EGU22-8578, updated on 02 Oct 2022
https://doi.org/10.5194/egusphere-egu22-8578
EGU General Assembly 2022

Data-driven data assimilation to better characterize both accuracy and uncertainty of climate projections: a case study with an idealized chaotic AMOC model

Pierre Le Bras[1,2], Pierre Ailliot[3], Noémie Le Carrer[1,2], Juan Ruiz[4], Florian Sévellec[1,5], and Pierre Tandeo[2]

[1]Laboratoire d'Océanographie Physique et Spatiale, IUEM, Univ. Brest, CNRS, IRD, Ifremer, Brest, France
[2]IMT Atlantique, Lab-STICC, UMR CNRS 6285, F-29238, France
[3]Univ Brest, CNRS UMR 6205, Laboratoire de Mathematiques de Bretagne Atlantique, France
[4]Centro de Investigaciones del Mar y la Atmósfera, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, CONICET-UBA, Buenos Aires, Argentina, UMI-IFAECI (CNRS-CONICET-UBA), Buenos Aires, Argentina
[5]Ocean and Earth Science, University of Southampton, Southampton, United Kingdom

The multi-model ensemble approach is applied in geosciences to provide better predictions or projections, by weighting the outputs from different dynamical models. Basically, the weighting procedure relies on the choice of a performance metric to measure the closeness of individual model outputs to actual observations. The highest weight is then given to the model that best matches the observations, and so forth. Model weights can be used to constrain both the mean and the uncertainty in future projections of climate models.

In this study, we seek to combine different parameterizations of an idealized three-dimensional chaotic model of the Atlantic Meridional Overturning Circulation. One of the parameterizations plays the role of the observations. Each parameterization is evaluated online in a data assimilation framework using the EnKF by comparing the forecasts with the observations.

Traditional data assimilation procedures require access to the model equations, resulting in significant computational costs to run multiple model simulations to obtain forecasts at each time step. Here, a machine learning approach is implemented to provide the forecasts (i.e., analog forecasting). For each parameterization, the classical way of producing the forecasts is, in our case, replaced by an already existing catalog of trajectory time evolutions (e.g., long-term simulations), allowing to statistically emulate the model dynamics. This data-driven methodology retains the benefits given by the classical EnKF (i.e., optimal initial conditions, uncertainties consideration), at low computational costs. For each model-parameterization, a local performance metric (namely, the contextual model evidence) is computed at each time step in order to compare observations and model forecasts. This metric, based on the innovation likelihood, is sensitive to differences in the model dynamics and takes into account both the uncertainties of the forecasts and of the observations. To validate the methodology, different case studies are performed with various sensitivity tests (e.g., changing the parameterization used for the observations).

The results of the proposed weighting scheme on projections are discussed considering different quality metrics compared to benchmark methodologies. These include the equally weighting approach (also called the “model democracy”) and the direct comparison between the climatological probability distributions of simulations and observations.