## IAGOS-DM MOCAGE data set

With their multi-decadal timescale and their geographical coverage reaching the hemispheric scale, the IAGOS measurements are an essential tool for assessing long-term simulations from chemistry-climate/chemistry-transport models. The data are mostly gathered in the upper troposphere – lower stratosphere, a key region regarding the impact of greenhouse gases like water vapour and ozone on the surface temperature, and also regarding the exchanges between the troposphere and the stratosphere. However, the IAGOS data set has still never been used in its whole ensemble to assess model simulations, notably because they are not gridded and their spatio-temporal resolution is far different from the simulations outputs. In the case of the reference simulation with specified dynamics (REF-C1SD) from the Chemistry-Climate Model Initiative (CCMI) experiment, these outputs are characterized by a monthly resolution and a  $\sim$  1 km vertical resolution at cruise altitudes.

The current data set consists in IAGOS data distributed on the MOCAGE (MOdélisation du Climat À Grande Échelle) CTM grid configurated for the REF-C1SD simulation, averaged through each month and each sampled grid cell, for the following observed quantities: ozone, carbon monoxide, temperature, and for the following quantities derived from observations: O<sub>3</sub>/CO ratio and potential temperature. It is therefore named IAGOS-DM\_MOCAGE, with the first suffix DM standing for « distributed on the model grid ». It is the first output from the Interpol\_IAGOS software presented in Cohen et al. (2020, to be submitted). Its goal is to enable an easy and direct comparison between the gridded data sets IAGOS-DM\_MOCAGE and its corresponding simulation output, the MOCAGE REF-C1SD. For each year and each layer, a repertory gathers the IAGOS-DM\_MOCAGE files (simply called « IAGOS\_yyyymm.nc ») and the MOCAGE-M files (called « REFC1SD\_yyyymm.nc »), i.e. the REFC1SD-MOCAGE files after applying a mask with respect to the IAGOS sampling. The suffix M stands for « masked ».

The final outputs from the Interpol\_IAGOS software consist in time series and mean seasonal (and yearly) climatologies. In the corresponding files, the time series have two temporal dimensions: one corresponds to the time series properly (from the 1<sup>st</sup> until the N<sup>th</sup> month) and the other one refers to the seasons (including the whole year as a « fifth season ») in order to facilitate the seasonal treatment.

All the output files are in the NetCDF format.