SPECIAL PROJECT PROGRESS REPORT

All the following mandatory information needs to be provided. The length should *reflect the complexity and duration* of the project.

Reporting year	July 2023-June 2024
Project Title:	BONSAI (Boosting eNsemble Size for Advanced Insights into climate predictability)
Computer Project Account:	SPITBEAL
Principal Investigator(s):	Alessio Bellucci
Affiliation:	Consiglio Nazionale delle Ricerche, Istituto di Scienze dell'Atmosfera e del Clima (CNR-ISAC)
Name of ECMWF scientist(s) collaborating to the project (if applicable)	………………………………………………………………… …………………………………………………………………
Start date of the project:	15/02/2022
Expected end date:	31/12/2024

Computer resources allocated/used for the current year and the previous one

(if applicable)
Please answer for all project resources

		Previous year		Current year	
		Allocated	Used	Allocated	Used
High Performance Computing Facility	(SBU)	8 Millions	4.8 Millions	7 Millions	393000
Data storage capacity	(Gbytes)	48000	756	80000	825

Summary of project objectives (10 lines max)

Special Project BONSAI aims to push the boundaries of large climate prediction ensembles by developing a prototype decadal prediction system based on a simplified dynamical model. By reducing the complexity of the model, including while preserving the essential elements necessary to replicate the fundamental characteristics of the observed climate and its variability, we can significantly increase the size of the forecast ensembles. This advancement allows for an order of magnitude expansion, enabling us to transition from the current standard resolution/complexity decadal prediction's ensemble size of around 10 members to hundreds of members. This yields several benefits. Firstly, it enhances the sampling of uncertainty related to initial conditions, resulting in more reliable predictions. Secondly, it enables more effective suppression of unpredictable noise, thus improving the signal-to-noise ratio and enhancing forecast skill. Ultimately, this enhancement contributes to a better understanding of the predictable aspects of the climate system and improves our ability to make accurate long-term climate projections.

Summary of problems encountered (10 lines max)

No specific problems were encountered during this reporting period.

Summary of plans for the continuation of the project (10 lines max)

In the following months, the set of decadal hindcasts will be further extended, so as to complete the 1991-2000 decade. The regional scale analysis of BONSAI DPS skill will be also extended. The promising results obtained over the North Atlantic subpolar gyre area motivate a more extensive regional skill assessment, focusing on specific sectors and state variables. In particular, the sources of predictive skill found in the North Atlantic area will be assessed by an in depth analysis of the simulated Atlantic Meridional Overturning Circulation (AMOC). Also, the analysis of specific climate extremes, initiated during the 3rd project year, will continue.

List of publications/reports from the project with complete references

Results from BONSAI have been reported in:

Ruggieri, P., Abid, M. A., Garcia-Serrano, J., Grancini, C., Kucharski, F., Pascale, S., and Volpi, D.: SPEEDY-NEMO: performance and applications of a fully-coupled intermediate-complexity climate model, EGU General Assembly 2024, Vienna, Austria, 14–19 Apr 2024, EGU24-14831, https://doi.org/10.5194/egusphere-egu24-14831, 2024.

Mele L. and Ruggieri P., Advances in understanding extreme meteorological events: a review of the UNSEEN methodology and its applications, in preparation.

Summary of results

In the present document, we report on the advancements made during the third reporting period (July 2023-June 2024) of the SP BONSAI. Having completed the set-up of the BONSAI Decadal Prediction System (hereafter BONSAI DPS) and performed an explorative set of 100-member ensembles of initialized decadal hindcasts in the project's 2nd year, the following year (subject of the present report) was mainly dedicated to a) extend the set of decadal hindcasts targeting the 1980-1990 decade, and b) assess the skill of BONSAI, targeting two regional case studies: sea-surface temperature in the North Atlantic Subpolar Gyre (Section 3) and heat extremes over the Mediterranean area (Section 4).

1. Extended set of decadal hindcasts

The set of decadal hindcasts was substantially extended. The 1980-1990 decade (with 100-member ensembles of 10-year predictions, for each start date) was completed. This effort required the use of 4.8 million SBUs on ATOS HPC facility. This extended dataset provides a solid basis for a first evaluation of BONSAI, with a total of 12 start-dates (1980 to 1990, and 1995) x 100 members x 10 years = 12,000 simulated years, now available for analysis. In the following months, this set will be further extended to complete the 1991-2000 decade.

2. **Global Mean SST bias attribution**.

A prominent warm bias in the global mean SST fields (GMSST), evident during boreal winter months (see 2nd Year Report, Fig. 2) emerged during the 2nd year of the project. As anticipated in the 2nd Year report, the origins of this bias were investigated in more detail during the early part of project year 3.
An initial hypothesis was formulated, attributing the bias to the use of a motionless atmosphere with standard thermal and moisture conditions for the initialization of SPEEDY. While the use of this initialization strategy was justified by the rapid dynamical adjustment of the motionless atmosphere after initialization, some concerns were raised that the imbalance at the air-sea interface might lead to some spurious equilibrium, ultimately causing the surface ocean bias. To test this hypothesis, a single member of the 1980 start-date ensemble was re-run starting from a more realistic SPEEDY initial state taken from a present day simulation (i.e., with the atmosphere dynamically spun-up and with thermodynamic conditions consistent with the BONSAI climatology). However, the test run showed no discernible differences with respect to any other member of the ensemble, revealing that the alternative initialization had no effect on the model bias (not shown). Therefore, a different analysis was performed, computing the mean SST for different domains, including Tropics (35S-35N), northern hemisphere extratropics (35N-90N), and southern hemisphere extratropics (90S-35S). Comparing the mean SST time series (taking the 1980 initialized ensemble as an illustrative case) for the three selected domains, it emerges that the bias in the GMSST originates in the southern hemisphere, showing a ~5-6°C seasonal range, compared to the observed 4°C range (ORAS4; Figure 1), dictated by the boreal winter months, consistent with the GMSST bias. Tropics and extra-tropical northern Hemisphere mean SSTs, on the other hand, show a considerably reduced bias.

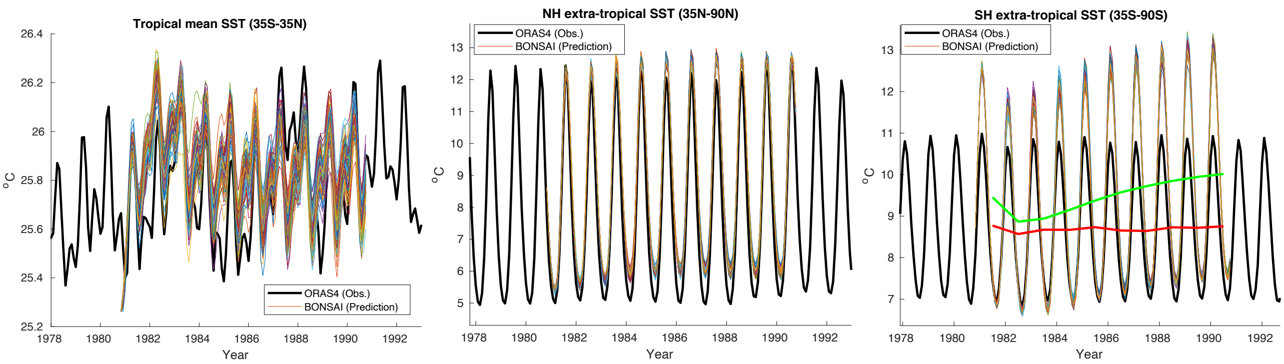

Fig. 1. Mean (area-weighted) SST evolution for the 100-member hindcast ensemble initialized on November 1980 (color curves) for Tropics (35S-35N) (left), Northern Hemisphere extratropics (35N-90N) (middle) and Southern Hemisphere extratropics (35S-90S) (right). The corresponding time series from ORAS4 reanalysis (member 1) is also shown (black solid curve).

3. Regional skill assessment: the North Atlantic Subpolar Gyre case

The North Atlantic Subpolar Gyre (SPG) is an area where initialized decadal hindcasts performed with state-of-the-art climate models generally reveal significant predictive skill (Yeager et al., 2012; Msadek et al., 2014; Robson et al., 2018). This makes it a particularly suitable study area to test the predictive ability of BONSAI.

In this section, the system's skill in predicting the SPG SST evolution is evaluated, against the ORAS4 reanalysis as a verification data set. For this purpose, a simple index (hereafter SPGI) is defined as the basin mean SST computed over the [60º-10ºW, 50º-60ºN] longitude-latitude box, a broad area in the northern North Atlantic representative of the SPG region.

Figure 2 shows raw, annual mean SPGI values from a sub-set of the initialized hindcast ensembles (shown with alternating colors to facilitate readability; for each start date, the ensemble mean forecast is also shown, in white) and the corresponding evolution of the observational counterpart (in black). Note that all ensemble members display a transient response after the initialization, characterized by a rapid temperature increase, reaching a peak around lead-years 3-4, followed by a more gradual decline. This transient behavior is common to all starting dates and reflects the well-known spurious model adjustment triggered by the full-value initialization procedure, also referred to as "coupling shock" (Balmaseda et al., 2009).

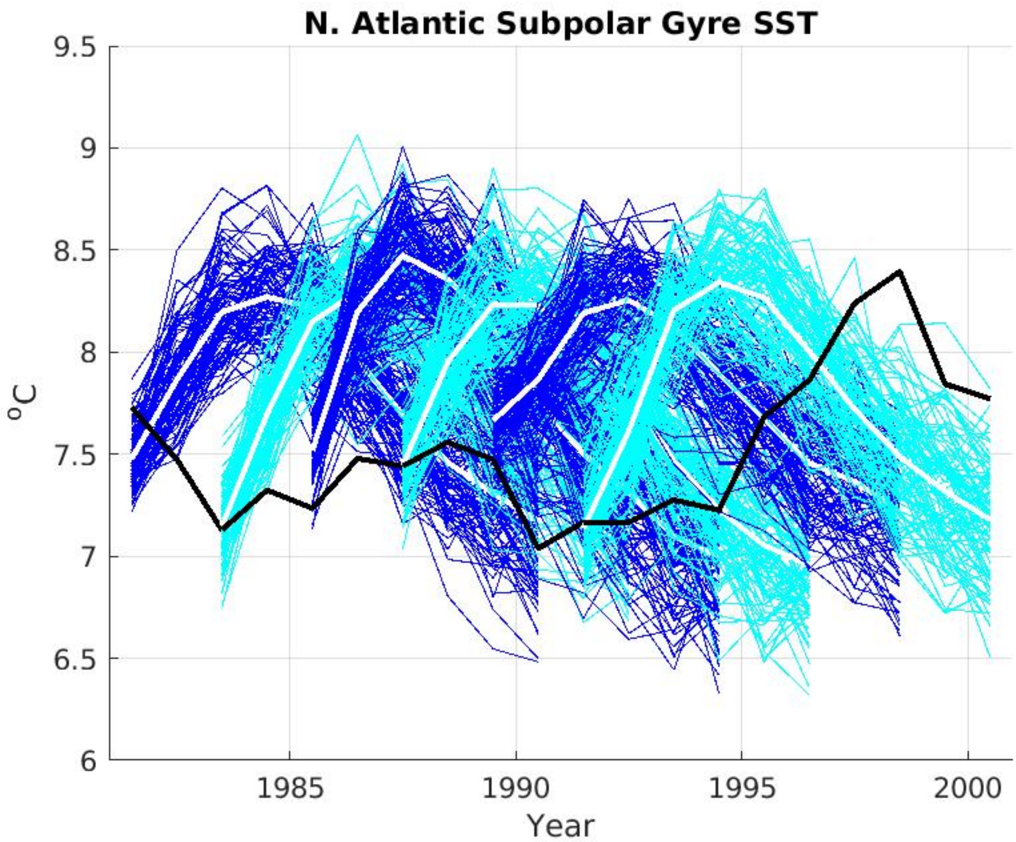

Fig. 2. Annual mean SST (ºC) in the SPG box, for the raw 100-member decadal prediction ensembles (blue and cyan curves; alternating colors for clarity; ensemble mean in white) and ORAS4 reanalysis (thick black curve). Note that only a sub-selection of the full set of performed hindcasts is shown.

To remove this spurious signal, a mean drift is calculated by averaging over the full set of decadal hindcasts, and subtracted from each ensemble member to obtain calibrated SPGI anomalies. The mean drift provides a lead-year dependent climatology representative of the 1980-2000 period. The corresponding 1980-2000 climatology is subtracted from ORAS4 values to obtain consistent observed anomalies. The simulated and observed SPGI anomalies are then shown in Figure 3, for the full set of decadal hindcasts. For each starting date, the ensemble member anomalies are shown (in red) together with the ensemble mean (in white) and ORAS4 reanalysis (in black). After de-drifting, the magnitudes of predicted SPGI anomalies show a reasonable consistency with observations, with the latter mostly lying within the range of simulated values. The ensemble mean forecasts show smaller amplitude variability compared to reanalysis, with the observed extreme values occurred in the mid-to-late-90s only marginally captured by the initialized hindcasts. However, for a few start-dates, the ensemble mean fluctuations appear to be in phase with the observed signal (notably, 1985, and 1988-to-1990).

Next, the skill associated with individual ensemble members is assessed, by looking at the frequency distribution of correlation values between single member realizations and the observed trajectory. In Figure 4, the correlation frequency histograms corresponding to different ensemble forecasts are shown. These reveal variously shaped distributions, suggesting that the predictability over the study area is subject to a certain degree of non-stationarity. In this respect, the level of symmetry characterizing a given distribution is a particularly insightful parameter: start-dates (or, equivalently, climate states) featuring a high symmetry in the correlation value distribution, show an equal likelihood associated with the occurrence of "good" (positive, significant correlations) and "bad" predictions (negative correlations). Large deviations from symmetry, on the other hand, are indicative of climate states proving to be either particularly predictable (positively skewed distribution) or unpredictable (negatively skewed distribution).

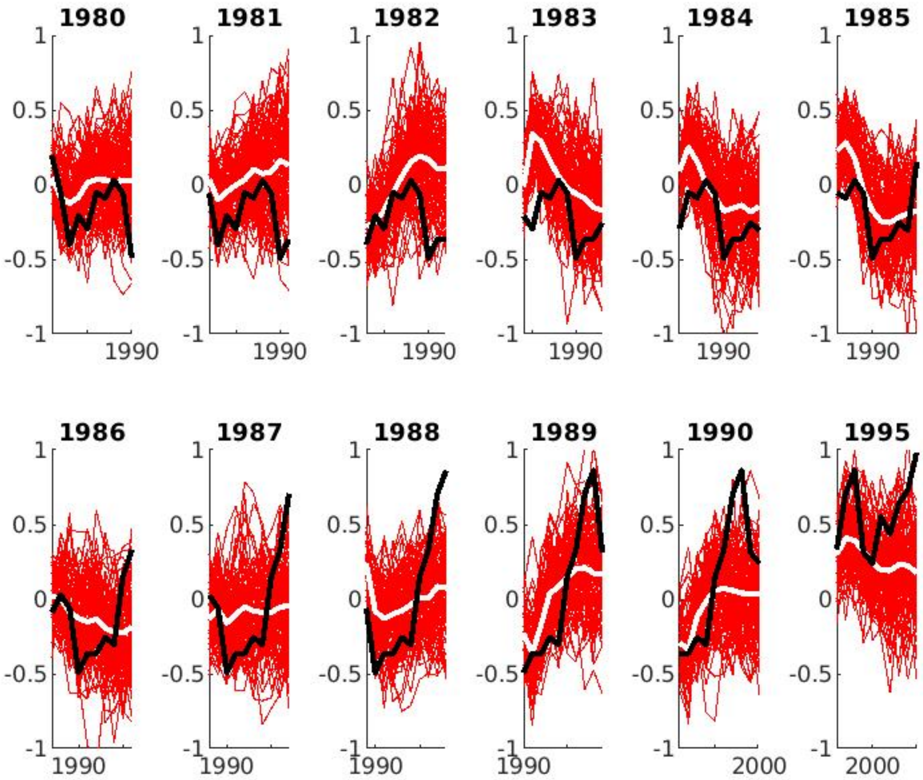

Fig. 3. Annual mean SST anomalies (°C) in the SPG box, for the drift-corrected decadal predictions (in red; ensemble mean in white) and ORAS4 reanalysis (thick black curve). The ORAS4 anomalies are computed relative to climatology over the reference period 1980–2000.

By scrutinizing the different start-dates in Fig. 4, it is evident the alternation between approximately symmetric and highly asymmetric (and positively skewed) distributions, while no occurrence of negatively skewed distribution is found. For example, correlations corresponding to start-dates 1980, 1981, 1982 and 1987 display a highly symmetric distribution, while start-dates 1983, 1985 and 1989 feature prominent modal peaks centered over correlation values exceeding 0.6-0.7, and relatively little or almost no occurrence of negative correlations. This result highlights the existence of "windows of opportunity" in decadal-scale climate predictability (Mariotti et al., 2020). According to this paradigm, there are specific epochs, and relative climate states, characterized by a higher degree of predictability, with a consistently larger occurrence of successful forecasts. In this specific SPG-based case study, early-80s appear to be a less predictable period, compared to late-80s/early-90s. It is worth noticing how this feature of the climate system can only be robustly assessed using BONSAI-like experimental settings, allowing to perform 100s-member ensembles of initialized reforecasts with the same climate model.

Based on the correlation distributions shown in Figure 4, it is possible to rank the individual ensemble members, and select the forecast sub-ensembles featuring the highest skill. In Figure 5, the mean forecast computed over the upper three, best ranking ensemble members are shown for each start-date. In line with what previously suggested, the largest discrepancies between predicted and observed SSTs occur during the early 80s, while there is a clear skill improvement emerging in the late 80s/early 90s, with most of the selected forecasts capturing years in advance the timing of the mid-90s rapid warming.

The analysis of SST prediction over the SPG in BONSAI DPS — which has a considerably lower resolution and degree of complexity than any model currently used in operational decadal predictions (Hermanson et al. 2022) — suggests that SST variability in the North Atlantic ocean can be predicted years in advance provided that the ocean component is accurately represented and initialized, with a relatively marginal role for other components of the climate system. A more comprehensive assessment of predictive skill over the SPG region in BONSAI will be completed in the following months. In particular, the relative roles of the ocean internal variability (including the AMOC and ocean heat transport) and external radiative forcings in determing the overall BONSAI skill will be assessed.

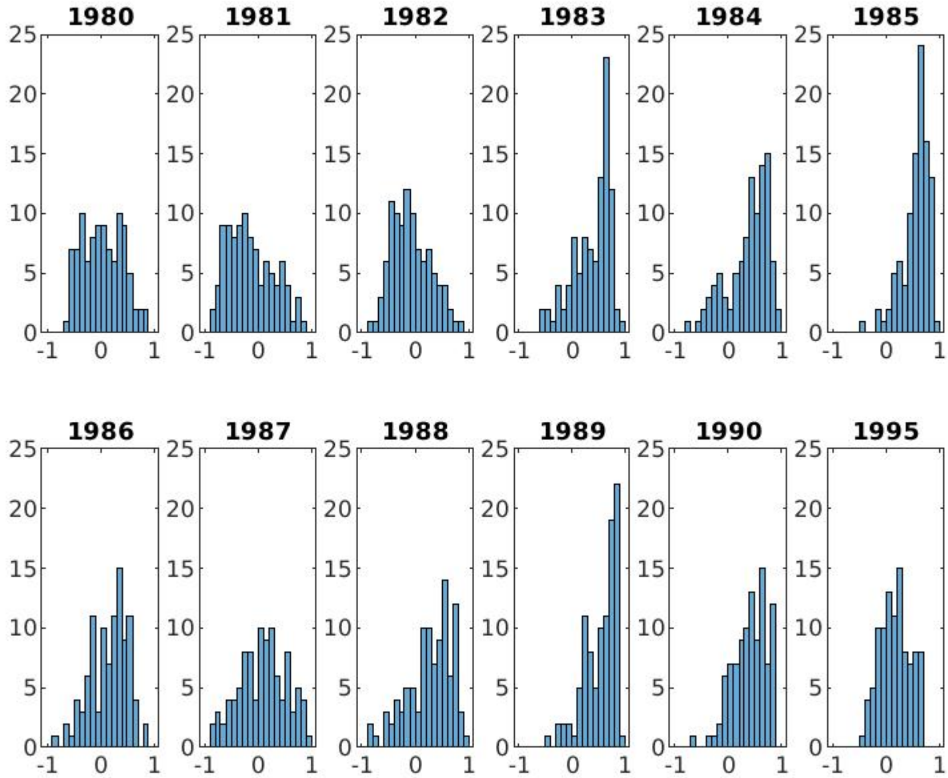

Fig. 4. Frequency distributions of Pearson correlation values between predicted and observed SST anomalies over the SPG box, for different start-dates. Correlation values in each histogram are binned into 0.1-wide intervals.

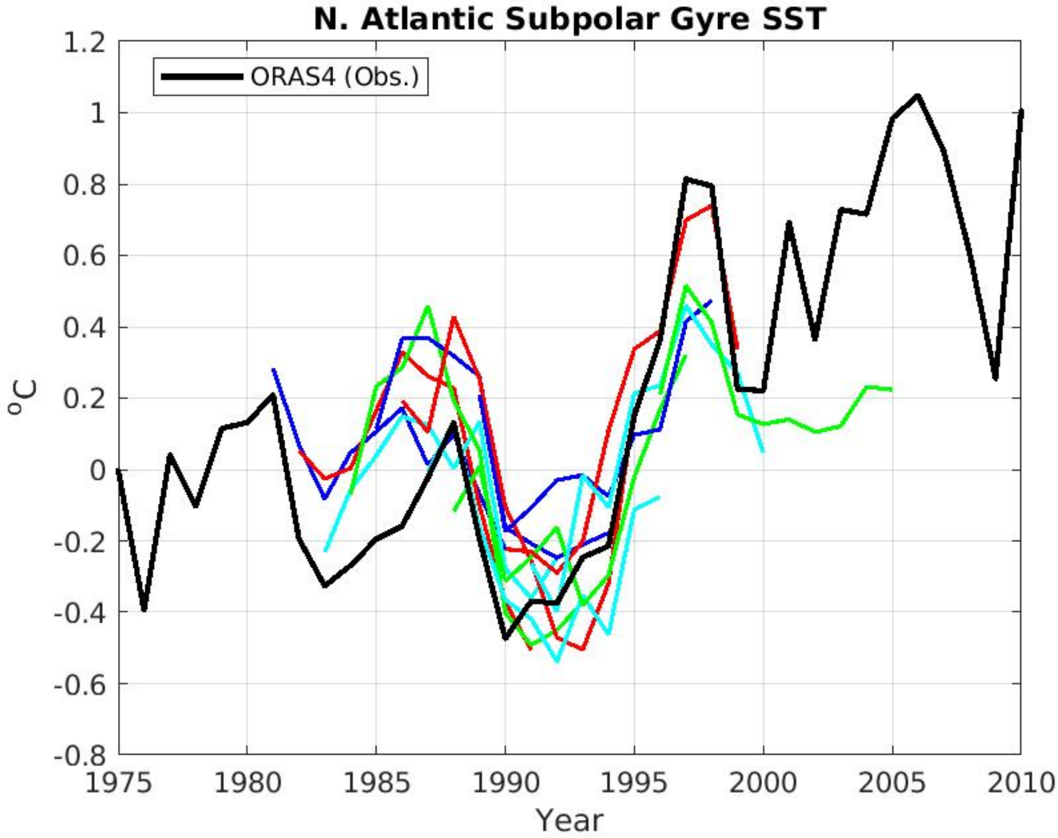

Fig. 5. Annual mean SST anomalies (°C) in the SPG box, for the drift-corrected decadal prediction (mean forecast over the upper-three best scoring ensemble members; alternating colors for clarity) and ORAS4 reanalysis (thick black curve). The ORAS4 anomalies are computed relative to climatology over the reference period 1980–2000.

4. Heat extremes in the Mediterranean.

The hindcasts produced so far in BONSAI were exploited in an ongoing study to characterize heat extremes over the Mediterranean region (Mele and Ruggieri, manuscript in preparation). The purpose of this study was to test the applicability of climate predictions to simulate plausible extreme events. As part of this effort the BONSAI hindcasts were used to compare temperature extremes in the Emilia-Romagna region (Italy; blue box in Fig. 6) simulated by the model with reanalysis data. Figure 6 provides an insight into the ability of the model to simulate the large-scale and synoptic conditions that led to observed extreme heat events in the target region encompassed by the blue box. The events on the right side of Figure 6 are identified as the most extreme cases in the ERA5 reanalysis (Hershbach et al., 2020), while those on the left side are chosen by defining extreme events in the hindcast (with a consistent definition) and by selecting those with a maximum correlation with the observed one in the Z500 field. This preliminary result reveals the ability of the model and the hindcasts to simulate realistic extreme events.

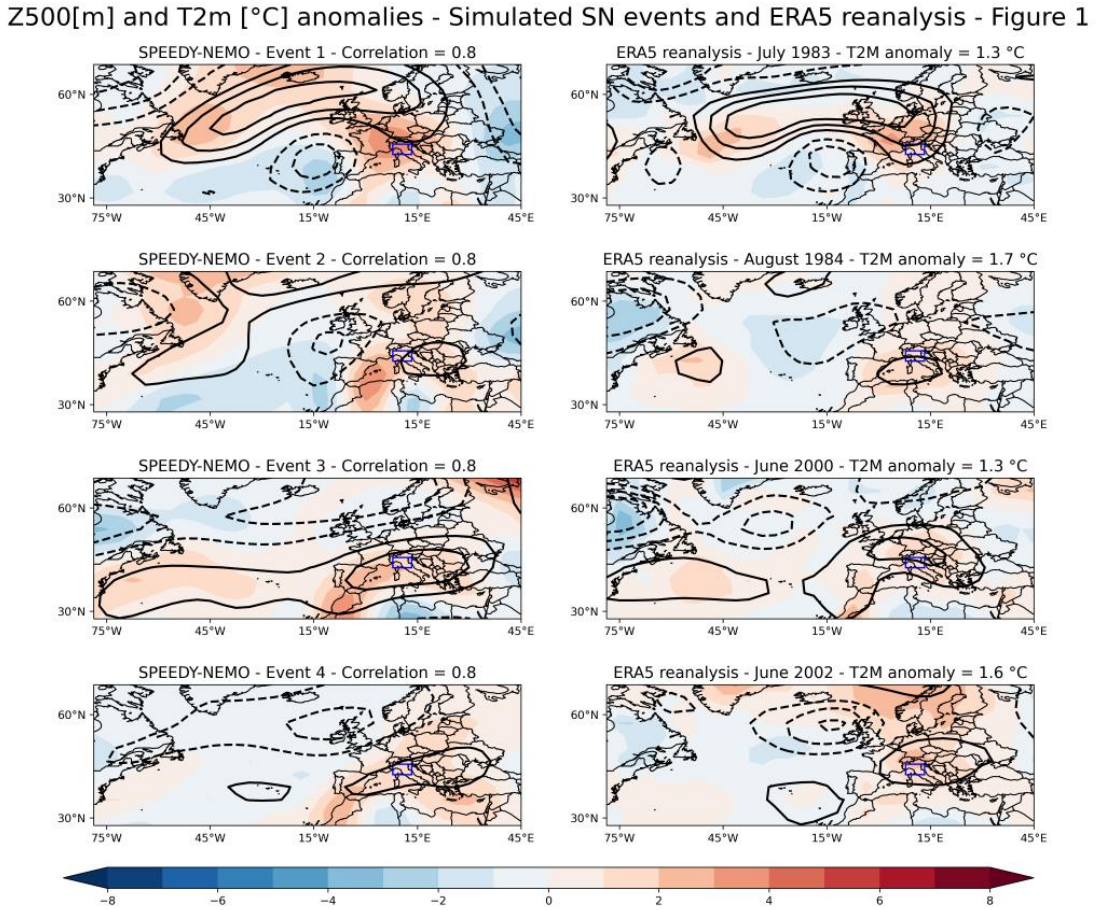

Figure 6. Contours of Z500 anomalies [m] and colormap of T2m anomalies [°C] in the clustering region: comparison between ERA5 reanalysis most severe events and SPEEDY-NEMO events with the highest Z500 anomaly correlation evaluated in this area. The 0 level Z500 anomaly contour has been omitted and the blue box shows the target region. Adapted from Mele L. and Ruggieri P. (in preparation).

5. References

Balmaseda, M. A., et al. (2009), Ocean initialization for seasonal forecasts, Oceanography, 22, 154–159.

Hermanson, L., and Coauthors (2022) WMO Global Annual to Decadal Climate Update: A Prediction for 2021–25. Bull. Amer. Meteor. Soc., 103, E1117–E1129, https://doi.org/10.1175/BAMS-D-20-0311.1.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. Quart. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10. 1002/qj.3803.

Mele L. and Ruggieri P., Advances in understanding extreme meteorological events: a review of the UNSEEN methodology and its applications, in preparation.

Msadek R, Delworth T, Rosati A, Anderson W, Vecchi G, Chang YS, Dixon K, Gudgel R, Stern W, Wittenberg A et al (2014) Predicting a decadal shift in North Atlantic climate variability using the GFDL forecast system. J Clim 27:6472–6496. doi:10.1175/ JCLI-D-13-00476.1

Robson, J., Polo, I., Hodson, D.L.R. et al. Decadal prediction of the North Atlantic subpolar gyre in the HiGEM high-resolution climate model. Clim Dyn 50, 921–937 (2018). https://doi.org/10.1007/s00382-017-3649-2

Yeager, S., A. Karspeck, G. Danabasoglu, J. Tribbia, and H. Teng (2012) A Decadal Prediction Case Study: Late Twentieth-Century North Atlantic Ocean Heat Content. J. Climate, 25, 5173–5189. https://doi.org/10.1175/JCLI-D-11-00595.1.