

EGU2020-7401

<https://doi.org/10.5194/egusphere-egu2020-7401>

EGU General Assembly 2020

© Author(s) 2023. This work is distributed under the Creative Commons Attribution 4.0 License.



## A boosting algorithm for Generalized Extreme Value distributions

Madlen Peter, **Alexander Pasternack**, and Henning Rust

Institute of Meteorology, FU Berlin, Germany

In weather and climate science statistical modeling is applied for manifold problems. Due to the increasing number of input variables, overfitting can easily deteriorate the performance for model predictions. In order to avoid this, it is often meaningful to apply model selection approaches. Since conventional approaches can be very time-consuming especially for many predictors, we are using the boosting approach, which combines model selection and parameter estimation. This iterative algorithm identifies and updates in each step only the most important coefficient, such that in the end most important predictor variables have non-zero coefficients and less relevant variables are ignored.

Boosting has been originally developed for classification problems but has also been extended and used for other applications; i.a. non-homogeneous gaussian regression. Based on the non-homogeneous boosting proposed by Messner et al. (2016), which is used to model mean and variance of a forecast distribution simultaneously, we have developed a boosting algorithm for a non-stationary Generalized Extreme Value distribution (GEV). Thus, it is possible to identify the most relevant predictor variables for location, scale and shape parameter concurrently. We apply this algorithm to various toy model simulations to assess the effect of this novel approach.