

How to benchmark your model? Making clarity on the existing tools

Composite Mapping Tools and Model Validation Tools

From FAIRMODE website

- https://knowledge4policy.ec.europa.eu/fairmode-forum-air-quality-modelling/fairmode-benchmarking-tools_en

PAGE | LAST UPDATED: 24 SEP 2026

FAIRMODE Forum for Air Quality Modelling

Page contents

Composite mapping tools

Model validation tools

Legacy versions (old versions of Delta Tools)

Composite mapping tools

These on-line tools are used for FAIRMODE intercomparison exercises, where data from different groups are available in a common repository. These on-line tools for benchmarking emissions and concentration are useful for "Assessment" purposes, focusing on:



- **Emission benchmarking:**
 - [Emission App](#): to compare different emission inventories and derive emission benchmarking indicators.
- **Concentrations benchmarking**
 - [Concentration App](#): to compute the long-term Model Quality Objective (MQO) comparing model results and observed values, on yearly values data.
 - [Best model App](#): to interactively display best models for each administrative polygon at NUTS2 level, featuring on-the-fly models mosaic display.
 - [Model viewers](#): to interactively overlap models and visualize concentration profiles charts along custom lines.

Please note that, to be able to upload maps in the **Concentration App**, you need to be authorized (to do so please write to our functional mailbox JRC-FAIRMODE@ec.europa.eu and we'll activate your "upload" access).

A recording on how to use Concentration app the tool is available here <https://webcast.ec.europa.eu/presentation-of-the-fairmode-concentration-dashboard-march-2026> (password: [YMjh5356rkiY](#)).

If you have any feedback or suggestions to improve the tools, please write to our functional mailbox JRC-FAIRMODE@ec.europa.eu.

Model validation tools

These tools are for model validation for individual assessment. The tools are currently developed in Python while the legacy tools were initially developed as IDL-based tools.



FAIRMODE benchmarking tools

- FAIRMODE offers a suite of tools to **benchmark air quality models** under the European AAQDs — organised in **two families**.
 1. **Composite mapping tools** — online apps for FAIRMODE **intercomparison exercises**, using data from many groups held in a common repository (emissions, concentrations, best-model, viewers).
 2. **Model validation tools (the Delta-tool)** — for **individual** validation of your own model against observations; in **Assessment** and **Forecast** versions.



Focus of the session: concentrations

1. Composite Mapping Tools

online · intercomparison across groups

- Run **online** (Jupyter / Voilà on the JRC platform); data from different groups in a common repository.
- **Concentration App** — concentrations benchmarking: long-term MQO from **yearly** model maps vs observed values.
- **Best-model App** & **model viewers** — on-the-fly composite map merging “best” concentrations, best model per NUTS2 polygon, MQI per NUTS2...
- Uploading maps need authorisation (contact JRC)

2. Model Validation Tools — Delta-tool

your own model vs observations

- For **individual** model assessment, not group intercomparison.
- Works on **time series** of model vs observed values.
- Computes the **full suite of indicators** (short- & long-term MQO / MQI / MPC) of the WG2 guidance.
- Two versions: **Assessment** and **Forecast**, implementing FAIRMODE guidance documents.



Today's focus

- **1. Composite Mapping Tools (Concentration App)** — compute the long-term MQO by comparing model concentration maps with observed **yearly** values.
- **2. Model Validation Tools — Delta-tool (Assessment)** — compute the **short- and long-term MQO** for your model from **time series** vs observations, with the complete WG2 indicator suite.
- Both support the fitness-for-purpose check of modelling applications under the AAQD (AAQD uncertainty).

▶ Live demo of both in a few minutes

- Few more words on the Model Validation Tools in the next slides



Model Validation Tools: Delta-Tool

Graphical interface (GUI)

online app

- Interactive **web app** (Jupyter / Voilà on the JRC platform), reached through EU-Login.
- Upload your model–observation data, choose pollutants and stations, get indicators and diagrams interactively.
- Ideal for quick, standardised benchmarking **without coding**.
- Demo in a few minutes
- <https://jeodpp.jrc.ec.europa.eu/eu/vaas/voila/render/fairmode/fairmode/DeltaTool.ipynb>

Python code

open source code for scripts

- The **same methodology** of the GUI as a Python package you can run locally or in batch.
- Reproducible, automatable and extendable — it is the **engine behind the online app**.
- Open source on code.europa.eu, with sample data and an online user guide.
- <https://code.europa.eu/jrcairqualitymodelling/deltatool-assessment>



Model Validation Tools: Delta-Tool (python code)

The new Python code

- A **Python re-implementation (porting) of the IDL Delta-Tool Light** (developed by Kees).
- **Open source** on code.europa.eu.
- Keeps the Delta-Tool Light **input formats**; computes short- & long-term MQO and the WG2 indicator suite.
- Shares a harmonised codebase with Forecast.
- It uses updated uncertainty curves (from AAQD)

vs Delta Tool Light (DTL) and CEN MQOR

- DTL was the **efficiency / batch** re-write of the Delta Tool for multi-model, multi-species, multi-domain, multi-period studies.
- MQOR (in R) implements CEN TS and is based on CEN uncertainty curves



Model Validation Tools: Delta-Tool (updates)

- Implements both fixed and indicative measurement uncertainty curves
 - Via configuration file (YAML) if all stations of the same type
 - Via startup.ini (13th column) if mixed stations
- With both fixed and indicative stations
 - Pooled summary
 - Per-type summary
- Indicators (WG2 guidance + <https://gmd.copernicus.org/articles/18/4231/2025/gmd-18-4231-2025.html>)
 - MQI (MQI_short, MQI_long)
 - Complementary Temporal Indicators (TI_B, TI_R, TI_σ) ... The TI_* *decomposes* the short-term MQI into 3 skills
 - Complementary Spatial Indicators (SI_RMSE, SI_B, SI_R, SI_σ) ... *spatial* variation across network
 - Spatial gradient MPIs (UT-UB, UB-RB) ... *expected* gradients between station types
 - Temporal-coherence MPIs (seasonal, week vs weekend, day vs night)



Conclusions

- We propose tools for
 - Composite Mapping Tools: Joint activities / intercomparison
 - Model Validation Tools: Your own model validation (via GUI or py code)
- Proposal to use Composite Mapping Tools for joint exercise:
 - later on in this session
- Questions or feedback?

