# SPECIAL PROJECT FINAL REPORT

All the following mandatory information needs to be provided.

| | |
|---|---|
| **Project Title:** | Understanding inter-annual to decadal predictability in the EC-Earth3 model and the potential benefits from perfect initialisation |
| **Computer Project Account:** | SPESICCF |
| **Start Year - End Year :** | 2024 - 2024 |
| **Principal Investigator(s)** | Etienne Tourigny |
| **Affiliation/Address:** | Barcelona Supercomputing Center – Earth Sciences Department<br>Plaça Eusebi Güell, 1-3, 08034 Barcelona (Spain) |
| **Other Researchers (Name/Affiliation):** | BSC: Markus Donat, Roberto Bilbao, Rashed Mahmood, Pablo Ortega Montilla, Miguel Castrillo, Vladimir Lapin, Aude Carreric<br>DMI: Rashed Mahmood |

The following should cover the entire project duration.

## Summary of project objectives

A major challenge for the advance of seasonal-to-decadal climate predictions is related to model initialization towards the observed state of the climate. In addition to imperfect estimates of the observed initial state due to limited observations, current initialization methods can cause 'initialization shocks' (1,2). We will perform a set of coupled climate model simulations to examine predictability in the absence of initialization shocks and model drift, and avoiding imperfect assimilation of the observed climate state (due to lack of observations), using the EC-Earth3 climate model. To this end, simulations will be started each year from 1960 to 2014 of a historical climate simulation covering the 20th and early 21st century. These idealised predictions will then be compared to existing prediction experiments initialised with (imperfect) observations, and we will estimate to what extent the model-predictability is affected by those initialisation-related shortcomings used for previous estimates, and whether perfect initialisation can avoid some of the inconsistencies between model and real-world predictability.

## Summary of problems encountered

There were no issues and we used all computing resources allocated for the project.

## Experience with the Special Project framework

Submission and reporting is well managed, and technical support for managing new users and fixing issues is very satisfactory.

## Summary of results

Resources allocated in this project were used to perform a perfect model initialization experiment using the EC-Earth3 model in which the model was initialized using climate states from the model's own historical simulation. For this, we previously carried out a historical simulation in order to save the climate states for initializing the climate predictions. This historical simulation was also used as reference when evaluating the perfect model predictions. Performing a new historical simulation (instead of using existing CMIP6 historical simulation) was necessary in order to use the same version of the model that was subsequently used for the perfect model initialized predictions. The initialized climate predictions were then performed using climate states from the reference historical simulation by adding small perturbations to the atmospheric initial conditions. We performed eleven year long hindcasts starting every year from 1960 to 2014 using CMIP6 historical forcings until 2014 and SSP2-4.5 from 2015. The ensemble size of the perfect model prediction consists of ten members. A detailed description about these model simulations have been documented in Mahmood et al. (2025a).

This so-called "perfect model" initialized prediction experiment provided unique opportunities to study some of the important issues related to initialized climate predictions. Since both the forecasts and the verification data come from the same model, the experiment eliminates model bias, initialization shock, and climate drift that affect real-world decadal predictions. Using this experiment we studied one of the most puzzling problems in climate predictions, the so-called "signal to noise paradox (or SNP in short)" (Scaife and Smith, 2018) in which the model predicts observation better than predicting itself despite physical inconsistencies between model and the observations. This is often attributed to model deficiencies. In Mahmood et al. (2025a), we highlighted the SNP paradox may have a different explanation which is related to the common analysis practices in which data is concatenated from individual independent initialized prediction (also called start-dates) to formulate a time-series. Such concatenation of the initialized predictions breaks the autocorrelation of the resulting time-series. We highlight here some of the findings from Mahmood et al. (2025a) here by identifying change in the autocorrelation of model members with respect to the autocorrelation of the reference and how this affects individual model member correlation with ensemble mean. We also highlight how this affects the ratio of predictable components (RPC) which is used to diagnose the signal-to-noise paradox.

We first evaluate the (potential) skill of the different ensembles in simulating near-surface air temperature (Figure 1) and sea level pressure (Figure 2). We find high ACC values for the Perfect_Init over most of the globe, suggesting that the perfectly initialized ensemble is highly skillful in predicting its reference run (Figure 1a). The skill in Perfect_Init extends to most regions of the globe including some regions where the hindcasts initialized with observations with the same EC-Earth3 model do not show much skill, such as the northeast and southeast Pacific and several land regions (cf Figures 1a,d). Similarly, for sea level pressure (SLP), we find that the Perfect_Init experiment shows higher and widespread skill (in predicting its reference run) over many regions of the globe compared to the skill of the real-world predictions (cf Figures 2a,d). No skill is found in the real-world predictions for SLP over the tropical and subtropical Indian and Atlantic Oceans and the bordering continental regions. Perfect_Init, on the other hand, shows positive ACC values over most of these regions although at some locations, ACC is not statistically significant at the 95% confidence level. These results indicate that there may be some decadal-scale predictability in the climate system of the EC-Earth3 model which is currently not being captured by the real-world (i.e., ClimPred_Init) predictions. To develop an understanding of the behavior of the different predictions, we first inspect the time series at one arbitrary grid point in the Pacific region as shown in Figure 3. It is immediately apparent how the individual ensemble members show very different characteristics compared to both the ensemble mean and the evaluation references—for both the Perfect_Init and the ClimPred_Init. The initialized ensemble members appear much noisier compared to both the transient reference simulation and also the observational reference time series.

The results in Figure 4 (panels a and c) show that when the difference in autocorrelation of individual members with respect to the autocorrelation of a reference is large the member correlation with ensemble mean is relatively much lower and the RPC values are larger (panels b and d). These results show that the model deficiencies may not just be the sole cause of the emergence of signal to noise paradox in initialized prediction. A detailed analysis and description of these results can be found in Mahmood et al. (2025a).

This perfect model experiment provides a unique opportunity to study many other different aspects of climate predictions. Therefore, simulations from this experiment are still being analysed, for example, to understand model specific predictability limits of extreme climate events and the results from this analysis would be reported in a peer-reviewed journal.

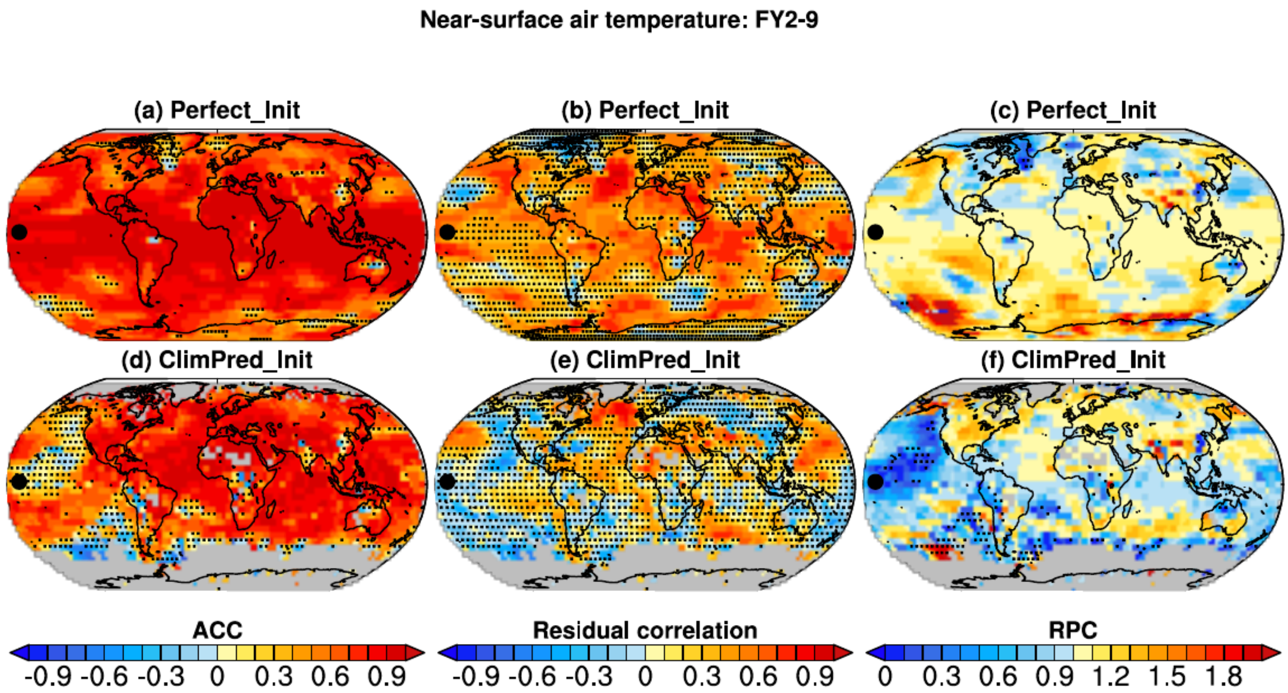


Figure 1. Skill assessment of different ensembles in simulating near-surface air temperature for the average of forecast years 2–9 based on (left) ACC, (center) residual correlation, and (right) the ratio of predictable components. The results for (top) Perfect_Init and (bottom) ClimPred_Init. The stippling represents regions where (a),(d) ACC and (b),(e) residual correlations are not statistically significant at the 95% confidence level. For (c),(f) RPC, the regions where ACC is negative are stippled. The black markers indicate the grid point used for plotting near-surface air temperature and SLP time series in Figure 3. **Figure from Mahmood et al. (2025a)**.

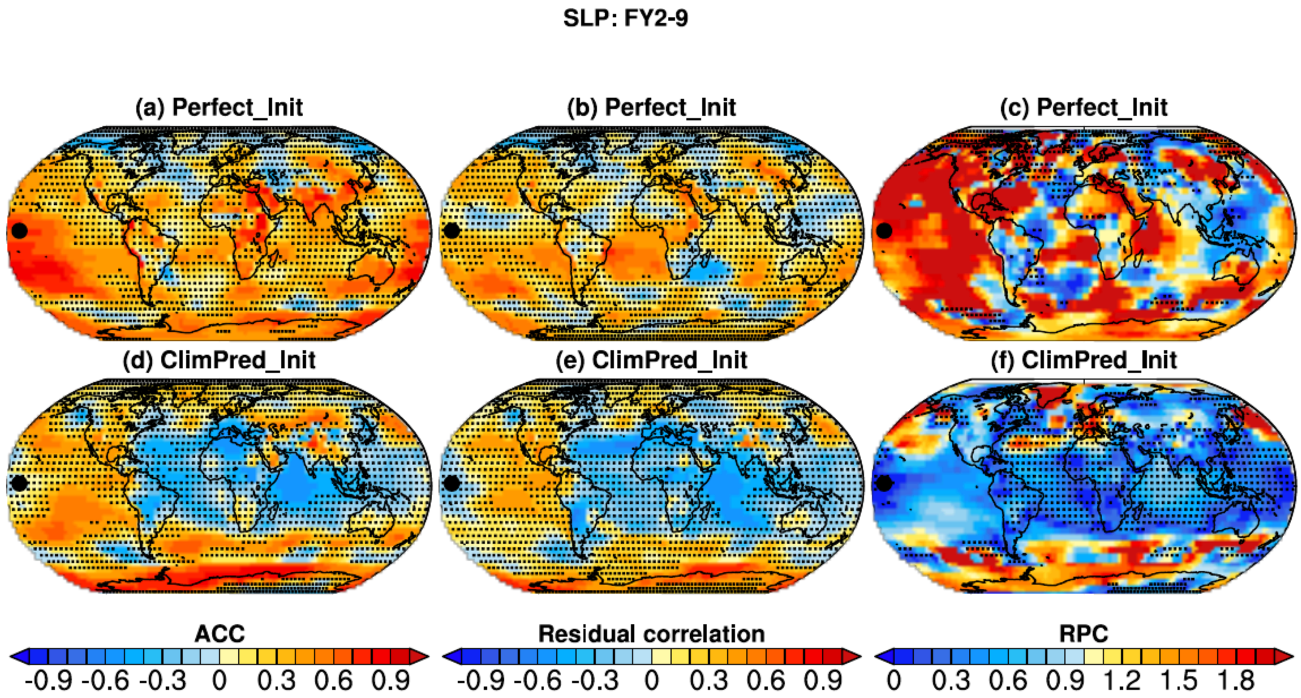


Figure 2. Same as Figure 1 but for sea level pressure. **Figure from Mahmood et al. (2025a)**.

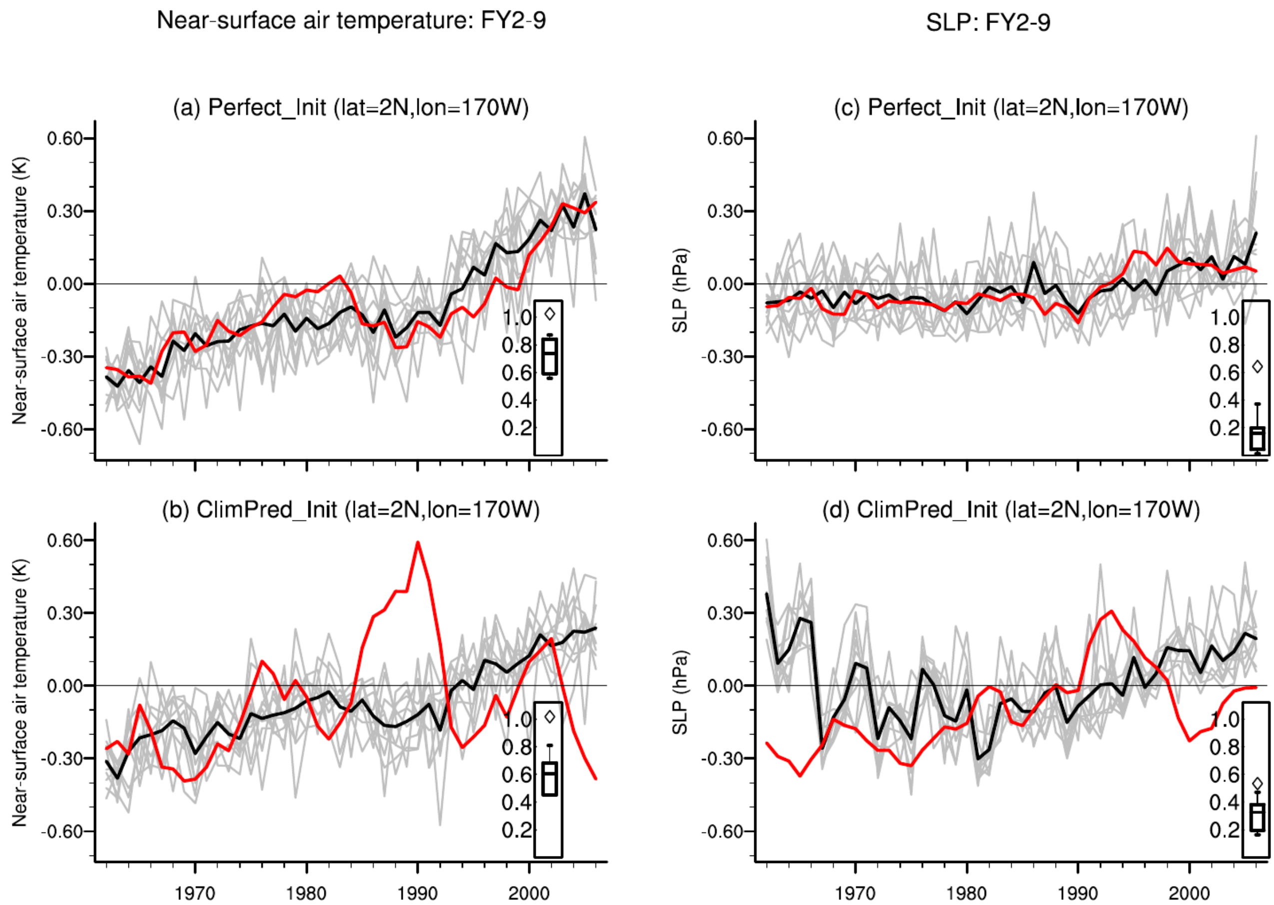

Figure 3. Time series of 8-yr average (a),(b) near-surface air temperature and (c),(d) SLP anomalies at one arbitrarily chosen grid point in the Pacific (2°N, 170°W). The gray lines represent the average of forecast years 2–9 in the individual members; the ensemble mean is shown in black. Red lines represent the evaluation reference (i.e., reference historical simulation for Perfect_Init and observations for ClimPred_Init). The boxplots show the ratio of lag-1 (i.e., 1-yr lag in start dates) autocorrelations between ensemble members and the evaluation reference. The diamonds represent the ratio of lag-1 autocorrelations between the ensemble mean and the evaluation reference. The years on the *x* axis represent the first year of the forecast years 2–9. **Figure from Mahmood et al. (2025a)**.

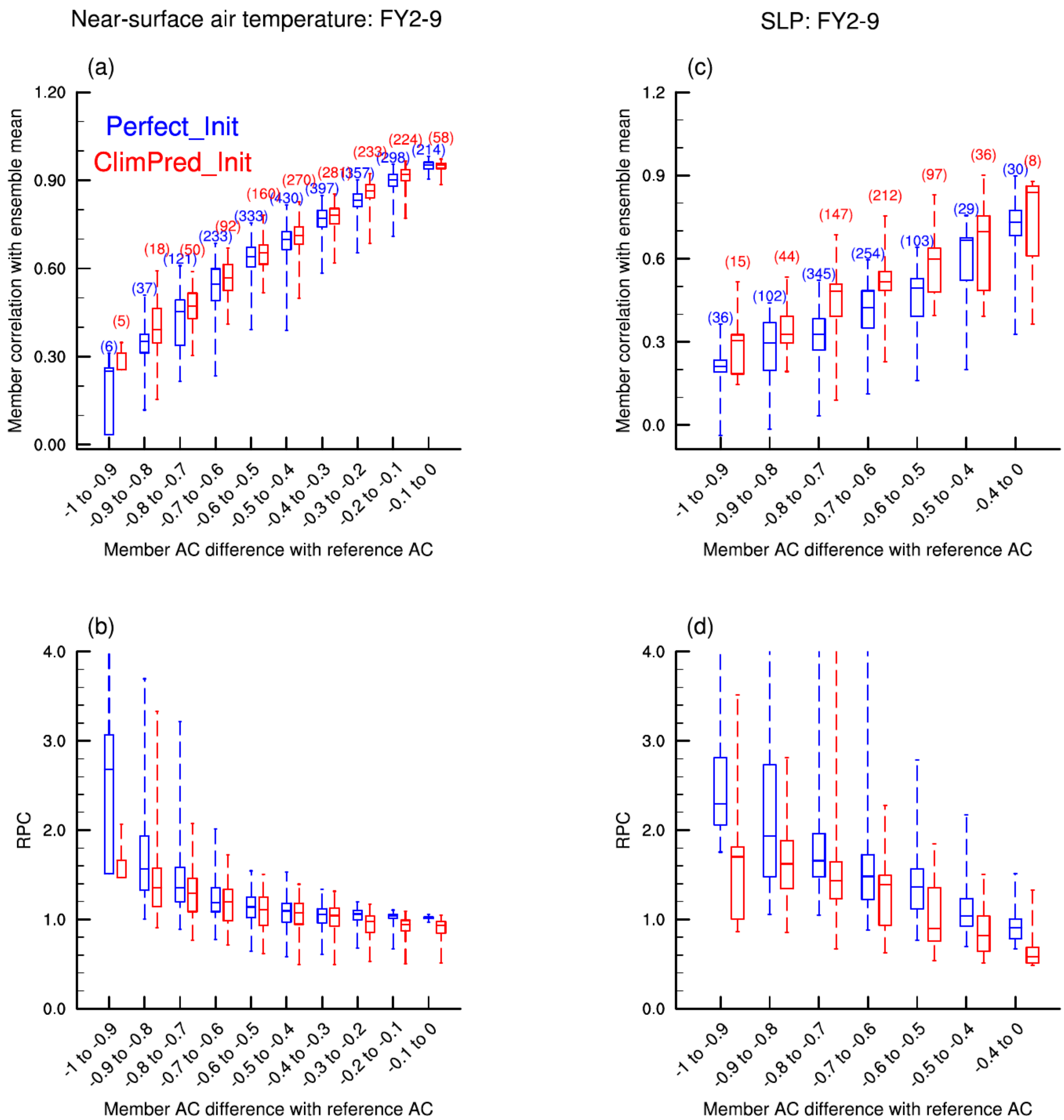

**Figure 4**. (a),(c) Model member correlation with the ensemble mean (shown on the y axis) binned according to the magnitude of the median difference in lag-1 autocorrelation of individual members with the autocorrelation of the reference over the globe for (a) near-surface air temperature and (c) SLP. (b),(d) RPC for the same groupings of regional autocorrelation differences for (b) near-surface air temperature and (d) SLP. The number of grid points included in each bin is shown at the top of the boxes, while bin ranges are shown on the x axis. Only grid points where ACC is positive and statistically significant at the 95% level have been used. The results shown are based on the average of forecast years 2–9. Perfect_Init (shown in blue) represents perfect model predictions while ClimPred_Init (shown in red) represents real-world initialized predictions. **Figure from Mahmood et al. (2025a)**.

## List of publications/reports from the project with complete references

A.A. Scaife, Smith, D., npj Clim Atmos Sci 1, 28 (2018). https://doi.org/10.1038/s41612-018-0038-4
Mahmood, R., M.G. Donat, F.J. Doblas-Reyes and E. Tourigny (2025a). A perfect-model perspective on the signal-to-noise paradox in initialized decadal predictions. Journal of Climate, 38, 2841–2851, https://doi.org/10.1175/JCLI-D-24-0381.1.
Mahmood, R., M.G. Donat, R. Bilbao, P. Ortega, V. Lapin, E. Tourigny and F.J. Doblas-Reyes (2025b). Multi-decadal initialized climate predictions using the EC-Earth3 global climate model. Earth System Dynamics, 16, 1923–1934, https://doi.org/10.5194/esd-16-1923-2025.
Mahmood R., Donat M.G., Christiansen B., Doblas-Reyes F. J., (2026), Spurious drifts in the ITCZ and associated atmospheric circulation in decadal climate predictions, Geophysical Research Letters, in review.

## Future plans

We plan to implement a new seasonal-to-decadal forecasting system using EC-Earth4, the next generation version of EC-Earth planned for contribution to CMIP7. We expect to apply for a Special Project to perform initial hindcasts useful for the development of the system.