# REQUEST FOR A SPECIAL PROJECT 2019–2021

**MEMBER STATE:** Italy

**Principal Investigator**[1]**:** Danila Volpi

**Affiliation:** Institute of Atmospheric Sciences and Climate, National Research Council (ISAC-CNR)

**Address:** Via Gobetti 101,
40129 Bologna,
Italy

**Other researchers:**

ISAC-CNR: Susanna Corti
BSC: Francisco Doblas-Reyes

**Project Title:**

Innovative iNitialisation techniques for multi-annual ClImate PredIcTions (INCIPIT)

| | | |
|---|---|---|
| If this is a continuation of an existing project, please state the computer project account assigned previously. | **SP** ___________________ | |
| Starting year:<br>(A project can have a duration of up to 3 years, agreed at the beginning of the project.) | 2019 | |
| Would you accept support for 1 year only, if necessary? | YES ☒ | NO ☐ |

**Computer resources required for 2019-2021:**
(To make changes to an existing project please submit an amended version of the original form.)

| | | **2019** | **2020** | **2021** |
|---|---|---|---|---|
| High Performance Computing Facility | (SBU) | 36 million | 36 million | - |
| Accumulated data storage (total archive volume)[2] | (GB) | 50000 | 50000 | - |

*Continue overleaf*

1 The Principal Investigator will act as contact person for this Special Project and, in particular, will be asked to register the project, provide annual progress reports of the project's activities, etc.

[2] If e.g. you archive x GB in year one and y GB in year two and don't delete anything you need to request x + y GB for the second project year etc.

**Principal Investigator:** Danila Volpi

**Project Title:** Innovative iNitialisation techniques for multi-annual ClImate PredIcTions (INCIPIT)

# Extended abstract

***Objectives***

The main goal of this special project is to enhance the transfer of observed information to the model during the initialisation of multi-annual forecasts. This phase of the climate prediction process is of utmost importance, because it has been shown that a correct initialisation can improve the forecasts up to a few years ahead. However, the systematic errors of the models make this task challenging, because of the discrepancy between the observed and model mean climate. The main consequences are incorrect propagation of systems and a quick loss of the observed information. This project will therefore test innovative initialisation techniques explicitly designed to tackle specific limitations detected in the methods currently in use.

***Introduction and motivation***

One of the most relevant scientific challenges of multi-year prediction is dealing with the intrinsic systematic error of the models. This is due to the fact that models describe complex real-world processes simplifying through a set of equations and algorithms. The main consequence of the model error is the difference between the model and observed mean state. Such a difference complicates the initialisation task, which consists in transferring the observed information to the model at the initial time of a forecast. In multi-year predictions, after initialisation the model drifts away from the real-world attractor towards its own biased attractor. To account for such a bias, a-posteriori bias correction (Fuĉkar et al. 2014) needs to be applied, which although introduces additional errors in the forecast, is unavoidable for the forecasts to be usable. On top of the correction of the bias, multi-annual predictions also have the additional challenge of disentangling the smaller magnitude of climate signal to be predicted, from the initial drift to be removed (Smith et al. 2013). The ultimate way of reducing model biases a priori is through model improvements, such as the increase of spatial resolution, or improvements of the parameterisations of the model. However, with the current models, knowing their systematic errors, ad hoc initialisations and corrections could be implemented to improve the prediction skill.

The first initialisation method ever tested involves replacing the model state at the initialisation time with the best estimate of the observed climate system (Pohlmann et al. 2009). The model error causes predictions to drift towards the model climate biases, which develop on different timescales (Doblas-Reyes et al. 2010) and, hence, depend on the forecast time. Various techniques for a-posteriori bias correction that take into account the forecast time, start-date or initial condition dependence of the bias have been designed and implemented (Kharin et al. 2012; Fuĉkar et al. 2014). This strategy is less appealing for multi-year predictions, because the magnitude of the predictable signal is smaller, and therefore could likely be removed during the inaccurate calculation of the bias.

An alternative to limit the drift is the flux correction (Magnusson et al. 2013a; Collins et al. 2006), that continuously corrects the model towards the observed attractor. In practice, it consists of additive adjustments that push the model solution towards the observed state. The third alternative is the anomaly initialisation method, which assimilates the observed anomaly variables onto an estimate of the model mean climate (Smith et al. 2013). Therefore, the philosophy of this strategy consists in allowing the model bias which is present in the model mean state, and only constraining the phase of the simulated variability towards the observed one. Previous studies (Hazeleger et al. 2013; Bellucci et al. 2015, Balmaseda et al. 2012) have applied these initialisation techniques to different models to highlight the relative strengths and limitations. The results show that in their standard implementation, at interannual time scales, the differences in skill of these techniques are small and limited to specific regions (Meehl et al. 2014). Magnusson et al. 2013b suggested that the best strategy could even be model dependent, because models have different biases.

*In this project we will test specific improvements to the standard anomaly initialisation techniques to assess whether tackling the detected limitations of the technique has a positive impact on the forecast skill.* The first limitation related to the use of the standard anomaly initialisation technique, is the risk of introducing anomalies recorded in the observed data whose amplitude does not fit in the range of the internal variability generated by the model. This could result in the model erroneously predicting extreme events, such as an intense El Niño or a pause in the thermohaline circulation. Volpi et al. 2017 addressed this issue by weighting the observed anomalies with the ratio between the model and observed standard deviation, to make their amplitude more consistent with the simulated variability. Such a refinement, together with the anomaly initialisation of the ocean density, showed improved skill in predicting the SST in some regions, the sea ice variability and the Atlantic Multidecadal Oscillation (AMO). The scaling technique that will be implemented in this project takes into account the non-normal distribution of the climate variables, which was not addressed in Volpi et al. 2017. The first year experiment will be devoted to the implementation and the assessment of the new scaling technique.

The second limitation which will be addressed in this project, is the geographical mismatch between the model and observed variability. There are two types of model biases that can affect the forecast skill: one is the bias affecting the amplitude of the signal. The second is the shift between the geographical position where the model develops the variability modes, and the actual observed position of the modes. To avoid the drift, the anomaly initialisation technique includes the amplitude bias in the initial state, by adding the model climatology to the observed anomalies. However, due to the displacement bias, the model might receive the information of the observed variability in a shifted position with respect to the location where the model would reproduce it. In those cases, the model might not interpret the strengthening of an observed variability mode as the strengthening of the corresponding variability mode in the model. As such, the information coming from the observed state will not be correctly propagated by the model, but either lost or, even worse, creating spurious perturbations. The second year experiment will implement and assess an innovative method to overcome such limitation.

This special project would run in parallel to the H2020 MSC project LISTEN, in which other initialisation techniques will be tested at both seasonal and multi-year scales, offering a rich pool of techniques for the comparison, on top of the existing standard initialised experiments.

***Experimental set-up***

The experiments will be carried out with the EC-Earth version 3.2 global coupled model. The atmosphere component will be initialised with the reanalyses ERA5 (Hersbach and Dee 2016) or ERAInterim (Dee et al. 2011) (depending on availability), the ocean component with NEMOVAR-ORAS4 reanalysis (Balmaseda et al. 2012), and the sea ice component with an updated version of the sea ice reconstruction from Guemas et al. 2013. However, initialising the ocean and sea ice components with ORA-20C (De Boisséson et al. 2018) will be taken into account, provided that the most recent starting dates does not exceed 2010 for a matter of availability. The experiments will run for 5 years with 16 ensemble members and 25 starting dates.

*The weight anomaly method: Quantile matching*

The first experiment will test a method to weight the observed anomaly and tackle the issue of introducing values that do not fit in the range of the internal variability generated by the model. This advanced method also takes into account the non-gaussianity of the climate variables. Figure 1 illustrates the implementation of the quantile matching method for the SST. In red, the cumulative distribution function (defined as the probability of a variable to take a value smaller than a specific given value) of SST is represented for a specific grid point, calculated with all NEMOVAR-ORAS4 members, over a reference period (in this example 1971-2000). Similarly, the SST cumulative distribution function for the historical simulation of the model is shown in black, for the same grid point. Assuming November 1960 as the target initial time, the red star indicates the value taken by NEMOVAR-ORAS4 at the initial date. The model is initialised with the model value (marked by a black star in the figure) whose cumulative distribution function matches with the one of the observed value at the initialisation time.

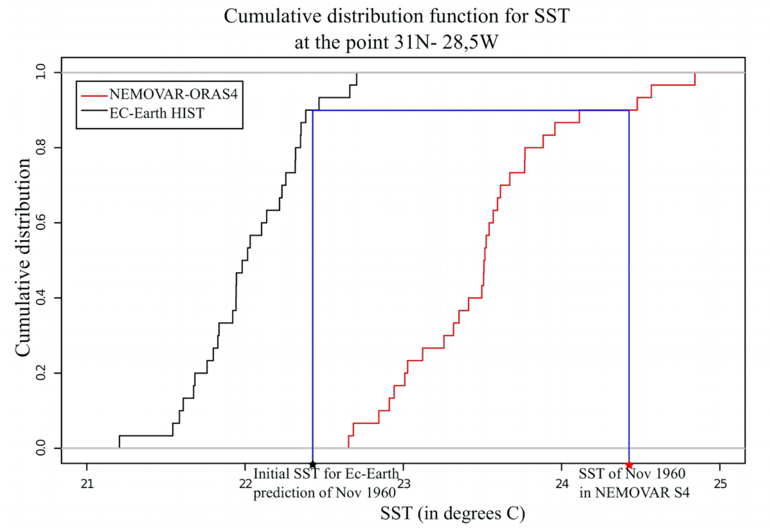


**Fig. 1** Example of the implementation of the quantile matching: cumulative distribution function for the SST at the point 31°N-28.5°W in red for NEMOVAR-ORAS4 and in black for a historical simulation of EC-Earth. The red star indicates the SST value of NEMOVAR-ORAS4 at such grid point on the 1st of November 1960. The prediction will be initialised with the value indicated by the black star: it corresponds to the value taken by the EC-Earth SST which has the same cumulative distribution as NEMOVAR-ORAS4 at the initialisation time.

*Addressing the spatial bias: the analog method*

The use of the anomaly initialisation technique in presence of the displacement bias implies the risk of introducing the information about the observed variability in a shifted position with respect to the location where the model would reproduce such variability. The analog method represents a possible alternative to overcome the issue. It consists in choosing an initial state that belongs to the model attractor and whose amplitude of the main variability modes is as close as possible to the amplitude of the corresponding observed variability modes, at the given initialisation time.

The implementation consists first of all in selecting a set of variability modes which are assumed to hold the most relevant climate variability, to represent the observed state. This implies the analysis of indices such as ENSO, the North Atlantic Oscillation (NAO), the Pacific Decadal Oscillation (PDO), AMO, the monthly mean intensity of the thermohaline circulation, the intensity of the subpolar and subtropical gyres, the strength of the Antarctic Circumpolar Current, etc. The chosen observed modes are computed using all the monthly mean data from the initialisation date and all the observed members over a reference period. Similarly, for the model the modes are estimated using all the members and the data of the initialisation month available from the historical simulations, in order to have the largest set of possible initial conditions. The initial state of a prediction is then chosen among the pool of the model states, as the one which has the set of model indices the closest to the observed indices at the initialisation time.

***Technical Characteristics***

The total number of coupled integrations at T255L91-ORCA1 will be thus 25 (starting dates) x5 (year-long) x16 (ensemble members) x2(experiments) = 4000 years. These can be split into 2000 years per each project year.

Scaling tests performed at ECMWF (in the framework of the SPNLTUNE special project, run by J. von Hardenberg) have determined that in the current configuration EC-Earth are optimal: T255L91-ORCA1: 144 cores for NEMO + 432 for IFS. One year of simulation is completed in about 18,000 SBUs/year. The total requirement will be **72 million SBUs** over two years.

The data storage volume required per year is 50000 gigabytes.

| Year | Experiment | Model Configuration | Experiment set-up | SBUs |
|---|---|---|---|---|
| Year 1 | Quantile mapping | 255L91-ORCA1 | 5 year-long, 16members, 25 starting dates | 36 million |
| Year 2 | Analog method | 255L91-ORCA1 | 5 year-long, 16members, 25 starting dates | 36 million |

*References*


- Balmaseda, M. A., K. S. Mogensen, and A. T. Weaver, 2012: Evaluation of the ECMWF ocean reanalysis ORAS4. Quart. J. Roy. Meteor. So., doi:10.1002/qj.2063.
- Bellucci, A., R. Haarsma, S. Gualdi, P. J. Athanasiadis, M. Caian, C. Cassou, E. Fernandez, A. Germe, J. Jungclaus, J. Kröger, D. Matei, W. Müller, H. Pohlmann, D. Salas y Melia, E. Sanchez, D. Smith, L. Terray, K. Wyser, and S. Yang, 2015: An assessment of a multi-model ensemble of decadal climate predictions. Clim Dyn, 44, 2787-2806. https://doi.org/10.1007/s00382-014-2164-y.
- Collins, M., B. B. B. Booth, G. R. Harris, J. M. Murphy, D. M. H. Sexton, and M. Webb, 2006: Towards quantifying uncertainty in transient climate change. Clim. Dyn., 365, 1957-1970.
- De Boisséson E., M. A. Balmaseda, and M. Mayer 2018. Ocean heat content variability in an ensemble of twentieth-century ocean reanalyses. Clim. Dyn. 50, 3783-3798. https://doi.org/10.1007/s00382-017-3845-0
- Dee D. P., S. Uppala, A. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Alonso-Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. Beljaars, L van de Berg, J-R Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A.J. Geer, L. Haimberger, S. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. W. Kållberg, M. Köhler, M. Matricardi, A. McNally, B. M. Monge-Sanz, J-J. Morcrette, B-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J-N. Thépaut, and F. Vitart, 2011: The ERA-INTERIM reanalysis: configuration and performance of the data assimilation system. Quart. J. Roy. Meteor. So ., 137, 553-597.
- Doblas-Reyes, F. J., A. Weisheimer, T. N. Palmer, J. M. Murphy, and D. Smith, 2010: Forecast quality assessment of the ensembles seasonal-to-decadal stream 2 hindcasts. ECMWF Technical Memorandum, 621.
- Fuĉkar N.S., D. Volpi, V. Guemas and F. J. Doblas-Reyes (2014). A posteriori adjustment of near-term climate predictions: Accounting for the drift dependence on the initial conditions. Geophysical Research Letters, 41, doi:10.1002/2014GL060815.
- Guemas, V., F. Doblas-Reyes, K. Mogensen, S. Keely, and Y. Tang, 2013: Ensemble of sea ice initial conditions for interannual climate predictions. Clim. Dyn., 43, 2813-2829, doi: 10.1007/s00382-014-2095-7.
- Hazeleger W., V. Guemas, B. Wouters, S. Corti, I. Andreu-Burillo, F. J. Doblas-Reyes, K. Wyser, and M. Caian, 2013: Multiyear climate predictions using two initialization strategies. Geophys. Res. Lett., 40 (9), 1794-1798, doi:10.1002/grl.50355.
- Hersbach H. and D. P. Dee 2016, ECMWF Newsletter No. 147.
- Kharin V., G. Boer, W. Merryfield, J. Scinocca, and W. Lee, 2012: Statistical adjustment of decadal predictions in a changing climate. Geophys. Res. Lett., 39 (L19705), doi:10.1029/ 2012GL052647.
- Magnusson L., M. Alonso-Balmaseda, S. Corti, F. Molteni and T. Stockdale, 2013a: Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors. Clim Dyn, 41: 2393. https://doi.org/10.1007/s00382-012-1599-2.
- Magnusson L., M. Alonso-Balmaseda, and F. Molteni, 2013b: On the dependence of ENSO simulation on the coupled model mean state, Clim. Dyn. 41: 1509. https://doi.org/10.1007/s00382-012-1574-y
- Meehl G.A., L. Goddard, G. Boer, R. Burgman, G. Branstator, C. Cassou, S. Corti, G. Danabasoglu, F. Doblas-Reyes, E. Hawkins, A. Karspeck, M. Kimoto, A. Kumar, D. Matei, J. Mignot, R. Msadek, A. Navarra, H. Pohlmann, M. Rienecker, T. Rosati, E. Schneider, D. Smith, R. Sutton, H. Teng, G.J. van Oldenborgh, G. Vecchi, and S. Yeager, 2014: Decadal Climate Prediction: an update from the trenches. Bull. Amer. Meteor. Soc., 95, 243–267, https://doi.org/10.1175/BAMS-D-12-00241.1
- Pohlmann H., J. Jungclaus, A. Köhl, D. Stammer, and J. Marotzke, 2009: Initializing decadal climate predictions with the GECCO oceanic synthesis: Effects on the North Atlantic. J. Clim., 22, 3926-3938.

- Smith D. M., R. Eade, and H. Pohlmann, 2013: A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction. Clim. Dyn., doi:10.1007/ s00382-013-1683-2.
- Volpi D., V. Guemas, F. J. Doblas-Reyes, E. Hawkins, and N. Nichols (2017). Decadal climate prediction with a refined anomaly initialisation approach, Climate Dynamics, 48, Issue 5-6, 1841-1853 doi:10.1007/s00382-016-3176-6.