REQUEST FOR A SPECIAL PROJECT 2024–2026

MEMBER STATE:	……………Italy…………………………………………………………………………….…..
Principal Investigator[1]**:**	………Valerio Lembo…………………………….…………………………………..……
Affiliation:	………CNR-ISAC………………………………………………………………………….……
Address:	…via Gobetti 101, 40125, Bologna, Italy..…………………………………….……
Other researchers:	Federico Fabiano (CNR-ISAC), Susanna Corti (CNR-ISAC), Paolo Davini (CNR-ISAC), Jost von Hardenberg (Politecnico di Torino)
Project Title:	Linear response theory and stochastic parametrizations for the investigation of response and sensitivity: a case study with an intermediate-complexity model

To make changes to an existing project please submit an amended version of the original form.)

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SP** ………….	
Starting year: (A project can have a duration of up to 3 years, agreed at the beginning of the project.)	2024	
Would you accept support for 1 year only, if necessary?	YES ☒	NO ☐

Computer resources required for project year:		**2024**	**2025**	**2026**
High Performance Computing Facility	[SBU]	*2869200*	*7012500*	*1980500*
Accumulated data storage (total archive volume)[2]	[GB]	91500	157500	186200

EWC resources required for project year:		**2024**	**2025**	**2026**
Number of vCPUs	[#]			
Total memory	[GB]			
Storage	[GB]			
Number of vGPUs[3]	[#]			

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] These figures refer to data archived in ECFS and MARS. 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.
[3]The number of vGPU is referred to the equivalent number of virtualized vGPUs with 8GB memory.

Principal Investigator: ……Valerio Lembo…………………………………………

Project Title: … Linear response theory and stochastic parametrizations for the investigation of response and sensitivity: a case study with an intermediate-complexity model……………………………………

Extended abstract

Motivation

Along with the increasing complexity of climate models, the necessity to advance our understanding of fundamental processes governing the climate system has motivated the development of model hierarchies (Held et al. 2005, Ghil and Lucarini 2020). The increase in computational capabilities has led scientists to cope with challenging technological issues, but this can by no means come at the expense of adopting simplified and conceptual models that address the grounding features of the dynamics of atmosphere, oceans and other components of the climate system. More complex models bear information about the behaviour of simplified models, and the other way round, so that a hierarchical model chain is needed. This reflects the multiscale nature of the system, and the need for parameterizations addressing for the dynamics that are not explicitly resolved by the model discretisation. Furthermore, a discrepancy between more complex models and simplified model statistics does not necessarily imply that higher order models perform better in reproducing the system's statistics than the lower order ones (e.g. Ghil 2015).

In order to fill the gap in the hierarchy ladder between conceptual models and fully developed General Circulation Models (GCM), Earth system Models of Intermediate Complexity (EMIC) have been developed (Claussen et al. 2002). These are three-dimensional (or 2.5-D, when the vertical structure is implied diagnostically from the horizontal fields) models, embedding all the relevant processes resolved by the discretized primitive equations, at a resolution coarser than operational GCMs and with some modules deliberately simplified. Several strategies have been developed, depending on the problem, from paleoclimate (cfr. POEM v1.0, with parametrised synoptic scales, Totz et al. 2018) to teleconnection studies (e.g. SPEEDY, an atmospheric GCM with coarse resolution and simplified parameterizations, Kucharski et al. 2006, often coupled with CLIO ocean and sea-ice model, Severijns and Hazeleger, 2010; Goosse and Fichefet, 1999).

Unlike other EMICs, the Planet Simulator (PlaSim; Fraedrich and Lunkeit, 2008; Fraedrich, 2012) is a 3-D primitive equations atmospheric GCM, featuring a number of spectral resolutions (from T21 to T63) and a slab-ocean at its boundary. This poses it on the one hand at the high-end of the EMIC complexity, on the other hand exposing it to the inadequacy of the slab-ocean boundary conditions, when sensitivity studies imply timescales of the order of sea-ice and oceanic internal variability. The problem has been recently addressed by several working groups, using the ocean and sea-ice modules of the UVic Earth System model (Weaver et al. 2001, Schmittner et al. 2010), or the ones from the GENIE v1.0 (Lenton et al. 2006; Holden et al. 2016).

A simple coupled version of PlaSim with a dynamic ocean has been tested at CNR-ISAC, tuning the coupling of PlaSim with the Large-Scale Geostrophic (LSG) ocean module. Despite the unavoidable limitation of working with a geostrophic ocean, this offers the unique opportunity to run a fully-tested stable version of a highly performing state-of-the-art EMIC for sensitivity and paleo-climatic studies.

Here, the sensitivity of internal variability to external forcings, as it can be distinguished from the long-term climate response in PlaSim-LSG, is studied. In order to make it an effective playground for testing model settings with more complex Earth system models, we take advantage of a recent PlaSim implementation (available on GitHub: https://github.com/ValerioLembo/PLASIM/tree/sppt_plasim), including the routines for Stochastically Perturbed Parametrization Tendencies (SPPT) scheme (Buizza et al. 1999) developed at ECMWF. The impacts of such new schemes will be evaluated under different horizontal/vertical resolutions, as detailed below.

The computational advantage provided by such EMIC-class models can be exploited to investigate the asymptotic response of the modelled climate in a systematic way. In order to do so, a particularly sound and rigorous methodology is the one relying on the linear response theory (LRT, Ruelle 2009), formalizing the analysis of the response of a non-autonomous dynamical system with chaotic properties to a given forcing. LRT has found several applications in the field of climate science (e.g. Zappa et al. 2020, Torres Mendonça et al. 2021, Basinski-Ferris and Zanna, 2023). An experimental protocol adopting the LRT in the context of climate modelling was introduced in Ragone et al. 2016, then applied to a state-of-the-art CMIP-class coupled model in Lembo et al. 2020. This relies on the fact that the first order

Green's function of any observable being part of the system can be straightforwardly retrieved with a suitable choice of the perturbing forcing, for instance a step forcing with abrupt change in a specified forcing agent at a given time during the simulation. If the LRT holds, in order to predict the asymptotic response of a chosen observable belonging to the system to a "sufficiently smooth" forcing (i.e. for practical purposes a relatively small perturbation of the background state), the first order Green's function can be convolved with the time modulation of the imposed forcing. As described by Ragone et al. 2016, several crucial aspects of the climate forced response can be effectively interpreted through the lens of LRT, such as the equilibrium climate sensitivity, the transient climate response etc., and a number of interesting features can be successfully predicted, such as the reduction of the Atlantic Meridional Overturning Circulation (AMOC) or the emergence of the North Atlantic cold bloc (cfr. Lembo et al. 2020).

The PlaSim-LSG model

PlaSim-LSG has been tested with four different horizontal resolutions in the atmosphere, while the oceanic component, LSG, was kept at constant 5x5 degrees resolution. The resolutions correspond to spectral truncations T21 (32x32 steps in the lonxlat grid space), T31 (48x48), T42 (60x60), T63 (96x96), respectively. The model timestep is rescaled accordingly, in order to ensure numerical stability. Results on the computational resources requested from the optimal combination of these parameters are summarized in Table 1, and were obtained from tests performed on several supercomputer platforms, specifically the TINTIN and WILMA machines at CNR-ISAC, and HPC systems Galileo100 and Marconi100 at CINECA. Concerning vertical resolution, 15 atmospheric vertical levels is the smallest numbers of levels for a fixed horizontal resolution that ensures a reasonable closure of the energy budget at the Top-of-Atmosphere. This results from tests performed on the Deutsches KlimaRechenZentrum (DKRZ) HPC systems in Hamburg.

PlaSim-LSG is set to run on multi-processors, if compiled accordingly. Scaling tests have been performed with two of the standard resolutions, namely T21 and T42, and results are shown in Figure 2. We emphasize that the scaling of the model has been found to be suboptimal, as increasing the number of nodes the CPU time only decreases by up to 1.5x/2x for T42/T21 when 8 nodes are considered.

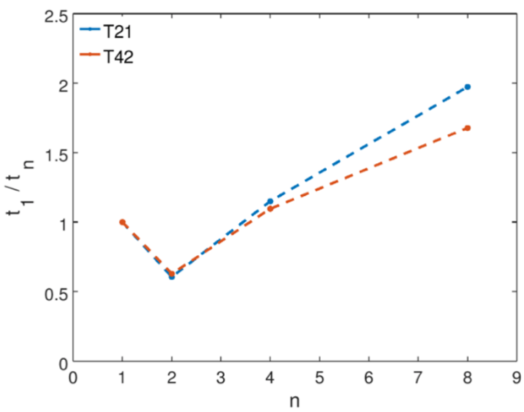

***Figure 1:** scaling of CPU time as a function of the number of cores used for the integration. The relative gain in computational time is computed as the ratio of the time required for sequential integration to the time actually employed. Examples are provided for two spectral resolutions, namely T21 (dashed blue), and T42 (dashed red).*

Packages necessary to build the model and run it are typically part of the core build of a Linux machine. This includes C and C++ compilers (e.g. gcc, g++), Fortran compilers (such as gfortran), the "make" utility, the X11 libraries. For multi-processor program execution, mpi, mpif90, mpicc, mpirun are also needed. For post-processing, the Climate Data Operators (CDO) package is required, as it is called by the model workflow at the end of every simulated year. NCO and NCVIEW libraries are also helpful in order to monitor the progress of the simulations.

***Table 1:** CPU time per year (in minutes, rounded to the nearest integer) and storage per year as a function of spectral truncation, vertical and temporal resolution. The number of longitudinal and latitudinal gridsteps for each spectral resolution is also provided.*

Spect. Trunc.	***Atm. Vert.***	***Timestep (min.)***	***Lon. steps***	***Lat. steps***	***CPU time x year***	***Storage x year***

	lev.				*(min)*	*(MB)*
T21	*15*	*45*	*32*	*32*	*2*	*24*
T31	*15*	*30*	*48*	*48*	*6*	*51*
T42	*15*	*20*	*60*	*60*	*19*	*598*
T63	*15*	*10*	*96*	*96*	*89*	*844*

Experimental design

The experimental design firstly consists of individual unforced runs (namely "control") aimed at generating a number of initial conditions for the ensemble runs that are performed across the project. These unforced runs share the same radiative boundary conditions characteristic of the pre-industrial era, as in CMIP simulations. In order to ensure that the initial conditions are independent and identically distributed, we sample them over a sufficient temporal distance, in order to avoid the autocorrelation determined by the timescales of oceanic variability (cfr. Ganopolski et al. 2002, Wunsch and Heimbach 2008). So far, a 4000-yr long control run has already been performed at T21 resolution, 2000-yr with T42, with the aim to prolong it to 4000 years. The numerical stability of T31 and T63 is currently being investigated in order to provide sufficiently long control runs for the remaining two resolutions.

The long-term climate response to increased CO2 concentrations is investigated by performing ensemble step forcing runs. These experiments consist of imposing an instantaneous change in CO2 concentrations, then running the model until it reaches statistically steady state. The initial conditions are taken from the control run, and a number of steps is considered, namely 2x, 4x and 8x the pre-industrial CO2 concentrations. We demonstrated the experimental design with the T21 resolution performing 20 ensemble runs over a 2000-yr integration period. The initial conditions were taken from the respective control simulation. Results are shown in Figure 2, the thick line denoting the ensemble mean evolution.

Given the ensemble of millennial runs obtained with the step forcing setting illustrated above (as shown in Figure 2 for T21), hereafter "predictor", the Green's function computed from that is convolved with the time modulation of a ramp-up forcing (hereafter "predictand"), in which the CO2 concentration is linearly increased by 1% every year and then kept steady for at least 1000 years. The predicted response is compared with an actual ramp-up 20-members ensemble model run for each magnitude of the step. The adherence of the prediction to the ensemble mean evolution is evaluated as a function of the ensemble members and resolution. We do indeed expect that:

- *The increasing magnitude of the step would likely degrade the predictive skill of the LRT protocol, given the emergence of non-linear behaviour;*

- *The increasing resolution would introduce non-linear feedbacks, making the predictive power of the Green's function from the predictor weaker and more prone to the interaction of different spatio-temporal timescales of internal variability;*

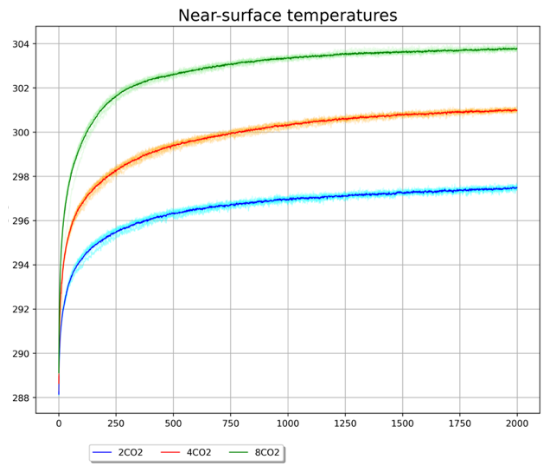

Figure 2: *Time series of near-surface air temperatures for T21 2000-yr 20-members ensemble simulations (xCO2_ens_t21) for step forcings at time t=0 2xCO2 (in blue), 4xCO2 (in red), 8xCO2 (in green). Thick lines denote ensemble mean, thin lines the evolutions of the individual runs.*

The other aspect that we chiefly aim at investigating in the project, is the impact of the mentioned subgrid-scale stochastic parameterization scheme, namely a Stochastically Perturbed Parametrization Tendencies (SPPT) scheme (Buizza et al. 1999). This was originally implemented as part of the operational Integrated Forecast Scheme (IFS) model at the European Centre for Medium-range Weather Forecast (ECMWF) (Palmer et al. 2009). In other models, the inclusion of SPPT in model equations has proven to be effective in better constraining the statistics of internal variability for a state-of-the-art GCM (e.g. Yang et al. 2019). We evaluate the outcome of this development in a version of PlaSim-LSG adapted by Frank Lunkeit (University of Hamburg) and publicly available on GitHub (see above). Since the tuning of a model with several different choices of the SPPT parameters is often computationally demanding, we consider that working with PlaSim-LSG is particularly advantageous, given that it outperforms state-of-the-art coupled GCMs by a few orders of magnitude in terms of necessary computational resources. Therefore, we investigate the configuration space in order to test the conservation properties against different parameters. Once the tuning is done on small integration periods (few months to 10 years), we perform for each spatial resolution a 50-members ensemble of model simulations initialised from the control run and with boundary conditions representative of the historical period, in agreement with CMIP forcing conditions for the 1850-2020 period. For each ensemble, the "deterministic" counterpart (with SPPT scheme switched to off) is developed, in order to assess the impact of including such parametrisation. Several aspects of the model climatology are diagnosed, with the help of existing evaluation tools developed in our group, namely the modes of atmospheric variability (i.e. weather regimes, cfr. Fabiano et al. 2021), the energy budgets, transports, material entropy production and water mass budgets (after Lembo et al. 2019; Eyring et al. 2020).

Summarizing, we plan to perform the following experiments:

1. CTRL_tx: the remaining single-run control simulations, that have not been previously provided (T31, T63). In order to ensure sufficient decorrelation of initial conditions, each run will consist of at least 5000 years;

2. HIST_ENS_tx_d: an ensemble of 50 realisations for each of the four resolutions, with the initial conditions provided by the CTRL runs, forced with radiative conditions representative of the 1850-2020 period (i.e. CMIP-historical forcing for 1850-2014, RCP8.5 for 2014-2020). The model setting will be the usual PlaSim-LSG coupling with deterministic tendency equations;

3. HIST_ENS_tx_s: same as in HIST_ENS_d, the only difference being the inclusion of the SPPT scheme in the tendency equations of PlaSim;

4. 1pct_xCO2_t21: a 2000-yr 20-members ensemble experiment with annual increase of CO2 concentration for each of the step forcing experiments performed so far (i.e. 2xCO2, 4xCO2, 8xCO2; cfr. Figure 2).

5. 2CO2_ENS_tx: a couple of ensembles with a fixed change in CO2 concentrations and corresponding ramp increase (as in 4.) for each of the three remaining resolutions (the runs with T21 are either developed in 4. or already available), in order to detect the emergence of non-linear feedbacks and the reliability of the linear prediction as a function of spatial resolution;

Configuration and justification of resources

The request number of SBU for the present project is organized according to the proposed simulations as follows:

Experiment Name	***Description***	***SBU /year***	***Total years***	***Total SBU***	***Year 1***	***Year 2***	***Year 3***
CTRL_t31	*Control run at T31*	*6*	*5000*	*30000*	*30000*	*0*	*0*
CTRL_t63	*Control run at T63*	*89*	*5000*	*443200*	*443200*	*0*	*0*
1pct_xCO2_t21	*Ensemble runs at T21 with 1% increase until 2x, 4x, 8x CO2*	*2*	*120000*	*240000*	*240000*	*0*	*0*
2CO2_ens_t31	*Ensemble runs at T31 with 2xCO2 step forcing and 1% increase until 2xCO2*	*6*	*80000*	*480000*	*240000*	*240000*	*0*
2CO2_ens_t42	*Ensemble runs at T42 with 2xCO2 step forcing and 1% increase until 2xCO2*	*19*	*80000*	*1520000*	*1520000*	*0*	*0*
2CO2_ens_t63	*Ensemble runs at T63 with 2xCO2 step forcing and 1% increase until 2xCO2*	*89*	*80000*	*7120000*	*356000*	*6764000*	*0*
HIST_ens_t21_d	*Ensemble runs of deterministic historical simulation at T21*	*2*	*8500*	*17000*	*0*	*0*	*17000*
HIST_ens_t21_s	*Ensemble runs of stochastic historical simulation at T21*	*2*	*8500*	*17000*	*0*	*0*	*17000*
HIST_ens_t31_d	*Ensemble runs of deterministic historical simulation at T31*	*6*	*8500*	*51000*	*0*	*0*	*51000*
HIST_ens_t31_s	*Ensemble runs of stochastic historical simulation at T31*	*6*	*8500*	*51000*	*0*	*0*	*51000*
HIST_ens_t42_d	*Ensemble runs of deterministic historical simulation at T42*	*19*	*8500*	*161500*	*0*	*0*	*161500*
HIST_ens_t42_s	*Ensemble runs of stochastic historical simulation at T42*	*19*	*8500*	*161500*	*0*	*0*	*161500*
HIST_ens_t63_d	*Ensemble runs of deterministic historical simulation at T63*	*89*	*8500*	*756500*	*0*	*0*	*756500*
HIST_ens_t63_s	*Ensemble runs of stochastic historical simulation at T63*	*89*	*8500*	*756500*	*0*	*0*	*756500*
5% Buffer				*59000*	*40000*	*8500*	*8500*
Total				*12395500*	*2869200*	*7012500*	*1980500*

In terms of storage space, we start from an already sizeable amount of data that have been produced from T21 and T42 control simulations, as well as from the ensembles of step forcing experiments with T21. This sums up in total to about 25 TB that will have to be made available for computations on the machine. Concerning the new simulations, we estimate that an output at monthly resolution for 1 year of simulations amounts to 24 MB, 51 MB, 598 MB, 844 MB for T21, T31, T42, T63, respectively (cfr. Table 1).

Multiplying for the number of proposed integrated year, as a function of resolution, we obtain:

- *3.2 TB for T21;*
- *5.2 TB for T31;*
- *58 TB for T42;*
- *86 TB for T63;*

These will be distributed across the duration of the project according to the Table at the beginning of the current proposal.

References

- *Basinski-Ferris, A. and Zanna, L.: Estimating freshwater flux amplification with ocean tracers via linear response theory, Earth Syst. Dynam. Discuss. [preprint], in review, 2023.*
- *Buizza, R., Miller, M., and Palmer, T. N.: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System, Quart. J. Roy. Meteorol. Soc., 125, 2887–2908, 1999*
- *Claussen, M., Mysak, L., Weaver, A. et al.: Earth system models of intermediate complexity: closing the gap in the spectrum of climate system models, Clim. Dyn., 18, 579–586, 2002*
- *Eyring, V. et al.: ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP. Geosci. Mod. Devel., 9(5), 1747–1802, 2016*
- *Fabiano, F., Meccia, V. L., Davini, P., Ghinassi, P., and Corti, S.: A regime view of future atmospheric circulation changes in northern mid-latitudes, Weather Clim. Dynam., 2, 163–180, 2021*
- *Fraedrich, K.: A suite of user-friendly climate models: Hysteresis experiments, Eur. Phys. J. Plus, 127, 53, 2012*
- *Fraedrich, K., and Lunkeit, F.: Diagnosing the entropy budget of a climate model, Tellus A: Dynamic Meteorology and Oceanography, 60:5, 921-931, 2008*
- *Ganopolski, A., and Rahmstorf, S.: Abrupt Glacial Climate Changes due to Stochastic Resonance. Physical Review Letters, 88(3), 2002*
- *Ghil, M.: A mathematical theory of climate sensitivity or, How to deal with both anthropogenic forcing and natural variability?, in Climate Change: Multidecadal and Beyond, edited by C. P. Chang, M. Ghil, M. Latif, and J. M. Wallace (World Scientific, Singapore), pp. 31–51, 2015*
- *Ghil, M., and Lucarini, V.: The physics of climate variability and climate change. Reviews of Modern Physics, 92(3), 35002, (2020)*
- *Goosse, H., and Fichefet, T.: Importance of ice-ocean interactions for the global ocean circulation: A model study, J. Geophys. Res., 104(C10), 23337– 23355, 1999*
- *Held, I. M.: The Gap between Simulation and Understanding in Climate Modeling. Bulletin of the American Meteorological Society, 86(11), 1609–1614, 2005*
- *Holden, P. B., Edwards, N. R., Fraedrich, K., Kirk, E., Lunkeit, F., and Zhu, X.: PLASIM–GENIE v1.0: a new intermediate complexity AOGCM, Geosci. Model Dev., 9, 3347–3361, https://doi.org/10.5194/gmd-9-3347-2016, 2016*
- *Kucharski, F., Molteni, F. and Bracco, A.: Decadal interactions between the western tropical Pacific and the North Atlantic Oscillation. Clim Dyn, 26, 79–91, 2006*
- *Lembo, V., Lunkeit, F. and Lucarini, V.: TheDiaTo (v1.0) – a new diagnostic tool for water, energy and entropy budgets in climate models. Geosci. Mod. Devel., 12(8), 3805–3834, 2019*

- *Lembo, V., Lucarini, V. and Ragone, F.: Beyond Forcing Scenarios: Predicting Climate Change through Response Operators in a Coupled General Circulation Model, Sci Rep., 10, 8668, 2020*
- *Lenton, T.M., Williamson, M.S., Edwards, N.R. et al.: Millennial timescale carbon cycle and climate change in an efficient Earth system model. Clim Dyn, 26, 687, 2006*
- *Palmer, T., Buizza, R., Doblas-Reyes, F., et al.: Stochastic Parametrization and Model Uncertainty, ECMWF, 598, 2009*
- *Ragone, F., Lucarini, V., and Lunkeit, F.: A new framework for climate sensitivity and prediction: a modelling perspective. Clim Dyn, 46(5–6), 1459–1471, 2016*
- *Ruelle, D.: A review of linear response theory for general differentiable dynamical systems. Nonlinearity, 22(4), 855–870, 2009*
- *Schmittner, A., Silva, T. A. M., Fraedrich, K., Kirk, E., and Lunkeit, F.: Effects of mountains and ice sheets on global ocean circulation, J. Climate, 24, 2814–2829, 2010*
- *Severijns, C. A. and Hazeleger, W.: The efficient global primitive equation climate model SPEEDO V2.0, Geosci. Model Dev., 3, 105–122, 2010*
- *Torres Mendonça, G. L., Pongratz, J., and Reick, C. H.: Identification of linear response functions from arbitrary perturbation experiments in the presence of noise – Part 1: Method development and toy model demonstration, Nonlin. Processes Geophys., 28, 501–532, 2021*
- *Totz, S., Eliseev, A. V., Petri, S., Flechsig, M., Caesar, L., Petoukhov, V., and Coumou, D.: The dynamical core of the Aeolus 1.0 statistical–dynamical atmosphere model: validation and parameter optimization, Geosci. Model Dev., 11, 665–679, 2018*
- *Weaver, A.J., Eby, M., Wiebe, E.C., et al.: The UVic earth system climate model: Model description, climatology, and applications to past, present and future climates, Atmosphere-Ocean, 39:4, 361-428, 2001*
- *Wunsch, C. and Heimbach, P.: How long to oceanic tracer and proxy equilibrium? Quaternary Sci. Rev., 27(7–8), 637–651, 2008*
- *Yang, C., Christensen, H. M. , Corti, S., von Hardenberg, J., and Davini,P.: The impact of stochastic physics on the El Niño Southern Oscillation in the EC-Earth coupled model, Clim. Dyn., 1-17, 2019*
- *Zappa, G., Ceppi, P. and Shepherd, T. G.: Time-evolving sea-surface warming patterns modulate the climate change response of subtropical precipitation over land. Proc. Nat. Acad. Sci., 117(9), 4539–4545, 2020*