CLIVAR Research Foci Proposal Template

Title: Confronting Earth System Model Trends and Multidecadal Variability with Observations

Acronym: CEMT-MV

1. Scientific Background and Rationale (max. 1500 words)

<u>Scientific Background</u>

For decades, climate scientists have used numerical models to project future climate. The reliability of these projections, relies on the accuracy of models in representing both the transient evolution of anthropogenically forced signals and the dynamics of natural climate variability. Unlike weather forecasting, with many opportunities to evaluate forecast skill, there is only one realisation of anthropogenic climate change for evaluating projection skill. Thus, climate models have been primarily validated in terms of process representation and statistical properties. This situation is now changing as signals, long-term trends, are emerging across many variables and regions and their connection to anthropogenically forced climate change versus internal variability are being assessed. This now places the modelling community in a unique position to confront Earth System Model Trends with observations.

This effort must however go hand in hand with reducing uncertainties in simulating multi-decadal variability (MDV), which confound the attribution of climate trends, particularly at a regional level. The key issue is the uncertainty in the contributions of external forcing and internal climate dynamics in driving MDV. For example, the North Atlantic cooling from the 1940's-1970's and subsequent warming, which has been completely explained by either external forcing and ocean circulation changes (Hu et al. 2024). This uncertainty not only complicates signal attribution, but it affects climate sensitivity estimates, and degrades both seasonal and decadal predictions (Beverly et al. 2024, Chikamoto et al. 2019). Fortunately, there are opportunities to reduce uncertainties in the natural characteristics of MDV, through the advent of paleo reanalysis and climate reconstructions, developing a more integrated view of MDV, and targeted climate prediction experiments.

Now more than ever it is important to robustly identify forced signals and internally generated MDV. While controversial, unforced MDV may have enhanced the recent northern hemisphere warming and it may now be changing phase, affecting even iconic regions like the Arctic (Fig. 1). Such a change could substantially offset warming in the next decade (Omrani et al. 2022).

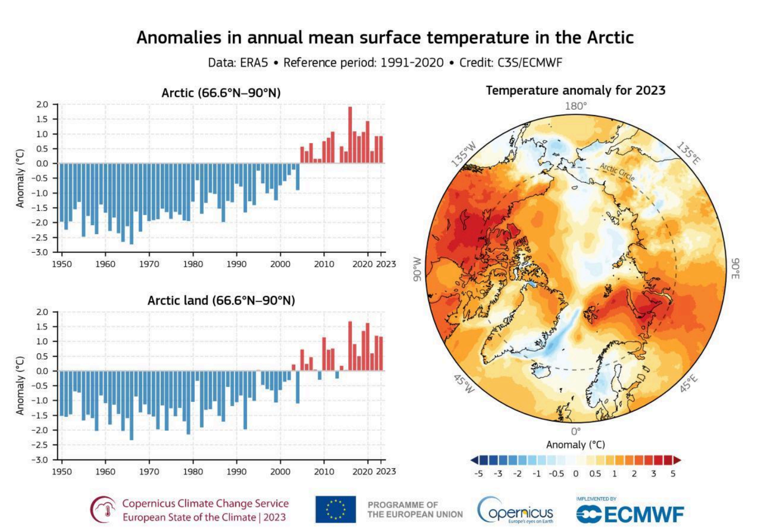

Figure 1. (Left) Annual average surface air temperature anomalies from 1950 to 2023 (top) for the Arctic as a whole (66.6°–90°N) and (bottom) over all Arctic land areas. (Right) Map of the annual average surface air temperature anomalies for 2023 over the north polar region. All anomalies are relative to the average for the 1991–2020 reference period. Data source: ERA5. Credit: C3S/ECMWF.

This RF aims to focus the international scientific community on these pressing issues so that we are better placed to inform policy makers and for maintaining clear and strong public discourse on the urgent need for climate mitigation and adaptation.

Previous activities

RF-CEMT-MV is motivated by several activities that have shown great community interest in the topic.

The online CLIVAR CDP 2022 annual workshop on external versus internal variability on decadal and longer time scales with close to 300 registrants (CLIVAR Report No. 02/2022); and the ICTP/CLIVAR 2023 summer school and workshop on Atlantic variability and tropical basin interactions. During these events, large uncertainties in the response of models to external forcing were identified as a key challenge that cannot be addressed with existing experimental protocols. Model errors and the short observational record were highlighted as major obstacles to reducing these uncertainties, while the development of constraints on regional climate change was seen as a priority. Recommendations included strengthening collaboration with paleo-climate researchers; the development of new approaches to constrain uncertainties; and novel process-based numerical experiments to identify dynamics and impacts of model biases.

As part of the WCRP Open Science Conference that took place in October 2023, co-chair Shaw led a concept paper on "Regional Climate Change: Consensus, Discrepancies and Ways Forward" (Shaw et al. 2024). In addition, the co-chairs Simpson and Shaw organized a US CLIVAR workshop, Confronting Earth System Model Trends with Observations, which brought together researchers with expertise on wide-ranging aspects of the Earth system to discuss the current state of the science in regards to historical trends in the Earth system and their representation in the ESMs that are used for future climate projections. This workshop led to a review paper published in science advances (Simpson, Shaw et al. 2025) which summarizes the state-of-the-science on this topic and identified research challenges and future directions in this area. Figure 2 below illustrates the three confronting cases found in the literature.

These efforts tackled similar problems but from different angles - the work on multi-decadal variability was primarily focused on decomposing historical trends into their forced and internal components, while confronting model trends focussed on the consistency of historical trends with the distributions produced by Earth System Models (ESMs). Ultimately, we must both ensure that ESMs can capture historical trends but can also do so for the right reasons with the correct proportion of forced versus internal variability.

Rationale

The success of the previous activities naturally leads to our proposal to develop a CLIVAR Research Foci (RF) that will

1. Devise appropriate metrics to compare models and observations — for trends and MDV
2. Further our understanding of internal and forced MDV (including multi-decadal trends) using paleoclimate reanalysis for the last millennium
3. Calibrate ESMs, predictions, and projections
4. Create updated outlooks of the state of climate signals and work towards the development of a public-facing web-portal.

The major outcomes of the RF are: (i) improved understanding of the relative roles of external-forcing and internal-dynamics in historical trends; (ii) developing a “state-of-the-signal” assessment which will summarize the current knowledge of historical forced trends and natural MDV and the abilities of climate models to reproduce them, which will inform the IPCC and the research community and will build on other activities (State of the Climate 2023, Forster et al. 2024); (iii) Community model analysis projects on the forthcoming CMIP7 assessment Fast Track; and (iv) development towards a web-based portal that will provide accessible information with expert guidance on the extent to which the current state-of-the-art ESMs are accurately representing historical forced trends and natural MDV. The RF would inform the next IPCC assessment report through peer reviewed publications and will also initiate resources that can provide up to date information between IPCC ARs.

2. Terms of Reference (max. 250 words)

Our main goal is to focus international research on identifying and predicting external-forced trends and internally-generated MDV, thereby providing a more informed assessment of projections for the coming decades. More specifically, RF-CEMT-MV will bring together scientists to focus on the following goals:

- Host workshops, webinars, and organize conference sessions to create awareness, synthesize research, and stimulate needed research
- Organize analysis, intercomparison, and coordinated numerical experiments to disentangle forced trends from MDV;
- Devise appropriate methods for assessing the performance of ESMs in predicting trends and MDV across a wide range of variables;
- Advance process level understanding in simulating historical trends and MDV, and foster new approaches to calibrate models predictions/projections;
- Inform on future observational efforts by identifying problem areas where improved process understanding is needed and demonstrating the need for consistent long-term climate records through identification of observational uncertainties;
- Write review/perspective papers to consolidate understanding and guide future research;
- Develop a “State-of-the-signal” summary assessment on the current knowledge of historical trends and the ability of climate models to reproduce them (build on the current “State of the Climate”); Work towards a web-based portal to present this information in an accessible, regularly updated, information with expert guidance on the extent to which ESM’s are accurately representing historical trends.
- hold summer schools and hackathons to educate and interest promising early career researchers from around the world;
- To report annually to CLIVAR.

The entire RF team will meet at least once every six months, with additional smaller meetings to arrange various RF activities.

3. Relevance to CLIVAR Science Plan (max. 100 words)

The RF will confront model trends and variability with observations, thereby tackling gaps in our understanding of natural and anthropogenic forcing of decadal to multi-decadal changes. We will advance process-level understanding of successes and discrepancies in simulating historical trends to inform improvement of S2D predictions and reduce uncertainties in long-term projections, globally and regionally. The RF will lead to "understanding mechanisms of climate variability and change, analysis and predictions of changes in Earth's climate". Finally, the RF will focus on the "development and evaluation of climate simulations and predictive capabilities" by examining discrepancies in simulating historical trends as well as successes.

4. Working Plan (max. 1000 words)

Component 1 Evaluating signal and MDV in ESMs

The recent review article of Simpson et al. 2025 assessed whether long term trends fell into the following categories: successes, partial successes, discrepancies, and uncertain situations and examples of three of these categories are shown in Figure 2. But the field is fast moving, so in this first stage of the RF we will continue to distill the literature on this topic and identify trends in ESMs that are not represented with fidelity. Through this we will also identify gaps - features of the climate system where the long term trends in ESMs have not yet been explored in detail, and start to explore them and will develop approaches for identifying and synthesizing the likelihood of discrepancies. This initial scoping analysis will be performed on existing large ensembles, such as those represented in the Multi-Model Large Ensemble Archive (https://www.cesm.ucar.edu/community-projects/mmlea) or the all forcing simulations provided through the EPESC Large Ensemble Single Forcing MIP (LESFMIP) activity (Smith et al 2022), as well as existing multi-model ensembles (CMIP3 to CMIP6) and the forthcoming CMIP7 simulations so that the evolution of model behavior can be explored. High resolution ensembles such as those from HighResMIP, the CESM1 MESACLIP, and the EERIE project will also be examined to explore resolution sensitivities and other experiments such as AMIP or experiments (such as those from the Tropical Basin Interaction coordinated experiments (TBI-coEX)) will be used to gain mechanistic understanding. These simulations will be compared to a wide array of observational products to incorporate observational uncertainty and will use a multi-variate approach to ensure robustness of conclusions across multiple connected observed variables.

To further refine the assessment of long-term trends, we will extend the analysis to systematically investigate the simulation of internally-generated vs externally-forced MDV. We will consider a range of methods, such as those proposed through the Forced Component Estimation Statistical Method Intercomparison Project (ForceSMIP) to estimate the forced component in the observed record along with statistical approaches such as observation-based large ensembles. Emphasis will be given to the holistic definition of modes of MDV that extend beyond simple indices, and for example include subsurface ocean parameters. Furthermore, we will investigate how the representation of MDV relates to the reliability of climate predictions and climate sensitivity. Through this analysis we aim to better understand model differences and develop post-processing techniques to produce more reliable climate predictions and projections. We will create a catalogue of available model and observational datasets and current statistical approaches that can be used to analyze them. This will include post-processed datasets (e.g., on a common grid) that will be accessible on a single analysis platform that all working group RF participants will have access to.

After beginning this scoping analysis with existing datasets, a key aim for this component is to be prepared to perform an analysis of CMIP7 data as soon as it starts to become available (the deadline for submission to the CMIP7 Assessment FastTrack is currently January 2027. This effort will be initiated through the organization of a hackathon, likely a coordination with other CMIP7 hackathon events, and the participants will be divided into sub-groups to tackle a subset of Earth System Variables and/or regions with a few focussed publications that will inform the IPCC report and place the CMIP7 ensemble within the context of prior model ensembles. In the capacity building hackathon exercise, early career scientists will gain experience in the analysis of long term trends in large climate datasets and analysis codes will feed into other components.

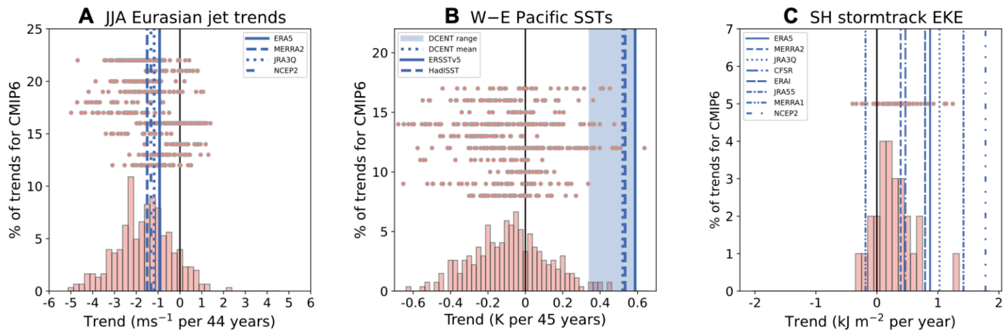

Fig. 1. Three examples of trends over the historical record. (**A**) Success. (**B**) Discrepancy. (**C**) Uncertain situation. (A) The 1980–2023 trends in the summertime Eurasian jet stream following the method of (*69*). It shows the trend in June, July, and August (JJA) area averaged 200-hPa zonal wind over 30°E to 120°E, 35°N to 45°N for (salmon pdf) 303 members from 41 CMIP6 models, (salmon dots) from bottom to top: five CMIP5-era large ensembles (CanESM2, CESM1-CAM5, CSIRO-Mk3-6-0, GFDL-CM3, and MPI-ESM) and six CMIP6-era large ensembles (ACCESS-ESM1-5, CanESM5, EC-Earth3, MIROC6, MPI-ESM1-2-LR, and CESM2). Observation-based data are shown in blue. (B) The 1979–2023 trends in annual mean difference between SST averages over 5°S to 5°N in the Western Pacific (110°E to 180°E) and Eastern Pacific (180°E to 280°E) as in (*53*) for (salmon pdf) 271 members from 41 CMIP6 models and (salmon dots) the same large ensembles as in (A). Observation-based data are in blue, and the blue range shows the minimum to maximum range of the 200-member ensemble of the DCENT observational dataset (*76*). (C) The 1979–2018 trends in SH winter (June-July-August) storm track activity as measured by 2.5- to 6-day band-pass filtered eddy kinetic energy vertically integrated and averaged over 40°S to 75°S for one member from 26 CMIP models and the CESM2 large ensemble [data taken from (*84*)].

Component 2 *New knowledge on MDV using reconstructions of the last millennium*

Climate reconstructions over the last millennium offer an exciting opportunity to deepen our understanding of MDV. Not only is the period sufficiently long to evaluate the natural drivers of MDV, but there is a wealth of data and extensive ensembles of ESM simulations. Furthermore, there are many efforts to expand paleo-proxy archives, rescue and digitize historical instrumental records, and develop multi-proxy reconstructions. There are also emerging paleo-climate reanalyses, providing a first dynamically-consistent, observationally-constrained picture of the period (e.g., Sun et al. 2025). We will conduct an intercomparison of available paleo-based reanalyses and multi-proxy reconstructions and paleo2K simulations to deepen understanding of MDV, identifying robust features and highlighting uncertainties.

In addition, we will work with the paleo-climate community to enhance reconstructions, through addressing limitations in the current techniques and data. For example, several methodologies have recently been tested to better account for oceanic dynamics in paleo-reanalysis, including the use of Linear Inverse Models, surrogate models, and even fully coupled reanalyses. We will also identify critical data gaps through analysis and experiments. For example, adding more proxy archives from the southern hemisphere and subsurface ocean could be crucial to constraining MDV in paleo-reanalysis.

Component 3 – Improving the simulation of MDV and forced trends

A concerted effort is required to improve the simulation of MDV and the forced trend. These share common errors that result mainly from the parameterisation of unresolved processes. To guide the development of improved models, we will conduct coordinated process-based experiments to understand the diversity in the simulated MDV and the externally forced response. These will consist of both free-running and initialised decadal predictions for different phases of MDV. We implement different forcings, including idealised and regional aerosol loadings, and may also perturb parameters. Coordinated multi-model experimentation is important to obtain robust results and identify sources of uncertainty and impact of model bias. In addition, limited experimentation with km-resolution climate models, which rely less on parameterisations, can also be fruitful. There is interest among members of the RF and beyond to perform such experiments.

This will stimulate the development of model calibration approaches utilizing historical observations and improved metrics for MDV. Comprehensive coupled data assimilation and history matching using machine learned emulators are two approaches to tune parameters, which have been tested in state-of-the-art models. These efforts can provide uncertainty bounds on parameter values, leading to more reliable predictions and better constrained climate change projections.

Component 4 State-of-the signal

The group will synthesize the work of components 1 to 3 into a "state-of-the-signal" summary, which will be a comprehensive summary of our current state of knowledge on historical trends and the relative contributions of internal variability and external forcing to them and their representation in Earth System Models. This peer reviewed publication will build on the review produced by Simpson et al. (2025) with new analysis and an up-to-date summary of the literature on this topic, with new results obtained from CMIP7 simulations. Our hope is that this will serve as a blueprint for future continued updates of the "state-of-the-signal".

Our ultimate goal is that this work will also feed into the development of a web-based portal for near real-time updates on the modelled representation of historical trends i.e., something that resembles climate reanalyzer (https://climatereanalyzer.org/) and the global temperature dashboard (https://dashboard.theclimatebrink.com/#global) but for the comparison of long-term trends between models and observations. The activities of this working group will inform on the design of this portal and we will seek resources from other funding opportunities to allow it to come to fruition and we will explore multiple possibilities for hosting such as site, such as at NSF NCAR or on a CLIVAR site if CLIVAR offered this support, or via commercial options.

<u>Timeline, assuming the RF starts Fall of 2026</u>

2026

- Establish task teams for the four activities.
- Online workshop to introduce the RF, raise awareness of existing observational and model datasets, including paleo-reanalysis, and discuss plans for systematic analysis.
- Start analysis of ESM data from older archives to prepare for CMIP7 release..

2027

- Hackathon on CMIP7 analysis
- Aim to submit manuscript(s) summarizing the representation of historical trends in CMIP7 simulations ahead of the AR7 March 15th 2027 deadline
- Initiate a more comprehensive "state-of-the-trends" summary for publication
- Develop initial list of multi-variate fingerprints of MDV modes
- Initiate paleo-data intercomparison, with data sharing
- Tier-1 coordinated model experiments for component 3 initiated.

2028

- Online/in-person mid-term workshop on progress within the RF
- Finalize "state-of-the-trends" publication
- Publish CLIVAR exchange articles on:
    - Approaches to bias correct predictions, constrain projections

- MDV in paleoclimate reanalysis
- Analysis of the RF model experiments

2029

- Start working on review paper on the progress of the RF
- Finalize all coordinated analysis and experiments, and data sharing with the wider community
- Publish community papers on the experiments.
- Summer school, focussed on the new paleo-climate reanalysis/reconstructions; coordinated model experiments; model calibration.
- Discuss approaches on portal design and content with the RF

2030

- Final workshop
- Finalize and publish review paper on RF findings

In addition, conference sessions at EGU, AGU, and AOGS will be organized annually.

5. Deliverables (max. 250 words)

- Peer reviewed synthesis, perspective, and review papers in support of IPCC AR7, including a "State-of-the-signal" summary assessment on the current knowledge of historical trends and the ability of climate models to reproduce them (building on the current "State of the Climate")
- Webinars for community engagement and workshops to coordinate and synthesize research and to design model experiments.
- Special sessions at international conferences and meetings on multi-decadal variability and model representation of long-term trends.
- A new set of approaches for postprocessing climate projections and for comparing model trends with observations. A curated set of postprocessed variables in a common format along with observational data that can be readily updated.
- A repository of analysis scripts for producing the above curated archive.
- Coordinated model simulations for the broader climate dynamics community.
- Recommendations on model calibration approaches to improve climate predictions and better constrain regional climate change
- If time and resources allow, a regularly updated web-based portal with expert guidance on observed trends and their representation in models (analogous to the climate reanalyzer https://climatereanalyzer.org/)

6. Relationship with other CLIVAR and WCRP activities as well as other international programs (max. 100 words)

The RF will connect various CLIVAR and international activities ongoing in different panels on MDV, predictability and sensitivity. We see synergies with the TROPICS working group, Forced Component Estimation Statistical Method Intercomparison Project, The Regional Aerosol Model Intercomparison Project, The Large Ensemble Single Forcing Model Intercomparison Project (LESFMIP), PAGES 2k Network, Cloud Feedback Model Intercomparison Project, Decadal Climate Prediction Project. We see areas in which our research overlaps and where there will be mutual benefit from sharing data and expertise.

The co-chairs are also active participants in WCRP and CLIVAR activities. Co-chair Keenlyside and Stuecker are members of the CLIVAR TBI RF. Co-chair Shaw is a member of the CLIVAR Climate Dynamics Panel and Working Group 1 of the Explaining and Predicting Earth System Change WCRP Lighthouse Activity. Co-Chair Simpson is a member of the SSC for US CLIVAR. Shaw and Simpson are both active participants in the WCRP EPESC/APARC LEADER activity.

7. Key References (max. 10 most relevant publications)

Beverley, J. D., M. Newman, and A. Hoell (2024), Climate model trend errors are evident in seasonal forecasts at short leads, npj Climate and Atmospheric Science, 7(1), 285, 10.1038/s41612-024-00832-w.

Chikamoto, Y., A. Timmermann, M. J. Widlansky, S. Zhang, and M. A. Balmaseda, 2019: A Drift-Free Decadal Climate Prediction System for the Community Earth System Model. J. Climate, 32, 5967–5995, https://doi.org/10.1175/JCLI-D-18-0788.1.

Forster, P. M., C. Smith, T. Walsh et al. 2024, Indicators of Global Climate Change 2024: annual update of key indicators of the state of the climate system and human influence, Earth System Data Science, https://doi.org/10.5194/essd-2025-250

Hu, Ingo Richter, Yuko Okumura, Natalie Burls, Noel Keenlyside, Rhys Parfitt, Katinka Bellomo, Alessio Bellucci, Riccardo Farneti, Alexey V. Fedorov, et al. Unraveling the Complexity of Global Climate Dynamics: Interactions among El Niño–Southern Oscillation, Atlantic Meridional Overturning Circulation, and Tropical Basins Across Different Timescales. Ocean-Land-Atmos Res. 2025;4:0096.https://spj.science.org/doi/full/10.34133/olar.0096

Omrani, N.-E., N. Keenlyside, K. Matthes, L. Boljka, D. Zanchettin, J. H. Jungclaus, and S. W. Lubis (2022), Coupled stratosphere-troposphere-Atlantic multidecadal oscillation and its importance for near-future climate projection, npj Climate and Atmospheric Science, 5(1), 59, 10.1038/s41612-022-00275-

Shaw, T. A., P. A. Arias, M. Collins et al. 2024: Regional Climate Change: Consensus, Discrepancies and Ways Forward, Frontiers Climate, https://doi.org/10.3389/fclim.2024.1391634

Simpson, I. R., T. A. Shaw et al. 2025: Confronting Earth System Model trends with observations, Science Advances, DOI: 10.1126/sciadv.adt8035

State of the Climate 2023, https://ametsoc.net/sotc2023/SoCin2023_FullReport.pdf

Smith, D. M. et al (2022) Attribution of multi-annual to decadal changes in the climate system: The Large Ensemble Single Forcing Model Intercomparison Project (LESFMIP), Front. Clim., 4:955414, doi:10.3389/fclim.2022.955414

Sun, H., Lei, L., Liu, Z., Ning, L., & Tan, Z.-M. (2025). An online paleoclimate data assimilation with a deep learning-based network. Journal of Advances in Modeling Earth Systems, 17, e2024MS004675. https://doi.org/10.1029/2024MS004675