



## **Using genetic algorithms to explore new predictor variables for statistical precipitation downscaling with analog methods**

Pascal Horton, Rolf Weingartner, and Olivia Martius

Oeschger Centre for Climate Change Research & Institute of Geography, University of Bern, Bern, Switzerland

Analog methods (AMs) allow for the prediction of local meteorological variables of interest (predictand), such as the daily precipitation, on the basis of synoptic variables (predictors). They rely on the hypothesis that similar situations at the synoptic scale are likely to result in similar local weather conditions. AMs can rely on outputs of numerical weather prediction models in the context of operational forecasting or outputs of climate models in the context of climate impact studies. The predictors archive is usually a reanalysis dataset.

Different meteorological variables from the NCEP reanalysis 1 were assessed after its release to identify the best predictors for daily precipitation. This former work provided a basis on the top of which more complex methods were developed by adding additional variables in a stepwise way. However, the first predictors of the method often remain the same, and the selection of new predictors is done manually. Nowadays, several new reanalysis datasets are available and were proven more skilful for analog methods than the NCEP reanalysis 1. The accuracy of several variables has significantly improved and more variables are now available than before. Therefore, the former selection of predictor variables might not be optimal anymore. Different variables from various reanalyses should be assessed, which can turn out to be a cumbersome task if done manually and extensively.

Genetic algorithms (GAs) were shown to successfully optimize the parameters of the AMs, such as the spatial domain on which the predictors are compared, the selection of the pressure levels and the temporal windows of the predictors, a weighting between predictors, and the number of analog dates to select. GAs can jointly optimize all parameters of AMs and get closer to a global optimum by taking into account the dependencies between parameters. Moreover, GAs can objectively infer parameters that were previously assessed manually, and can take into account new degrees of freedom. The mutation operator of GAs was identified as a key element for this application, and new variants were developed that proved efficient, such as the chromosome of adaptive search radius, which takes no control parameter.

Therefore, we propose using GAs to explore the variables from three reanalyses (MERRA-2, ERA-interim, CFSR) and select the most relevant ones, along with the appropriate analogy criteria. Although the expert's expertise remains necessary to supervise the selection of predictors, GAs facilitate the exploration of large datasets. The first tests proved the potential of this approach with the selection of unexpected – but yet relevant – combinations of variables and analogy criteria.