Geophysical Research Abstracts
Vol. 20, EGU2018-6894, 2018
EGU General Assembly 2018

Advanced score for the evaluation of prediction skill

André Düsterhus (1) and Sabrina Wahl (2,3)

(1) Institute of Oceanography, Center for Earth System Research and Sustainability (CEN), University of Hamburg, Germany (andre.duesterhus@uni-hamburg.de), (2) Meteorological Institute, University of Bonn, Germany, (3) Hans-Ertel Center for Weather Research, Climate Monitoring Branch, Germany

Deterministic, ensemble and statistical predictions deliver large datasets and with it a challenge for verification and evaluation. These datasets can be regional or global variable fields but also multidimensional time series. An important part of the verification procedure is the selection of a score and with it the test hypothesis. In seasonal and decadal prediction there are currently mainly two methodologies used to evaluate and compare datasets. The Anomaly Correlation Coefficient (ACC) evaluates the linear relationship of two datasets, while the root-mean-square error (RMSE) analyse the absolute deviation of one dataset from the other. Both methods rely on assumptions, among others normality assumptions of the data or uncertainty. As a consequence, skill can be under or overestimated when these procedures are used on non-normal-distributed data. Furthermore, ACC and RMSE are hardly comparable to each other, as their results are given in different ranges and units.

We have developed a new score to evaluate predictions, which is independent of the distribution of the data (or does not require any distributional assumptions). The score bases on a metric from image processing, the Earth Mover's Distance (EMD), which is applied to multi-categorical contingency tables. Using a categorisation allows to evaluate clearly non-gaussian datasets like precipitation or climate indices, which are common on many time scales of predictions. Moreover, by selecting different strategies for the categories, the EMD score is able to either replicate the results from the ACC and RMSE or estimate a score without the normality assumption.

Our contribution will focus on the creation of the categorical EMD-score and its application on different large datasets. We show that even meteorological variables, which are widely assumed following a normal-distribution, like surface temperature for a seasonal prediction system, are effected by the normality effect. In a comparison we present how to generate spatial patterns with this new score, which are similar to the ACC and RMSE and so bring both procedures to the same basis and making them comparable. We also show the application on non-normal distributed variables and the future potential of these kind of scores.