
UNSEEN Analog Weather Generators

Jonathan Wider^{*1,2,3} and Jakob Zscheischler^{1,2,3}

¹Department of Compound Environmental Risks, Helmholtz Centre for Environmental Research UFZ,
Leipzig – Germany

²Department of Hydro Sciences, TUD Dresden University of Technology, Dresden – Germany

³Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig – Germany

Abstract

Daily analog weather generators leverage the constraints that large-scale atmospheric circulation poses on the state and evolution of regional weather. These generators resample a historical dataset while ensuring that successive sampled days have consistent large-scale atmospheric fields, thereby also ensuring temporal consistency between local impact variables to some extent. Compared to other types of weather generators, resampling-based methods have the advantage that same-day dependencies between variables and between locations are automatically captured correctly. However, the generated time series are limited to observed ranges, and even "close" analogs in the historical record are relatively far from each other. We overcome these limitations by constructing an analog weather generator using ECMWF extended ensemble forecast hindcast ("reforecast") data, inspired by the UNSEEN approach of using ensemble simulations to assess and anticipate extremes beyond observed records. Advantages of the chosen ECMWF reforecast dataset include its high spatial resolution and large set of states from a relatively constant climate, while model biases remain limited because the forecasts are initialized from reanalysis data. The spread of the ensemble forecasts also provides a natural comparison point for the quality of analogs. We tune the parameters of the weather generator by optimizing its predictive quality in a forecast setting, and propose sampling the datasets in blocks as a way to trade off the weather generator's flexibility against more accurate temporal dependencies.

We apply our methodology to simulate weather over Europe. Analogs are defined in terms of geopotential height at 500 hPa and computed over an extended region including parts of the North Atlantic. Evaluations are performed on key properties of the generated weather time series, such as their annual cycle and several types of univariate and multivariate extreme events. The weather generator is particularly valuable for applications requiring an accurate representation of dependencies-both between variables and across space. This is the case for a number of different types of compound events.

Keywords: analog, analogue, UNSEEN, resampling

^{*}Speaker