EGU21-2979
https://doi.org/10.5194/egusphere-egu21-2979
EGU General Assembly 2021

The potential of machine learning for modeling spatio-temporal properties of water isotopologue distributions in precipitation

Kira Rehfeld[1,2], **Jonathan Wider**[1,2], Nadine Theisen[1,2], Martin Werner[3], Ullrich Köthe[2], and Nils Weitzel[1]

[1]Institute of Environmental Physics, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany (krehfeld@iup.uni-heidelberg.de)

[2]Interdisciplinary Center for Scientific Computing, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany

[3]Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Germany

Tracing the spatio-temporal distribution of water isotopologues (e.g., $H_2{}^{16}O$, $H_2{}^{18}O$,$HD^{16}O$, $D_2{}^{16}O$), in the atmosphere allows insights in to the hydrological cycle and surface-atmosphere interactions. Strong relationships between atmospheric circulation and isotopologue variability exist, mitigated by fractionation during phase transitions of water. Isotopic gradients correlate with precipitation amount, temperature, with distance to source areas of evaporation and often follow topographic features. Isotope-enabled general circulation models (iGCMs) have been established to explicitly simulate the processes that lead to these distributions, in response to the changes in radiative forcing, boundary conditions, and including effects of internal variability of the climate system. However, few of these iGCMs[1,2] of varying complexity exist to date and isotopic tracers decrease their computational efficiency.

Here, we evaluate the potential of replacing the explicit simulation of the isotopic component in the water cycle by statistical learning for offline model evaluation at interannual to multi-millennial timescales. This is challenging. While the relevant fractionation processes are well understood, the climate system is a chaotic, nonstationary system of high dimensionality. Therefore, successful statistical prediction requires the (so far elusive) understanding of the timescale-dependent relationships in the climate system. We present a case study on the feasibility of this approach.

We focus on the impact of variable selection (primarily surface temperature, precipitation and sea-level pressure) and boundary conditions (CO_2 concentrations, ice sheet distribution). We also compare different approaches to dimensionality reduction, and compare the performance of different machine-learning approaches including simple linear regression, random forests, Gaussian Processes and different types of neural networks. The accuracy of the predictions is evaluated using regional and global area-weighted mean squared errors across training and evaluation data from individual GCM simulations and across climatic states. We find a high spatial variability of prediction accuracy, modest in many locations with the presently employed approaches. We obtain encouraging results for the prediction of isotope variability in Greenland and the Antarctic.

References

[1] Tindall, J. C., P. J. Valdes, and Louise C. Sime. "Stable water isotopes in HadCM3: Isotopic signature of El Niño–Southern Oscillation and the tropical amount effect." *Journal of Geophysical Research: Atmospheres* 114.D4 (2009)

[2] Werner, Martin, et al. "Glacial–interglacial changes in H 2 18 O, HDO and deuterium excess–results from the fully coupled ECHAM5/MPI-OM Earth system model." *Geoscientific Model Development* 9.2 (2016): 647-670.