EGU2020-5279
https://doi.org/10.5194/egusphere-egu2020-5279
EGU General Assembly 2020

Evaluation of CMIP6 simulations of temperature extremes using proper evaluation methods, observations and reanalyses

Thordis Thorarinsdottir[1], **Jana Sillmann**[2], and Marion Haugen[3]
[1]Norwegian Computing Center, Oslo, Norway (thordis@nr.no)
[2]CICERO Center for International Climate Research, Oslo, Norway (jana.sillmann@cicero.oslo.no)
[3]Norwegian Computing Center, Oslo, Norway (marionh@nr.no)

Climate models aim to project future changes in important drivers of climate including atmosphere, oceans and ice, and their interactions. A comprehensive evaluation of climate models thus requires evaluation methods, or performance measures, that are flexible, specific and can address also extreme events. Climate models have traditionally been assessed by comparing summary statistics or point estimates that derive from the simulated model output to corresponding observed quantities using e.g. RMSE. However, it has been argued persuasively that probability distributions of model output need to be compared to the corresponding empirical distributions of observations or observation-based data products. Observation-based gridded datasets for climate extremes, despite having limitations, are particularly useful and necessary to assess model performance with respect to extremes. We discuss proper performance measures for comparing distributions of model output against corresponding distributions from data products that are flexible and robust enough to handle the particular aspects of extremes such as limited data availability. The new measures are applied to evaluate CMIP5 and CMIP6 projections of extreme temperature indices over Europe and North-America against the HadEX2 data set as well as the ERA5 and ERA-Interim reanalyses. Several models perform well to the extent that when compared to the HadEX2 data product, these models' performance is competitive with the performance of the reanalysis. While the model rankings vary with region, season and index, the model evaluation is robust against changes in the grid resolution considered in the analysis.