Model-data comparison challenges in paleo-climate analyses:
Towards an evaluation toolbox for transient climate model simulations

Jean-Philippe Baudouin\textsuperscript{1}, Oliver Bothe\textsuperscript{2}, Manuel Chevalier\textsuperscript{3,4}, Nils Weitzel\textsuperscript{1}, Anne Dallmeyer\textsuperscript{5}, Chris Brierley\textsuperscript{6}, and Kira Rehfeld\textsuperscript{1}

\textsuperscript{1}Institute of Environmental Physics, Heidelberg University, Heidelberg, Germany
\textsuperscript{2}Helmholtz-Zentrum Geesthacht Centre for Materials and Coastal Research, Geesthacht, Germany
\textsuperscript{3}Institute of Geosciences, Sect. Meteorology, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany
\textsuperscript{4}Institute of Earth Surface Dynamics, Geopolis, University of Lausanne, Switzerland
\textsuperscript{5}Max Planck Institute for Meteorology, Hamburg, Germany
\textsuperscript{6}Department of Geography, University College London, London, United Kingdom

VPICO - EGU2021
27 April, 2021
Introduction - Goals

- The issue → to address the challenges of transient simulations in model-data comparisons:
  - Large range of timescales, from monthly to multi-millenial
  - Variables non-simulated by timeslice simulations (vegetation, ice sheets, carbon cycle, ...)
  - Large diversity of proxy data (type, time period, time resolution, ...)

- The objective → to facilitate both inter-model, and model-proxy data comparison by:
  - Providing adaptable tools
  - Providing consistent metrics
  - Producing standardised figures and interactive visualisations

- This is still work in progress, and the toolbox will be released by the end of PalMod-Phase2, in 2022
Data flow

Model → Databases → Proxy

Post-processing → Pseudo-proxy → Quantitative metrics and ensembles → Plots & Interactive Visualisations → Web portal
The toolbox will not come with a proxy or model database.

Instead, it will provide adaptable loading functions to read the various formats of proxy database and CMORised model data.

A harmonisation of the various data types will be performed from within the toolbox coding environment using specifically designed classes and objects.

The toolbox could also be used to expand the databases through various post-processing.
Post-processing & Methods included

- Packages already published: sedproxy and Paleospec among others.

- Methods already published but without packaging:
  - Biomization (Anne Dallmeyer, vPICO & Chenzhi Li, vPICO)
  - Bayesian quantification of the model-data deviation (Nils Weitzel, vPICO)

- Adaptation of currently existing packages:
  - CVDP-PMIP by Chris Brierley & Anni Zhao (vPICO)

- Work in Progress:
  - Ensemble metrics (Oliver Bothe)
  - Other quantitative metrics of the model-data deviation (Manuel Chevalier)
Interactive website

• The toolbox will come with a webserver to display PalMod results. It will also provide the code for the users to host their own results.

• Example with R Shiny App →

• Other possibilities: Plotly, Bokeh (Python), mpld3 (Python/javascript)
Software

• Collection of functions and packages.

• Multi-language, in order to use already published code.
  • Some changes are still needed to adapt the code to the toolbox standards.

• Focus on R and Python for new developments.

• Use of ESMValTool standards to prepare for a future integration.