EGU24-18866, updated on 16 Jan 2025
https://doi.org/10.5194/egusphere-egu24-18866
EGU General Assembly 2024

A Gaussian framework for optimal prediction of extreme heat waves

Valeria Mascolo[1], Alessandro Lovo[1], Corentin Herbert[1], and **Freddy Bouchet**[2]
[1]ENS de Lyon, CNRS, Laboratoire de Physique, F-69342 Lyon, France (valeria.mascolo@ens-lyon.fr)
[2]CNRS, École Normale Supérieure, IPSL, Laboratoire de Météorologie Dynamique, France

Heat waves are a growing issue in the current climate, causing damage to human societies and other living beings. As climate warms, heat waves are one of the extreme events that will be exacerbated by the rising average temperatures. Understanding the mechanisms that drive heat waves is hence of vital importance to analyze them and to make predictions to mitigate their impacts. However, the rarity of any extreme event makes it particularly hard to study, especially if one wishes to understand the relation between predictors and probability, for instance using machine learning techniques, which are notoriously data-hungry. A possible solution is to use climate model simulations, but they introduce biases and can be prohibitively expensive to run for appropriate dataset lengths.

In this work, we introduce a simple and powerful statistical model, that is able to skillfully predict rare heat waves in a regime of lack of data, as it is the case for the ERA5 reanalysis dataset.

We focus on two-week heat waves over France, and we notice that composite maps of very extreme events are similar to the ones of much less extreme ones. This holds true for an increasing hierarchy of model complexity, including ERA5. We can thus analyze very extreme events by looking at less rare ones, having the advantage of increasing the available statistics. This effect can be explained assuming that the set of predictors and the heat wave amplitude are jointly gaussian. The prediction task can be thus rephrased into estimating the conditional probability of an extreme heat wave happening conditioned on the state of the set of predictors. This quantity, called committor function, is generally hard to estimate given the high dimensionality of the variables involved. Our assumption performs a dimensionality reduction, where an estimate is obtained through an optimal linear projection. The projection map will suggest the most relevant predictors.

We first demonstrate that this Gaussian approximation performs well on a 8000 years run of a model of intermediate complexity, PlaSim, for both analysis and prediction tasks when 500 hPa geopotential height, temperature and soil moisture are used as input fields. When we compare with a machine learning estimation of the committor, which is possible here thanks to the huge amount of available data, our Gaussian model is less skillful, but not by much. We then apply the Gaussian assumption to the 500hPa geopotential height field of the ERA5 dataset, outperforming the neural networks and extending the predictability horizon by several days.

The prediction method we propose is simple and effective. It opens possibilities to achieve skilfull predictions that outperform all known approaches, for instance machine learning, notably when data is scarce. Our method also serves as a better baseline than the climatology to benchmark more complex approaches. Since our statistical model is also interpretable, this framework has potential to go beyond prediction skill only and, thanks to the optimal projection map, foster the study of fast and slow drivers, and the effect climate change has on them.