Geophysical Research Abstracts
Vol. 21, EGU2019-15695, 2019
EGU General Assembly 2019

Objective precursor detection algorithm - making long-lead predictions of eastern U.S. hot days from sea surface temperatures

Sem Vijverberg and Dim Coumou

Vrije Universiteit Amsterdam, Institute for Environmental Studies (IVM), Water & Climate Risk, Utrecht, Netherlands (s.p.vijverberg@vu.nl)

Recently it was shown that a specific sea surface temperature (SST) pattern, called the Pacific Extreme Pattern (PEP), can give significant skill for 'hot-day' events in the eastern U.S. at up to 50-day lead time (McKinnon et al, 2016). However, the applied method to identify this pattern was not strictly objective and the authors did not test their predictive model on independent data. We confirm that a forecasting method based on the PEP pattern can provide skillful hindcasts. However, when applied to independent data this skill drops substantially. In this work, we present an objective precursor detection scheme which indeed finds an SST pattern at long lead time similar to PEP. Using our optimal precursors, we achieved improved significant ROC scores on the order of 0.61 at a lead time of 50 days when validated on independent data.

We have improved important aspects of the approach by McKinnon et al., namely, our method enables (1) extraction of important patterns in an automotive an objective matter (2) it quantifies the importance of each precursor region. An impressively improved skill is achieved and that can be attributed to the following. Firstly, instead of using a spatial subset, our method takes in a global SST dataset and extracts the precursor regions automatically. This avoids discarding tele-connective distant precursor regions. Secondly, it also selects optimal precursors for each individual lag. Thirdly, we calculate the spatial mean of our precursor regions and use these time series to train a logistic model. This training allows to give weights to each region, which corresponds to the importance of the precursor-region for the event.

Our new method renders significant (tested using bootstrapping) skill up to 50 days (ROC = 0.61) for eastern U.S. 'hot-day' events, compared to only 6 days (ROC = 0.56) using the approach by McKinnon et al.. When predicting at greater lead times, the precursor-regions shift towards more far-away regions, in particular the tropics.

Recognizing that the interaction, or conditional relationships, between the precursor-region time-series are key, we will also present forecast results based on Random Forests and Neural Network based methods. The logistic regression is nevertheless useful to quantify the relative contribution of each individual precursor time-series, while the trained neural network is expected to render the best predictions by accounting for their non-linear interaction as well. Since our presented scheme is generic, it can be applied to a variety of event time series where long-lead predictive skill is expected.