FAIRMODE WG5: Bias projection exercise

Stijn Janssen, Joana Soares, Enrico Pisoni, Alexander de Meij & Philippe Thunis January 17, 2025

Context

When setting up an air quality plan or a roadmap, the new AAQD requests for an assessment of absolute concentration levels that can be benchmarked with limit or target values. These concentration levels must be assessed in an (unknown) future projection under certain policy scenarios.

To estimate absolute concentration levels, the bias observed in modelling results should be eliminated as much as possible. Since the bias can only be quantified at station locations, the removal of a residual bias over the entire modelling domain has proven to be challenging. An additional challenge emerges in model simulations for future scenarios. The bias observed in the reference situation should be projected into the future to correct the results of the scenario projection. This is process of bias projection is depicted in Figure 1 below.

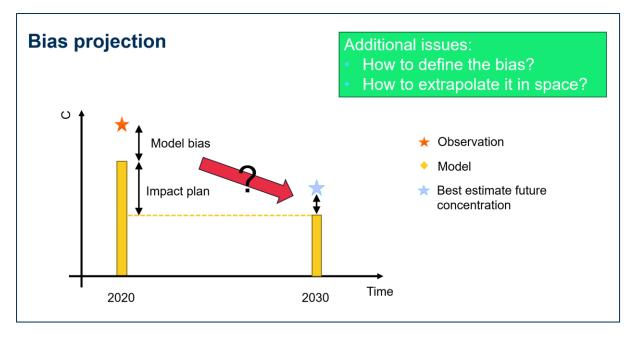


Figure 1: Schematic representation of a bias projection into a future scenario.

Discussions within WG5 made clear that Member States have developed different methodologies to deal with these challenges. To benchmark these methodologies, WG5 is setting up an intercomparison exercise. To minimize the participants effort and to maximize the output of the benchmarking study, an intercomparison based on synthetic data is proposed. Such an approach offers the opportunity that a known "truth" is available for benchmarking.

Setup of the benchmarking study

The layout of the benchmarking study is schematically depicted in Figure 2.

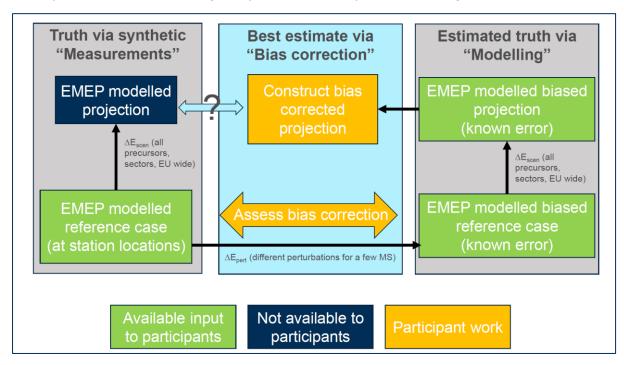


Figure 2: Layout of the synthetic benchmark study.

The EMEP model is used to simulate a reference case and a projection for the full European domain (left part of the scheme in Figure 2). An emission reduction ΔE_{scen} of 50% is applied for all precursors and all sectors in all European Member States. However, the non-anthropogenic emissions are kept constant in the projection. Based on these EMEP model simulations, values are extracted at the location of 3263 fixed monitoring stations in Europe and provided to the participants. These modelled results are considered as the "truth" with related observations in the European monitoring network.

The same EMEP model is used to create perturbated results for both the reference case and the future projection (right of the scheme in Figure 2). To obtain these perturbated results a ΔE_{pert} is applied on a selection of precursors and sectors for a few Member States. ΔE_{pert} can be variable per precursor, sector and Member State.

The perturbated results, reflecting model simulations that deviate from the truth (as happens in reality), are available to the participants. In addition, observations of the reference case at the station locations, reflecting the observed truth are available as well.

Based on the observations and the modelled simulations of the perturbed reference case, a bias correction field can be derived. By combining the bias information and the simulated (perturbated) reference case and projection, a bias corrected reference case and a bias corrected projection can be defined. Both derived data sets can then be benchmarked with the known "truth" of the reference case and the future projection.

Modelling results are available for PM2.5, NO2 and O3 as daily values for the full year and as an annual means. Both data sets are consistent.

Description of the exercise

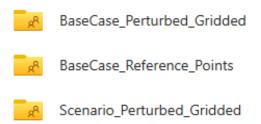
To participate in the benchmarking exercise, the following steps must be performed:

- 1. Download all the data (green blocks in Figure 2) from a JRC server.
- 2. Derive a bias corrected reference case based on your preferred methodology. Note that in this procedure, <u>no new model simulations must be performed</u>. Only post-processing of the available modelling results is required.
- 3. Derive a bias corrected projection based on your preferred methodology.
- 4. Create output of Step 2 and 3 as gridded NetCDF or georeferenced GeoTIFF files of the annual means of the three pollutants for the entire European domain. It should be noted that the perturbations have been applied in some dedicated regions, but we leave it up to the participants to correct where appropriate.
- 5. Deliver the output of Step 4 to the WG5 leads. Your data will be compared with the synthetic truth. This will be used in a benchmarking of the methodologies.
- 6. In a last step, the participants will be asked to briefly describe the bias correction and bias projection methodology they applied, for further reference and interpretation in the exercise.

Description of the data sets

The data is available at a JRC server via this <u>link</u>. Note that a login via a known email address (no password) is required. Contact WG5 leads¹ to get your email address verified.

Under the DATA_PUBLIC folder, two data sets are available: one for daily values, one for yearly values. Participants are free to work with the daily or the annual mean values. Each time aggregation contains the following structure:



Within each folder, data is available for the three pollutants PM2.5, NO2 and O3. Gridded data is provided as NetCDF, point data as csv files.

Timing

The following timeframe is proposed:

- January 17, 2025: data set is available for participants at the JRC server.
- Second half of February 2025: online TEAMS meeting to discuss first results, problems, concerns... A Doodle will be circulated to all participants.
- March 7, 2025: Presentation of the exercise and potential first results at the FAIRMODE Plenary Meeting in Prague.
- April 2025: data delivery by all participants.
- May 2025: online workshop to discuss first results of the benchmark exercise.
- Fall 2025: compilation of scientific publication describing the exercise.

¹ stijn.janssen@vito.be