Which climate models capture the variability and forced response in observed temperatures: a large ensemble comparison

Nicola Maher, Laura Suarez-Gutierrez, and Sebastian Milinski
Max Planck Institute for Meteorology, The Ocean in the Earth System, Hamburg, Germany (nicola.mahe@mpimet.mpg.de)

We evaluate how large ensembles of ten coupled climate models represent the observed internal variability and response to external forcings in historical surface temperatures based on a novel methodological framework. This framework allows us to directly attribute whether discrepancies between models and observations arise due to biases in the simulated internal variability or rather in the forced response, without relying on assumptions to separate both signals in the observations. The largest discrepancies occur due to overestimated forced warming in some models during recent decades. The areas where most models, a maximum of nine, adequately simulate observed temperatures are the North Atlantic, Tropical Eastern Pacific, and the Northern Hemisphere land areas. In contrast, none of the models considered offers an adequate representation over the Southern Ocean. Our evaluation shows that CESM-LE, GFDL-ESM2M, and MPI-GE perform best at representing the internal variability and forced response in observed surface temperatures both globally and regionally.