Evaluating probabilistic forecasts of extremes using continuous ranked probability score distributions

Maxime Taillardat[1], Anne-Laure Fougères[2], Philippe Naveau[3], and Raphaël De Fondeville[4]

[1]Météo-France, Toulouse, France (maxime.taillardat@meteo.fr)

[2]Univ. Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, F-69622 Villeurbanne, France

[3]Laboratoire des Sciences du Climat et de l'Environnement, UMR 8212, CEA-CNRS-UVSQ, IPSL & U Paris-Saclay, Gif-sur-Yvette, France

[4]Swiss Data Science Center, ETH Zürich and EPFL, Switzerland

Verifying probabilistic forecasts for extreme events is a highly active research area because popular media and public opinions are naturally focused on extreme events, and biased conclusions are readily made. In this context, classical verification methods tailored for extreme events, such as thresholded and weighted scoring rules, have undesirable properties that cannot be mitigated, and the well-known continuous ranked probability score (CRPS) is no exception.

Here, we define a formal framework for assessing the behavior of forecast evaluation procedures with respect to extreme events, which we use to demonstrate that assessment based on the expectation of a proper score is not suitable for extremes. Alternatively, we propose studying the properties of the CRPS as a random variable by using extreme value theory to address extreme event verification. An index is introduced to compare calibrated forecasts, which summarizes the ability of probabilistic forecasts for predicting extremes. The strengths and limitations of this method are discussed using both theoretical arguments and simulations.