EGU21-9375
https://doi.org/10.5194/egusphere-egu21-9375
EGU General Assembly 2021

Automatic input variable selection for analog methods using genetic algorithms

Pascal Horton and Olivia Martius
Institute of Geography, Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland
(pascal.horton@giub.unibe.ch)

Analog methods (AMs) are statistical downscaling methods often used for precipitation prediction in different contexts, such as operational forecasting, past climate reconstruction of climate change impact studies. It usually relies on predictors describing the atmospheric circulation and the moisture content of the atmosphere to sample similar meteorological situations in the past and establish a probabilistic forecast for a target date. AMs can be based on outputs from numerical weather prediction models in the context of operational forecasting or outputs from climate models in climatic applications.

AMs can be constituted of multiple predictors organized in different subsequent levels of analogy that refines the selection of similar situations. The development of such methods is usually a manual process where some predictors are assessed in different structures. As most AMs use multiple predictors, a comprehensive assessment of all combinations becomes quickly impossible. The selection of predictors in the application of the AM often builds on previous work and does not evolve much. However, the climate models providing the predictors evolve continuously and new variables might become relevant to be considered in AMs. Moreover, the best predictors might change from one region to another or for another predictand of interest. There is a need for a method to automatically explore potential variables for AMs and to extract the ones that are relevant for a predictand of interest.

We propose using genetic algorithms (GAs) to proceed to an automatic selection of the predictor variables along with all other parameters of the AM. We even let the GAs automatically pick the best analogy criteria, i.e. the metric that quantifies the analogy between two situations. The first test consisted of letting the GAs select the single best variable to predict daily precipitation for each of 25 selected catchments in Switzerland. The results showed great consistency in terms of spatial patterns and the underlying meteorological processes. Then, different structures were assessed by varying the number of levels of analogy and the number of variables per level. Finally, multiple optimizations were conducted on the 25 catchments to identify the 12 variables that provide the best prediction when considered together.