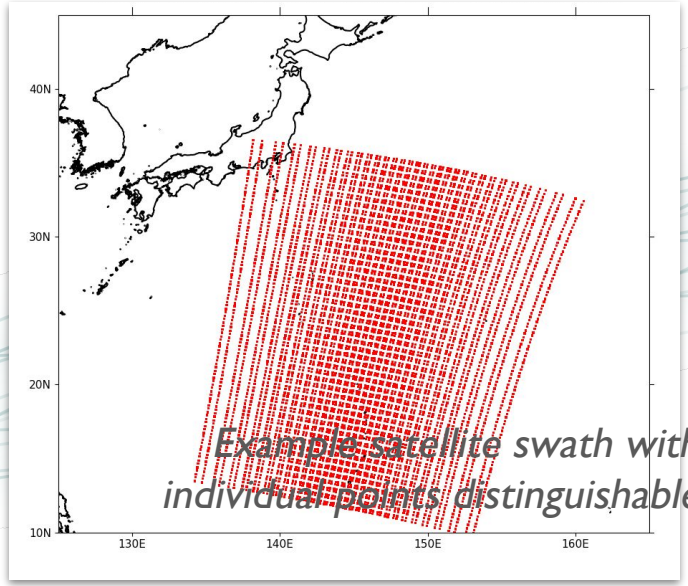
Images from: (left) Schematic of a Discrete Climate Model Grid, from Earth Magazine and Kotamarthi et al., (right) KNMI via Copernicus: Sentinel-5P Satellite Missions Catalogue

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

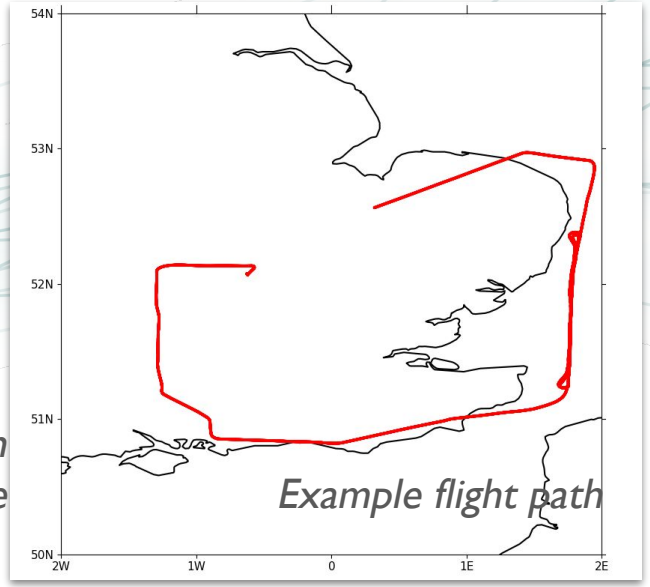
The challenge of integrating models and observations

- To account for the discrete nature relative to a gridded output and higher-temporal resolution of the observations to the models, we **need to interpolate** to compare values for points on observational paths or swaths (or other geometries), to **co-locate**
- Huge volumes of data are produced in both cases, so comparison is data-intensive with storage and processing concerns as well as the mathematical/scientific challenge

Images: generated with cf-plot



Example satellite swath with individual points distinguishable



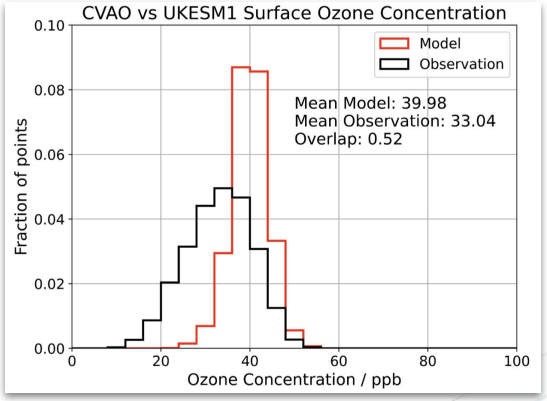
Example flight path

- *In-situ* = "in the original place"
- Observations VISION can support include in-situ cases e.g. research flights and ships, buoys, argo plus remote sensing e.g. satellites

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

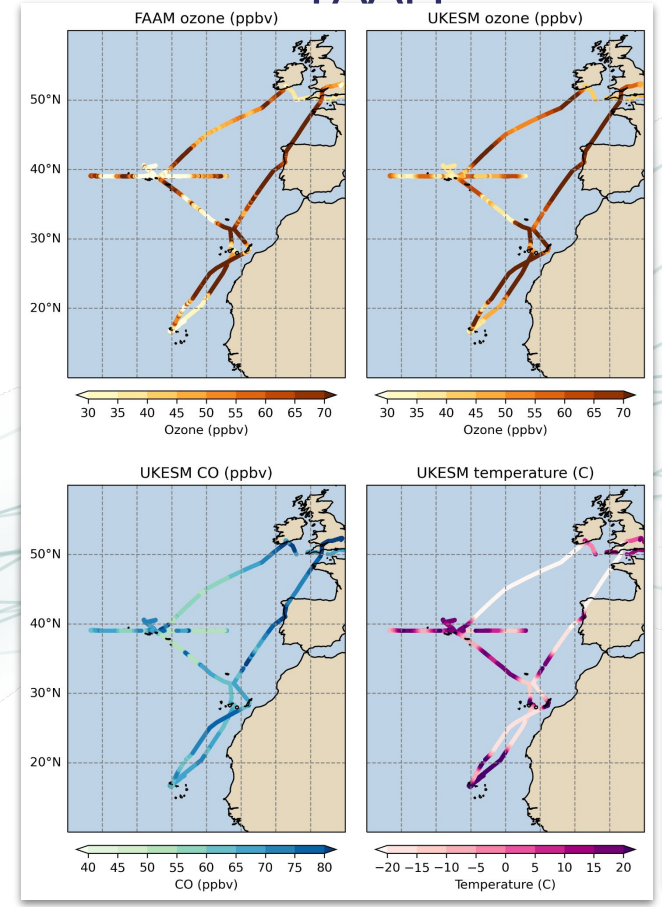
Pre-TWINE work: ISO Simulator i.e. Toolkit Version 1.0*

Cape Verde

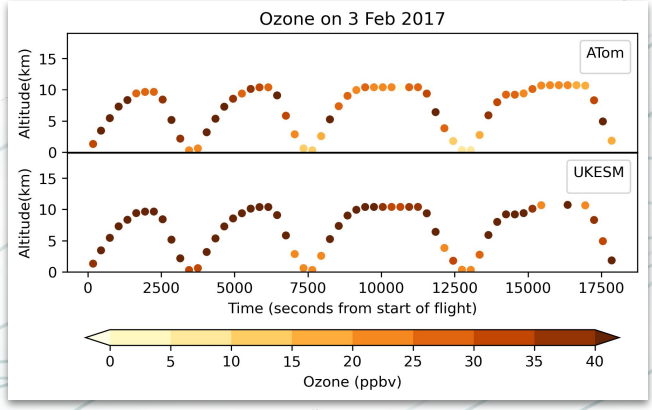


*"Virtual Integration of Satellite and In-situ Observation Networks (VISION) v1.0: In-Situ Observations Simulator". *Russo et al., GMD* (**2025**)

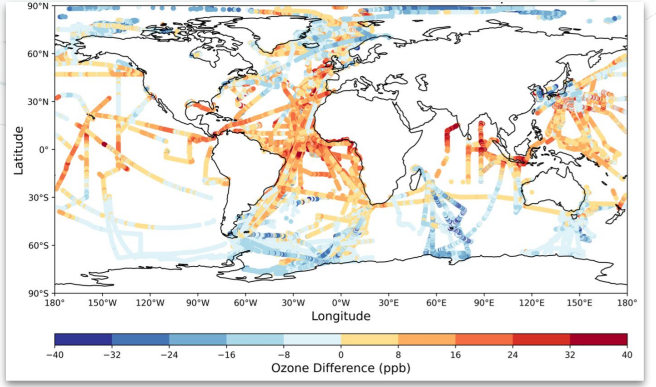
FAAM



ATom

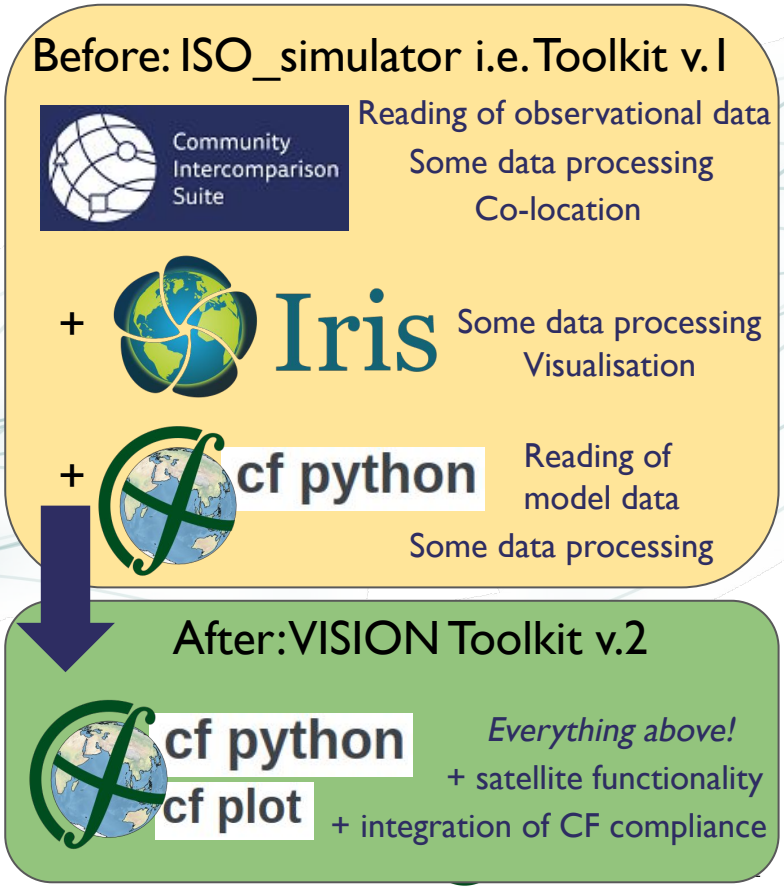


TOAR2: ships + buoys



How/why to improve on the ISO Simulator i.e. Toolkit v.1

- Very difficult to (get the required environment to) install due to depending on a library which is no longer maintained, CIS, causing environment conflict issues
- Uses a mix of libraries with overlapping functionality: CIS, Iris and cf-python
- Lacks satellite capability! Remote sensing application was identified early on as a way to greatly enhance the usefulness of the toolkit, so would be a major enhancement
- Didn't make use of metadata standards to streamline the data processing



National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

▸ VISION Toolkit Version 2

UNIVERSITY OF CAMBRIDGE | Yusuf Hamied Department of Chemistry

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

FAAM

UNIVERSITY OF LEEDS

Toolkit v.2: accessible, user-friendly, performant and stable

Climate and Forecasting Conventions

+

Python

=

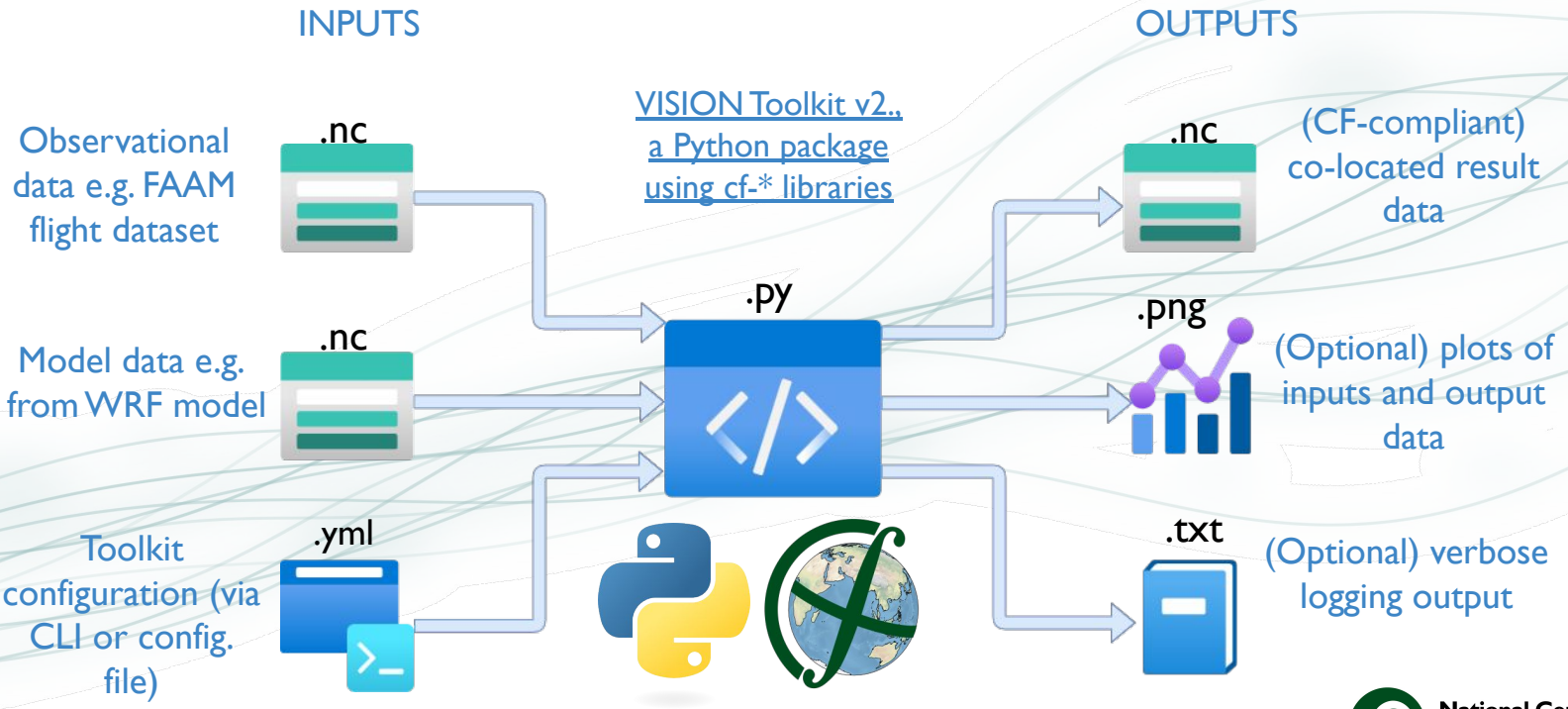
cf python (data analysis) + cf-plot (data visualisation)

- An **accessible** software package that is **easy and fast to use**, and is **portable** so can be used by anyone, with *any CF-compliant* model and observational dataset.
- Written in python with **cf-python** and **cf-plot**, data analysis and visualisation libraries developed and maintained within NCAS.
- Relies on standardised CF-compliant data for both model and observational inputs. We are providing **CF-netCDF usage guidance** to clarify this. Any ESM with sufficiently CF-compliant output is supported by the tool e.g. UM, WRF

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Toolkit v.2 as a Python package and command-line tool

Basic usage: run `$ visiontoolkit --config-file="<file path>"`



National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

The command-line interface

Customise run with `$ visiontoolkit <cli arguments and options>`

```
usage: visiontoolkit [-h] [-v] [-c CONFIG_FILE] [-p PREPROCESS_MODE_OBS]
                     [--preprocess-mode-model PREPROCESS_MODE_MODEL] [--orography OROGRAPHY]
                     [-s START_TIME_OVERRIDE] [-o OBS_DATA_PATH] [-m MODEL_DATA_PATH]
                     [--chosen-obs-field CHOSEN_OBS_FIELD] [--chosen-model-field CHOSEN_MODEL_FIELD]
                     [--skip-all-plotting] [-d OUTPUTS_DIR] [-f OUTPUT_FILE_NAME]
                     [--history-message HISTORY_MESSAGE] [--halo-size HALO_SIZE]
                     [-i SPATIAL_COLOCATION_METHOD] [-z VERTICAL_COLOCATION_COORD]
                     [--source-axes SOURCE_AXES] [--plotname-start PLOTNAME_START]
                     [--show-plot-of-input-obs] [-t] [--cfp-cscale CFP_CSCALE]
                     [--cfp-mapset-config CFP_MAPSET_CONFIG]
                     [--cfp-input-levs-config CFP_INPUT_LEVS_CONFIG]
                     [--cfp-input-general-config CFP_INPUT_GENERAL_CONFIG]
                     [--cfp-input-track-only-config CFP_INPUT_TRACK_ONLY_CONFIG]
                     [--cfp-output-levs-config CFP_OUTPUT_LEVS_CONFIG]
                     [--cfp-output-general-config CFP_OUTPUT_GENERAL_CONFIG]
                     [--satellite-plugin-config SATELLITE_PLUGIN_CONFIG] [--regrid-z-coord REGRID_Z_COORD]
                     [-r REGRID_METHOD] [--chosen-obs-fields CHOSEN_OBS_FIELDS]
                     [--chosen-model-fields CHOSEN_MODEL_FIELDS]
```

Example configuration file

Run with `$ visiontoolkit --config-file="<path to YAML file below>"`

```
{ Optional: assume observations time sampling starts at this time not actual initial time
    "start-time-override": "1998-02-21 11:50:00",
    "obs-data-path":"../data/compliant-data/core_faam_20170703_c016_STANCO_CF.nc",
    "model-data-path":"../data/2025-maria-um-hybrid/*.pp",
    "orography":"../data/2025-maria-um-hybrid/orography.pp",
    "chosen-obs-field":false,
    "chosen-model-field":"air_temperature",
    "outputs-dir":"toolkit-outputs/um-hybrid-height-faam-stanco-1",
    "output-file-name":"um_hh_faam_stanco_1_vision_result.nc",
    "cfp-cscale": "WhiteBlueGreenYellowRed",
    "cfp-mapset-config":{
        "lonmin":-2,
        "lonmax":2,
        "latmin":50,
        "latmax":54,
        "resolution":"10m"
    },
    "cfp-output-levs-config":{
        "min":250,
        "max":290,
        "step":2
    }
}
```

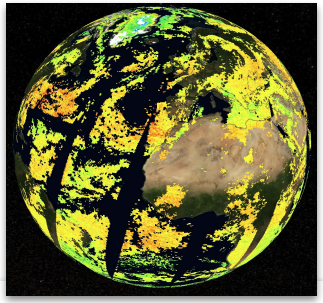
Data inputs and outputs, including which fields (variables) to select from each

cf-plot plotting configuration for optional plots (preview of input, plus output)

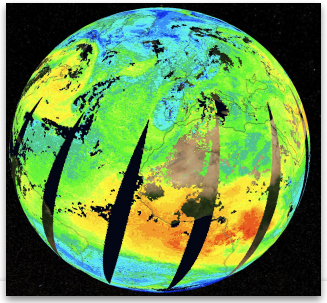
National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Overview: developing the satellite simulator component

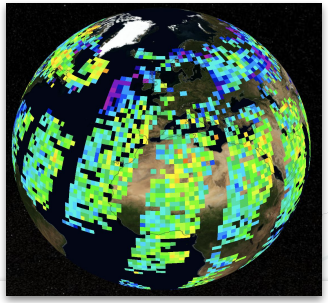
Methane



CO

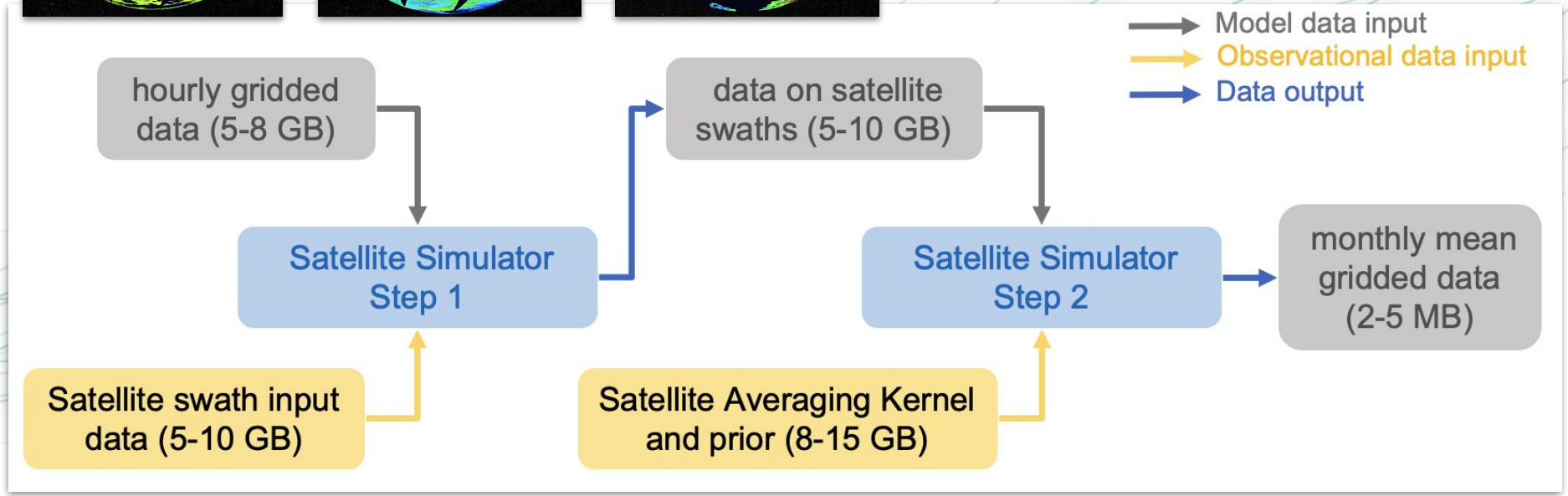


Ozone



'Plugin' toolkit code pre- and post-processes satellite data

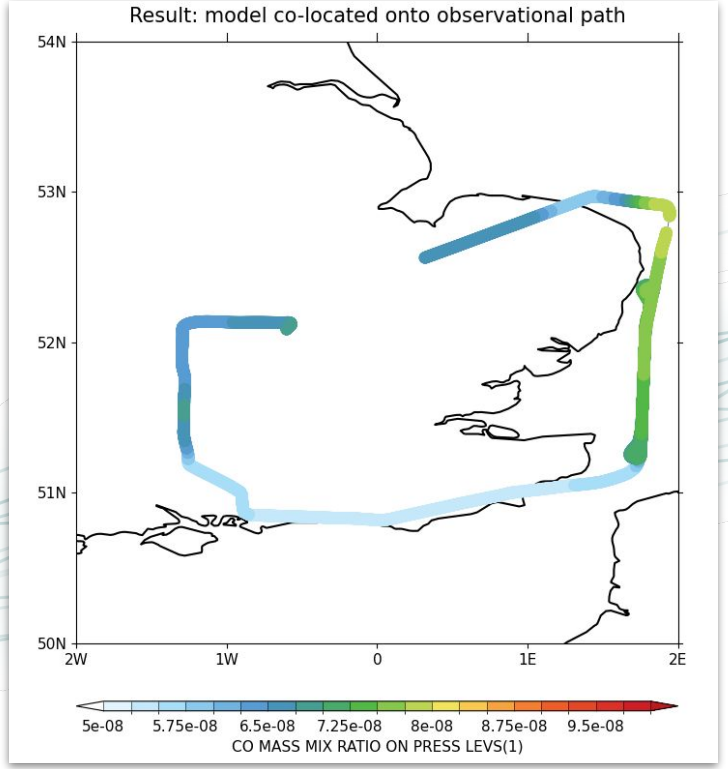
Remote Sensing Group Data Viewer:
https://gws-access.jasmin.ac.uk/public/rsg_share/webpages/rsg_data_viewer/



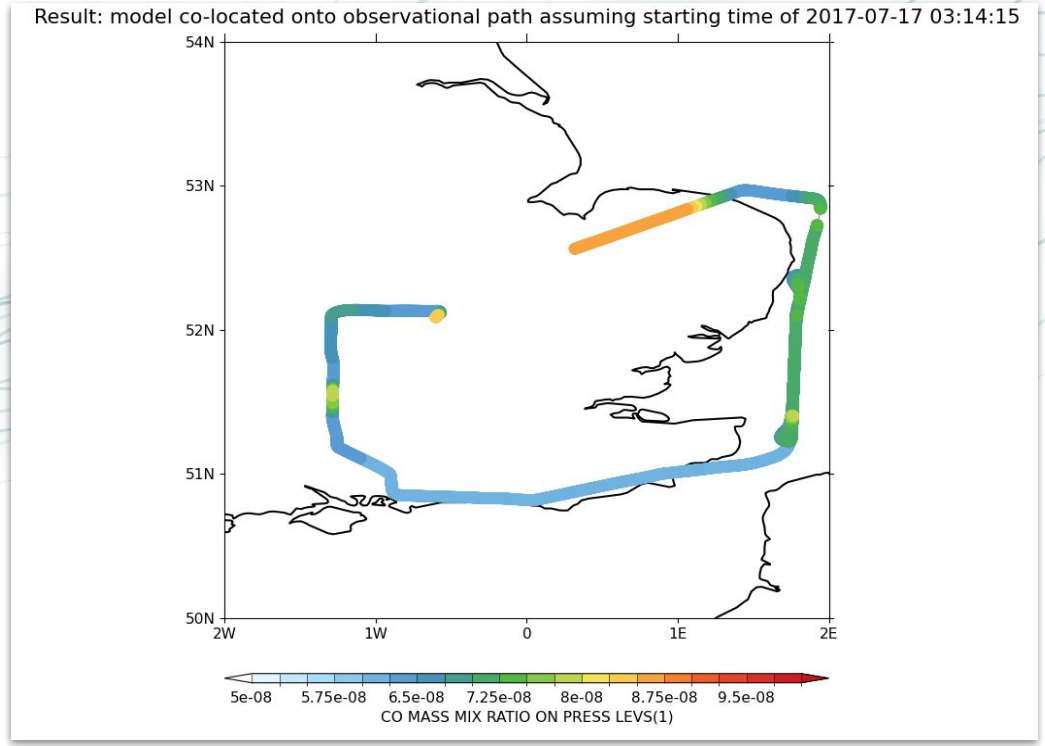
Toolkit v2. flight path example output plots

UKESM model data co-located onto FAAM specific flight track observational data:

Co-located to match the actual flight time



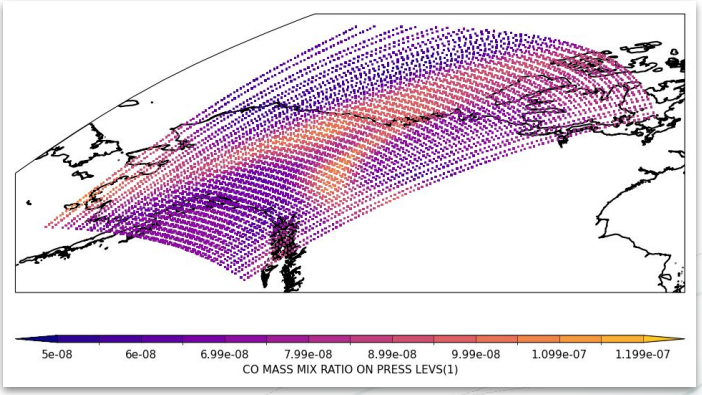
Set (by an input datetime)to assume a different start time of the flight



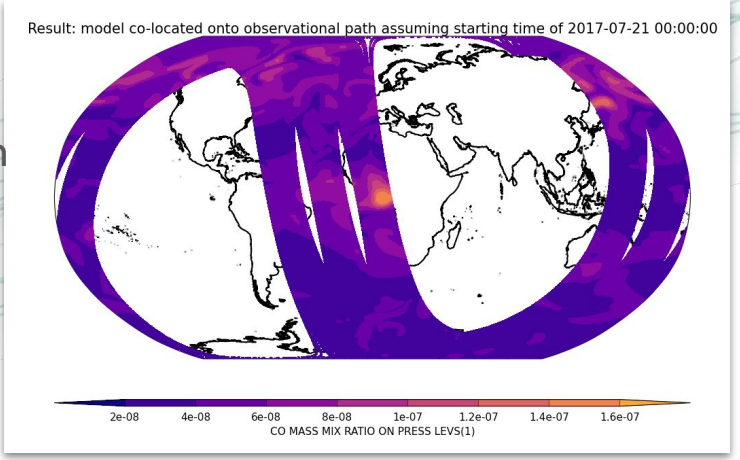
Toolkit v2. satellite example output plots

UKESM model data co-located onto the satellite observational data:

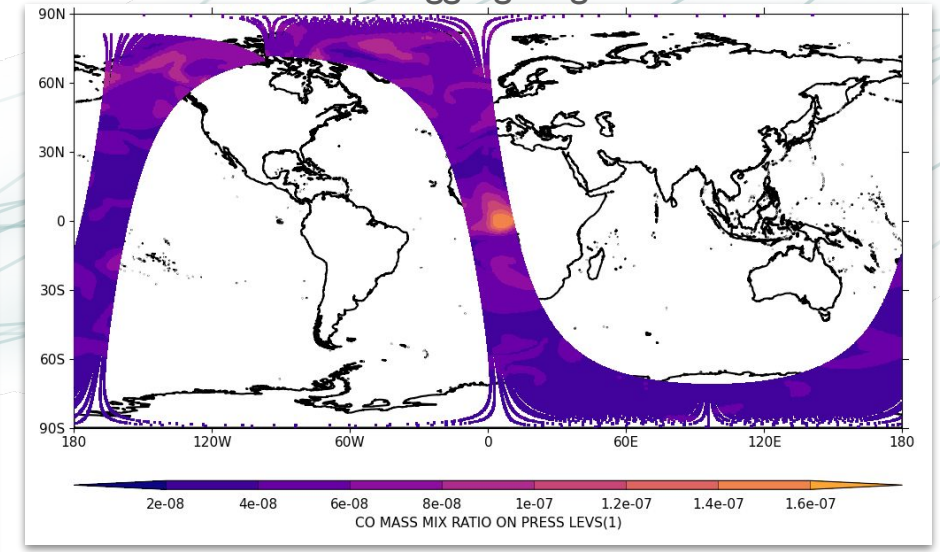
A *single swath* from *single data file*



Multiple orbits from reading a *full directory* of files



One *full orbit* of swaths, from toolkit natively *reading many files (by glob pattern)* like that shown to the left and aggregating



National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

- Existing applications of the VISION Toolkit, as examples

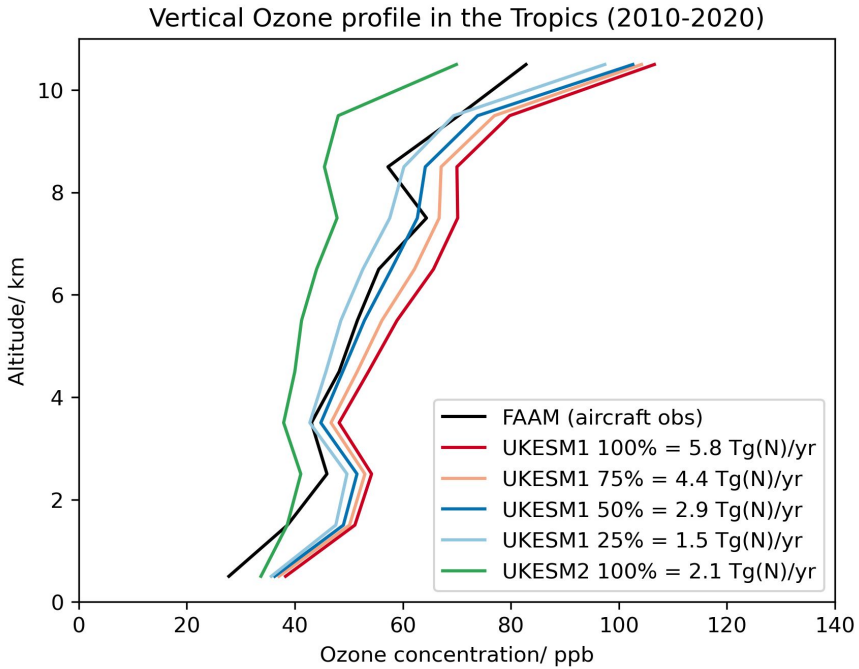
UNIVERSITY OF CAMBRIDGE | Yusuf Hamied Department of Chemistry

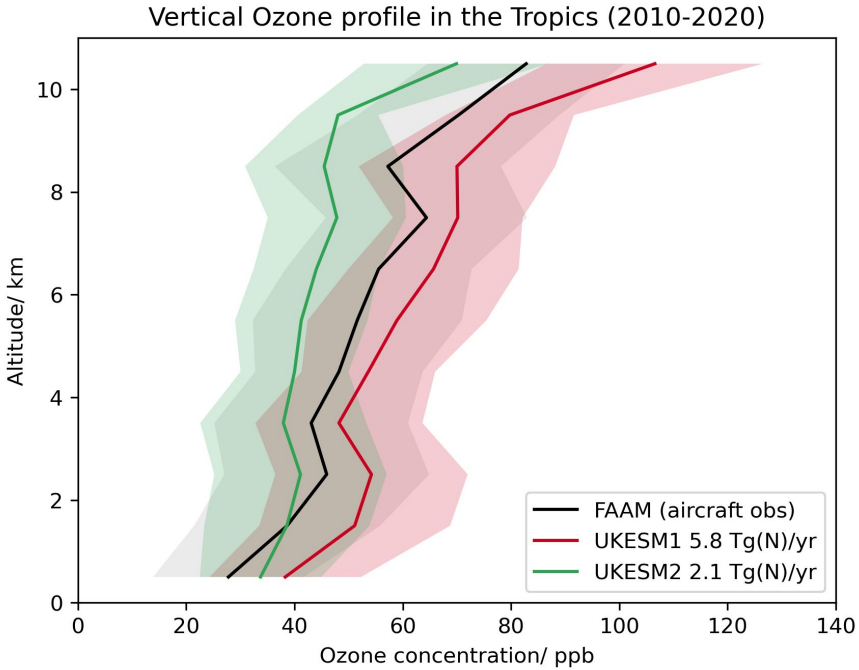
National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

FAAM

UNIVERSITY OF LEEDS

Improving UKESM2 using the VISION Toolkit

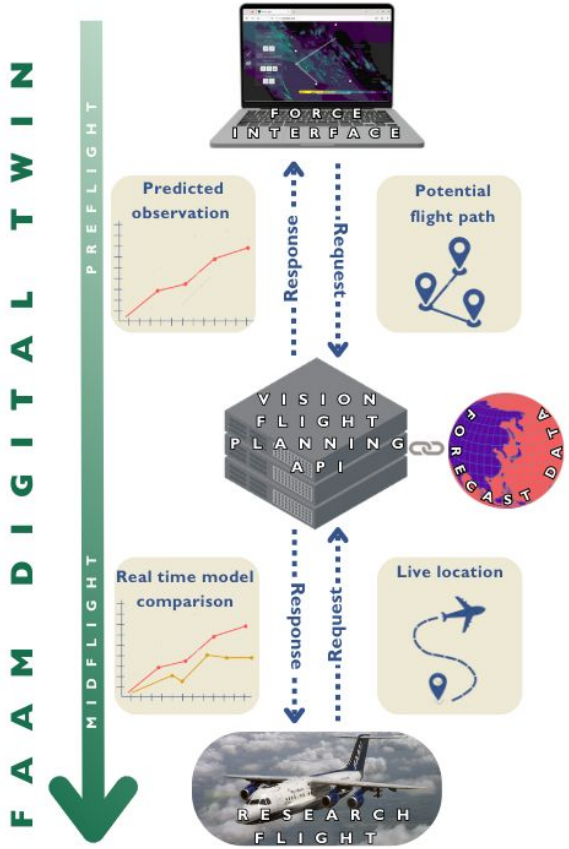


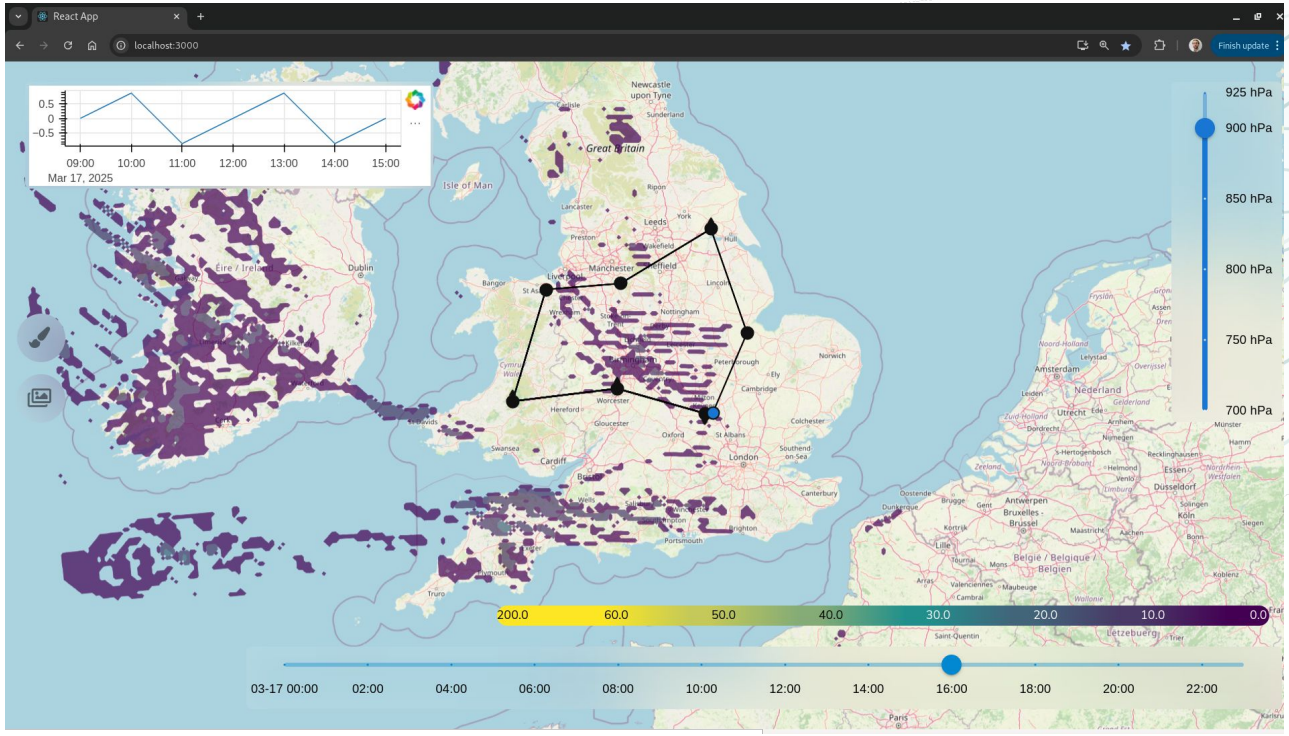


The VISION Toolkit is being used to improve the prototype configuration of UKESM2 by comparing against historical FAAM data. This has highlighted deficiencies with the model lightning NOx emissions parameterisation.

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Developing the FAAM Digital Twin in FORCE





National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

Status and Summary

UNIVERSITY OF CAMBRIDGE
Yusuf Hamied Department of Chemistry

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

FAAM

UNIVERSITY OF LEEDS

Status of the work and v.2 toolkit

- I agreed to give this talk with optimistic timing and unfortunately v.2 is not quite ready and released, but we expect it will be released and installable in standard Python library ways e.g. `pip install visiontoolkit` within ~2 months: end of June latest
- For now, tomorrow I am going to prepare a pre-v.2 version which can be installed locally (via cloning the repo and running a command) so you can try it out in the meantime (satellite averaging kernel algorithms are still being finalised so not included for now)
- The code lives in a GitHub repository at: https://github.com/NCAS-VISION/VISION-toolkit-v2-development, with basic documentation including of the CLI and Python API at: https://ncas-vision.github.io/VISION-toolkit-v2-development/build/html/index.html - all open-source and viewable, with an Issue Tracker for public use

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

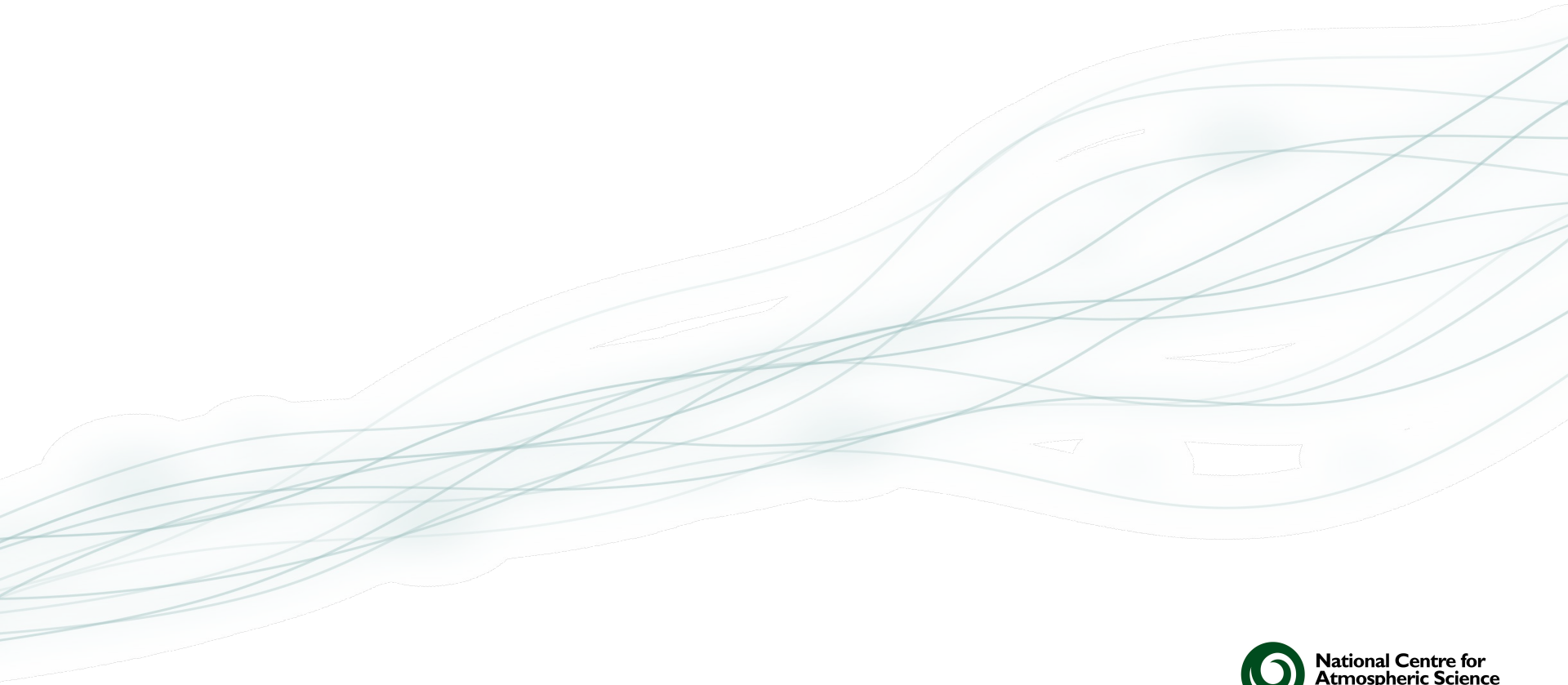
Summary: VISION (as of spring 2025)

- VISION project: based on initial work detailed in Russo. et al 2025 (https://gmd.copernicus.org/articles/18/181/2025/gmd-18-181-2025.html), with 15 months NERC-TWINE funding a cross-NCAS/NCEO team has produced a v.2 toolkit & FAAM digital twin
- The Toolkit is a Python library designed to mainly be used via the command-line interface (CLI)
- The basic idea is it takes as inputs model and observations data and co-locates via interpolation the model output onto the observational paths - sufficiently CF-compliant model and obs are supported
- Code at github.com/NCAS-VISION/VISION-toolkit-v2-development, documentation at ncas-vision.github.io/VISION-toolkit-v2-development/build/html/index.html
- Version 2 expected to be formally released by the end of June latest. A v.2 pre-release version will be made tomorrow, available for use, with installation instructions on the repo README. **Please try it out and/or get in touch with any thoughts, feedback or questions!**
- The VISION project already has: been used towards improving the prototype configuration of UKESM2; been integrated into FORCE; & improved the existing NCAS tools cf-python and cf-plot
- **Get involved! Please ask us how VISION can best help you with your own work**

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- Extra slides/images

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL



National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL