Decadal predictive skill assessment - ensemble and hindcast sample size impact

Frank Sienz, Wolfgang Müller, and Holger Pohlmann

Max-Planck-Institut for Meteorology, Hamburg, Germany (frank.sienz@mpimet.mpg.de)

Hindcast, respectively retrospective prediction experiments have to be performed to validate decadal prediction systems. These are necessarily restricted in the number due to the computational constrains. From weather and seasonal prediction it is known that, the ensemble size is crucial. A similar dependency is likely for decadal predictions but, differences are expected due to the differing time-scales of the involved processes and the longer prediction horizon.

It is shown here, that the ensemble and hindcast sample size have a large impact on the uncertainty assessment of the ensemble mean, as well as for the detection of prediction skill. For that purpose a conceptual model is developed, which enables the systematic analysis of statistical properties and its dependencies in a framework close to that of real decadal predictions. In addition, a set of extended range hindcast experiments have been undertaken, covering the entire 20th century.