Use of the MQO in project data WG4

Dublin, Ocober, 2024

What can we learn from Antwerp?

National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

- Project data, like the results of the Antwerp benchmark,
  can serve different purposes:

  - Use the project data to test aspects of the MQO.
  - I.e. effect number of stations, robustness.

  - Use MQO as *model validation*?
 - Is the MQO an alternative for other validation
    methods?

Use data to test the MQO

- A critical issue of the MQO is the available number of data points, i.e. number of MQI.
- If we have a sufficient large number of actual datapoints, we can test the effect of using all data or a limited number.

- WG4/Antwerp experiments: 28 locations with useful data to calculate MQI.
- Use uncertainty of indicative measurements in the calculation of the MQI of the Antwerp data (parameters status spring 2024).

- Make many random selections of an increasing number of MQI for each model and check the MQO → what fraction of the runs fulfill / pass the MQO?
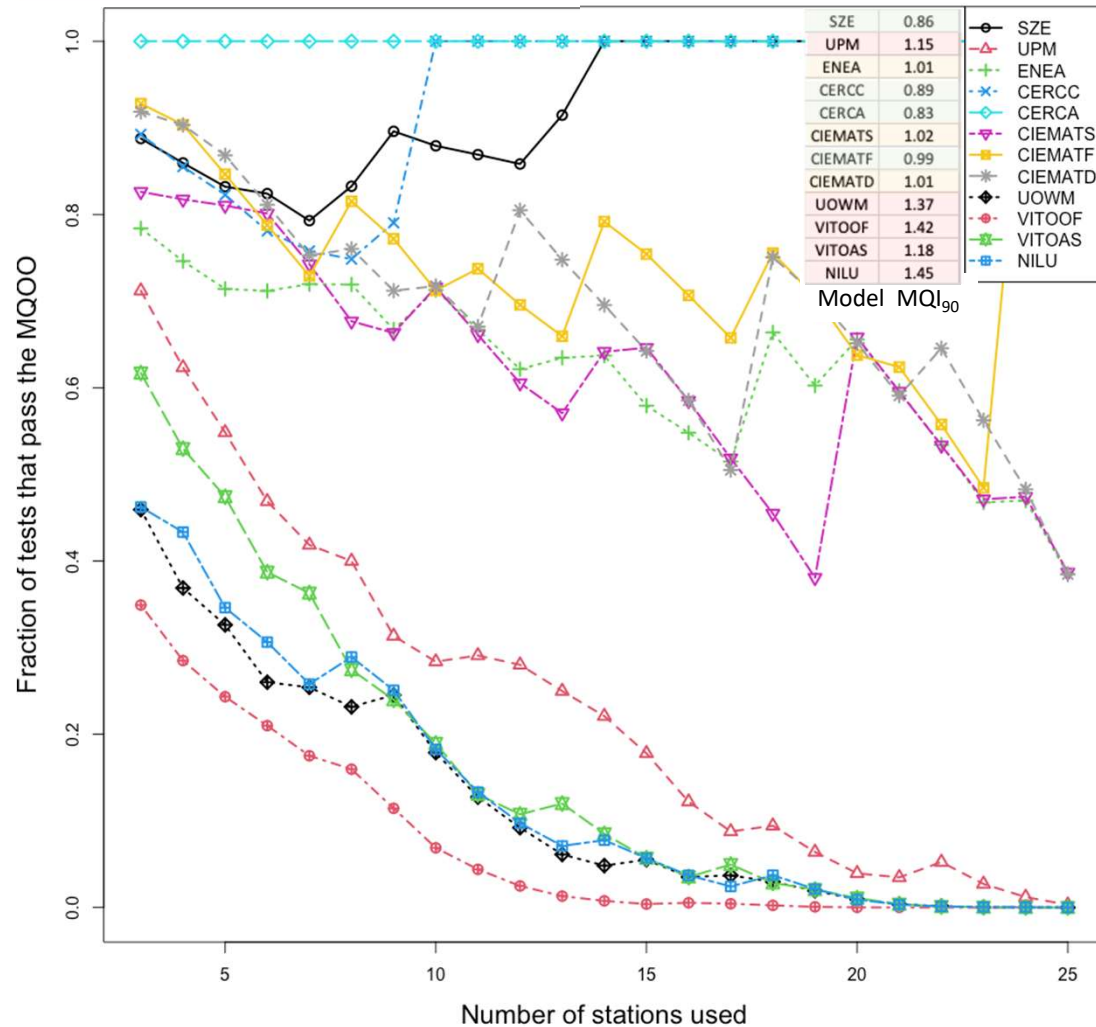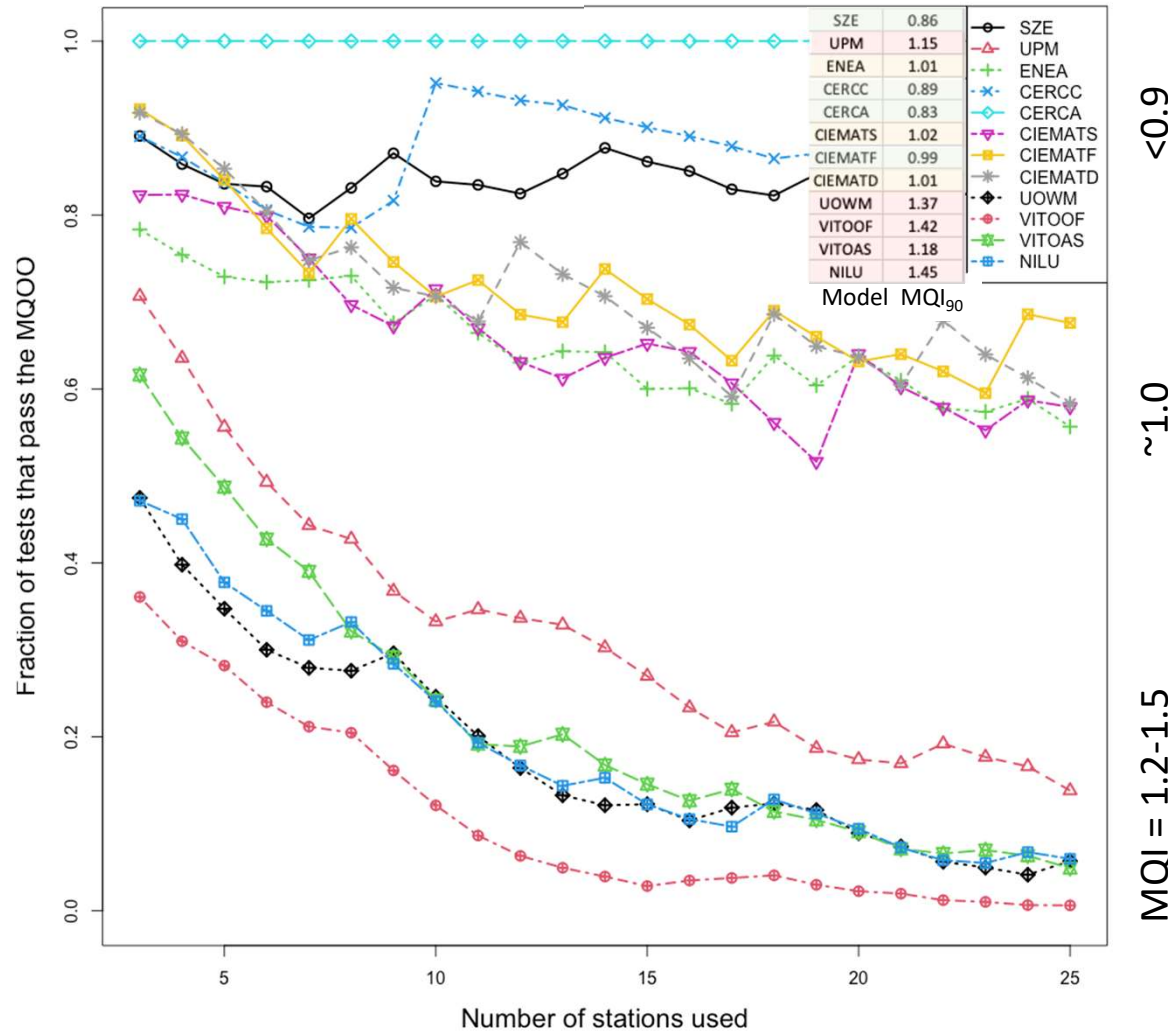
- Result of using 10000 random selections of the MQI obtained in the WG4/Antwerp experiments.
- Every value in the data set could only be drawn <u>once</u>.
- Some models clearly (almost) always pass the MQO.
- Some models are "iffy".
- And some models fail, but need at least 10 MQI to be reasonably certain.
- More MQI is better.

- Result of using 10000 random selections of the MQI obtained in the WG4/Antwerp experiments.
- Every value in the data set could be drawn <u>more than once</u>.
- Some models clearly (almost) always pass the MQO.
- Some models are "iffy".
- And some models fail, but need at least 10 MQI to be reasonably certain.
- More MQI is better.
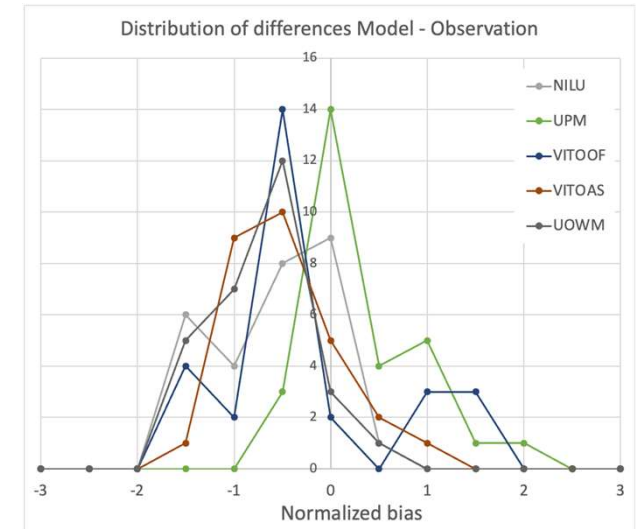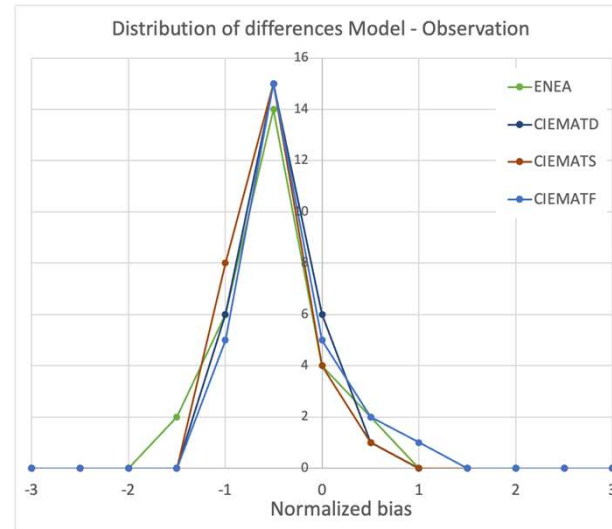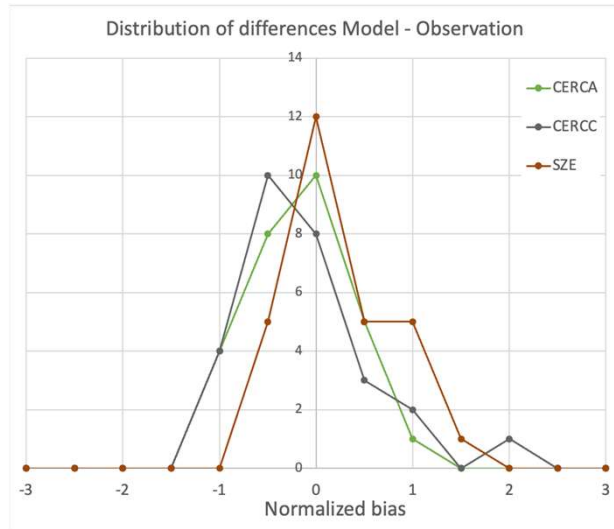- With MQI>100, the models ENEA, CIEMATD/F/S drop below 50% pass.

- The MQO is a binary test: pass / not pass.
- It does not give a clue what the problem is when a model does not pass …
- Plot the normalized bias, i.e. (Model-Observation) / scaled uncertainty.
- Bins of 0.5 ug/m3.

Model	MQI_{90}
SZE	0.86
UPM	1.15
ENEA	1.01
CERCC	0.89
CERCA	0.83
CIEMATS	1.02
CIEMATF	0.99
CIEMATD	1.01
UOWM	1.37
VITOOF	1.42
VITOAS	1.18
NILU	1.45

National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

Use data to validate model

| Model	MQI$_{90}$
SZE	0.86
UPM	1.15
ENEA	1.01
CERCC	0.89
CERCA	0.83
CIEMATS	1.02
CIEMATF	0.99
CIEMATD	1.01
UOWM	1.37
VITOOF	1.42
VITOAS	1.18
NILU	1.45
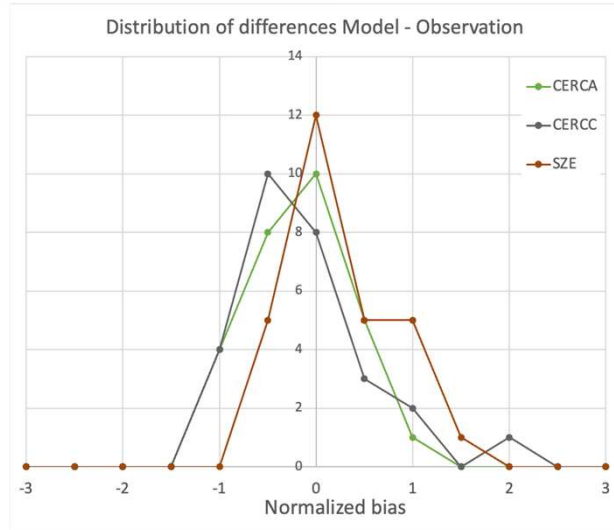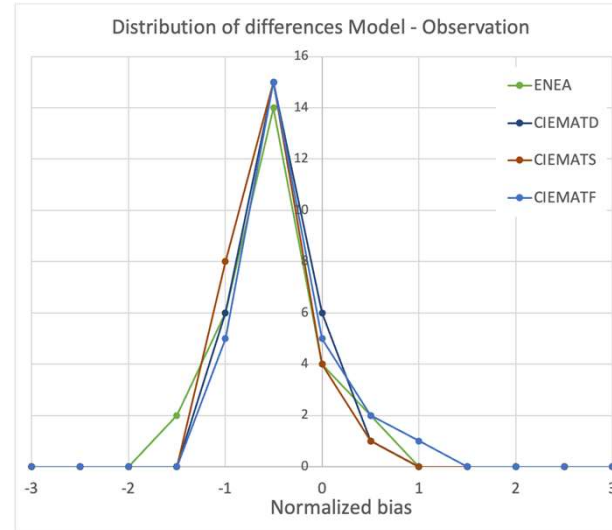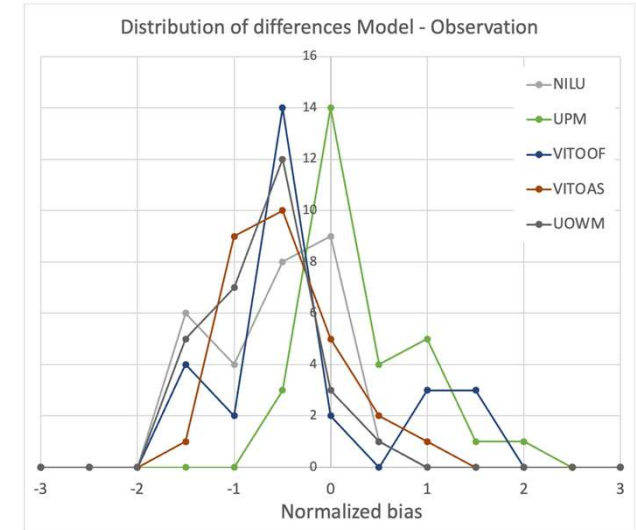
MQI < 0.9   ~1.0   MQI = 1.2-1.5

- The biases of the models that almost always pass the MQO are distributed around zero and the distributions are not very wide.
- The models in the middle group have a nice small distribution, but offset from zero → bias effect?
- The models that clearly fail the MQO have wide distributions of biases, partly offset → other issues? More information is needed.

7
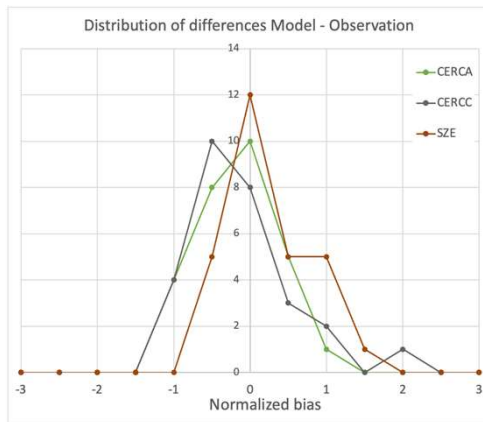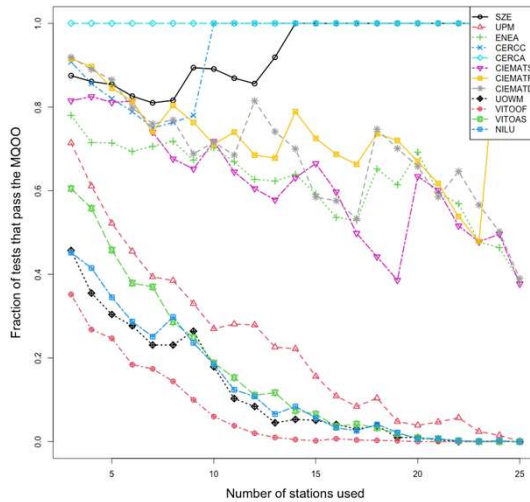
- Experimental data sets as collected by WG4, comparing observed concentrations to model results, are very important to test/verify the use of MQI/MQO.
- At least 10 MQI seem required to evaluate the MQO.

- Validation of models requires more information than simple pass/fail.
- The MQI/MQO can help during the process of model validation but are only a small aspect of what is needed for a full validation of models.

Questions?

Use of project data | FAIRMODE October, 2024