An evaluation of the CESM Decadal Climate Prediction Large Ensemble with application to regional studies

J. J. Rosa-Cánovas, M. García-Valdecasas Ojeda, P. Yeste, E. Romero-Jiménez, S. R. Gámiz-Fortis, Y. Castro-Díez, M. J. Esteban-Parra

Department of Applied Physics, University of Granada, 18071 Granada, Spain.

Corresponding author: jjrc@ugr.es

Abstract

The decadal climate prediction (DCP) is one of the major challenges addressed by the research community focused on climate studies during the last years. DCPs try **to fill the gap between seasonal-to-interannual predictions and multidecadal-to-centennial climate change projections** by taking advance of not only the forced climate change signal provided by boundary information, but also the initialization of the climate system components which exhibit longer memory, such as the ocean.

Climate modelling for DCP is a **very expensive activity in terms of computing resources** since many initialized experiments are needed to properly assess the predictive skill at such time scales. In the context of dynamical downscaling (DD), this problem becomes even more important. Hence, **the aim of this study is to evaluate the sea surface temperature (SST) from the Decadal Climate Prediction Large Ensemble (DPLE) and to explore the issue of reducing the number of ensemble members in consideration to make DD more affordable.**

The Decadal Prediction Largue Ensemble

➤ Initial dates:

Each DCP is initialised on November 1 and spans 122 months. A new experiment starts every year from 1954 to 2015, so there are 62 different initial dates.

➤ Members of the ensemble:

An ensemble of 40 members per initial date is generated by randomly perturbing the atmospheric initial conditions.

➤ Model components:

· **Atmosphere model**: Community Atmosphere Model, v5 (CAM5).
· **Land model:** Community Land Model, v4 (CLM4).
· **Ocean model:** Parallel Ocean Program, v2 (POP2).
· **Ice model:** Los Alamos Sea Ice Model, v4 (CICE4).

➤ Initialisation procedure:

The **ocean and ice components**, those exhibiting longer climate memory, were initialised with fields obtained from a **forced ocean–sea ice (FOSI) simulation** driven by atmospheric reanalysis data. Input data were mostly taken from the **Coordinated Ocean-Ice Reference Experiment (CORE)**.

Data and methodology

➤ DPLE SST data:

· **Ensemble members under study:** from 001 to 010, only those which provide variables to appropriately perform a DD simulation. The 40-member average (AVG) has been also calculated for comparative purposes.

· **Period of study:** from 1979 to 2008, encompassing 30 decades initialised every year.

· **Drift correction:** lead time-dependent drift corrected with ERA5 SST data as reference.

➤ Reference SST data for analysis:

· Extended Reconstructed Sea Surface Temperature, v5 (ERSSTv5).
· Hadley Centre Sea Ice and Sea Surface Temperature, v1 (HadISST1).

➤ Methodology:

Members from 001 to 010 (along with AVG) have been **ranked in terms of their skill in reproducing the reference SST**. It has been assessed by using **two metrics**: the root mean squared error (**RMSE**) and the anomaly correlation coefficient (**ACC**), both averaged over the **EUROCORDEX region.** Since the purpose is to address the performance at decadal time scales, the study has been mainly focused on the **lead time ranging 2-9 years** from the start date. The first year has been separately analysed to avoid skill enhancement due to initialisation.

Results: Ranking the ensemble members

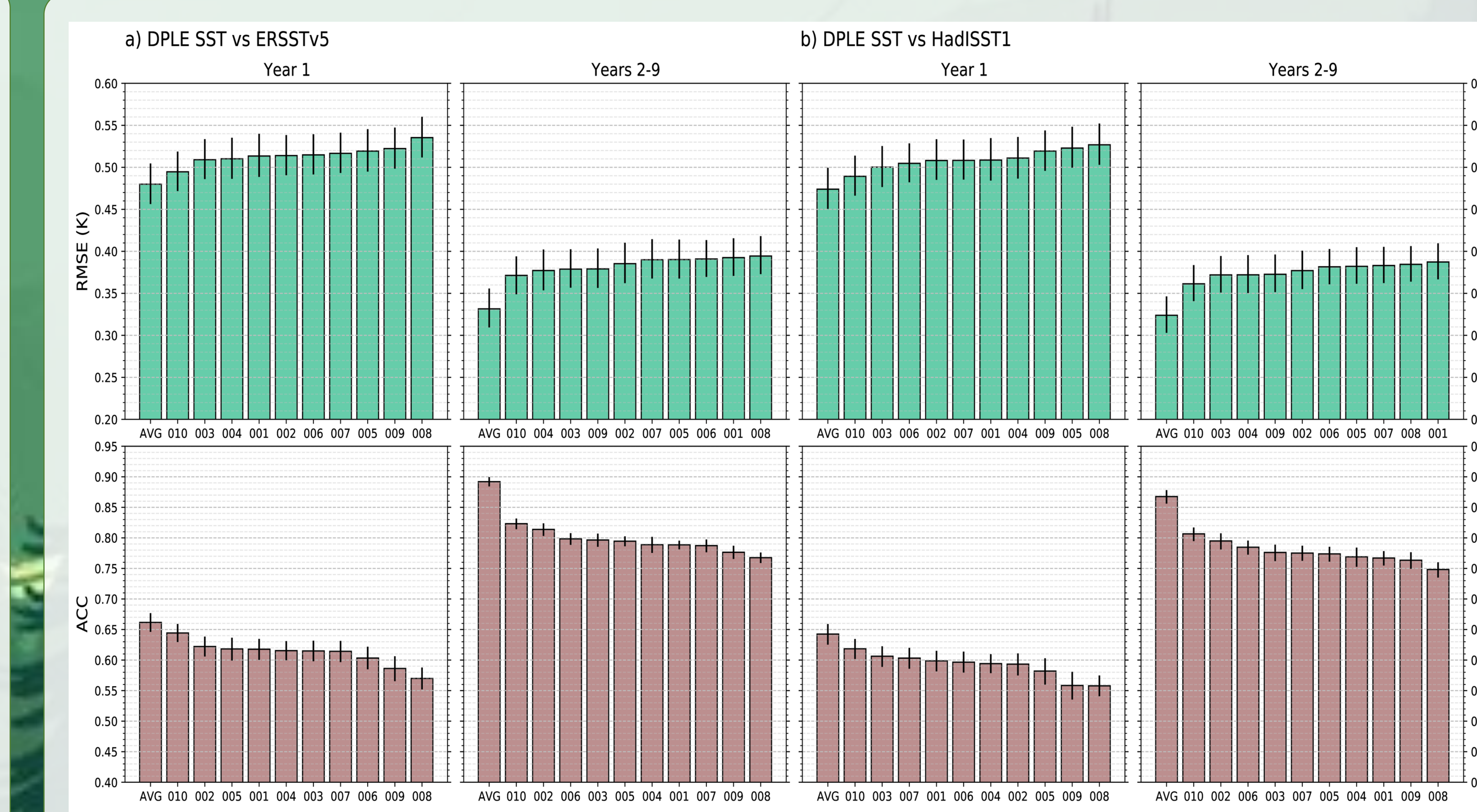

Figure 1: Ensemble members ranked in terms of their skill (descending order). Bars denote the spatially weighted averages for each metric over the EUROCORDEX region. Vertical black lines indicate the 95 % confidence interval for the sample averages (10000 bootstrapped samples). In a), ERSSTv5 is the reference dataset; in b), HadSST1 is.

Results: RMSE and ACC spatial fields for some members

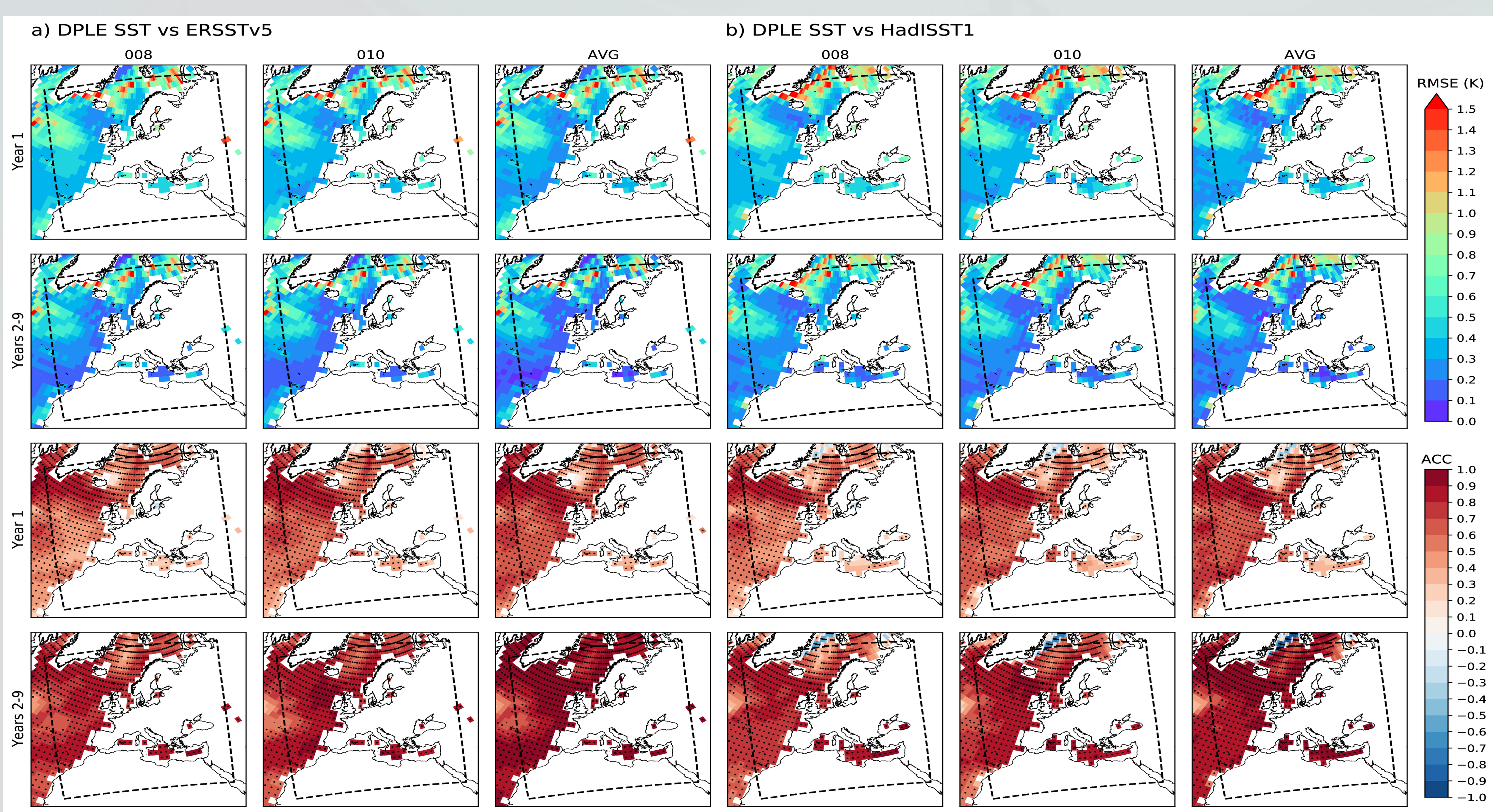

Figure 2: Skill metrics for member 010 (the most skilful), member 008 (the least skilful) and AVG. Dashed lines delimit the EUROCORDEX region. In ACC plots, black dots denote statistical significance at the 95 % confidence level for positive values (10000 bootstrapped samples). In a), ERSSTv5 is the reference dataset; in b), HadSST1 is.

Discussion and concluding remarks

➤ The performance in the years 2-9 is better than in the first year.

Since the average over the interval spanning years 2-9 from start date captures the observed warming trend better than the first simulated year, it was expected the highest scores were observed in the former case.

➤ Differences among member performances are not very remarkable.

Spatially averaged metrics in Figure 1 are similar when member performances are individually studied. Although differences are small, some patterns can be identified. The 40-member average clearly outperforms the whole set of single members, as expected, especially in ACC. The member 010 is the single member which shows the largest skill in reproducing SST magnitude and variability. On the other hand, the member 008 presents the lowest scores in (almost) all cases. Nevertheless, confidence intervals of single members use to overlap each other. The same situation is depicted by Figure 2.

A subset which encompasses the "best", the "worst" and an "intermediate" ensemble members could help to qualitatively address the uncertainty issue in a DD framework. As no member clearly outperforms the rest, at least in this case, this study could be extended to other long-memory fields. Anyway, this subset cannot replace at all the DD of the whole ensemble if a quantitative assessment of the ensemble uncertainty is required.

➤ Decreasing skill over Greenland outskirts.

Even though this topic falls out of the scope of this work, it is worth mentioning that the skill in reproducing observed SST exhibits drops in spurious points, especially when HadISST1 is the reference dataset, in the northernmost regions.

Acknowledgements:

- JJRC acknowledges the Spanish Ministry of Science, Innovation and Universities for the FPI fellowship (PRE2018-083921).
- This research has been carried out in the framework of the project CGL2017-89836-R, funded by the Spanish Ministry of Economy and Competitiveness with additional FEDER funds.