Response to Reviewer 2

Title: Climate Modes evaluation datasets from CMIP6 pre-industrial control simulations and observations

General assessment

This manuscript presents a dataset of climate variability indices derived from pre-industrial control (piControl) simulations of multiple CMIP6 models. The authors compute a suite of well-known modes of climate variability (including ENSO, AMO, IPO, NAO, SAM, IOD, and SDM) using established methodologies, and provide these indices as a consolidated archive intended to facilitate studies of internal climate variability. Additionally, the authors compare the spatial structure and temporal variability of these modes with observations, aiming to showcase consistency or biases with respect to observations.

The work is technically competent and clearly reflects substantial effort. Additionally, the manuscript is generally well written. However, after reviewing both the manuscript and the dataset itself in light of ESSD's evaluation criteria, I conclude that the submission does not align with the journal's requirements for significance, uniqueness, and added value to the scientific community. More specifically, I consider that the produced data substantially overlaps with already existing, broadly distributed, and thoroughly documented data. Besides, I consider that those aspects of the data presented in the manuscript that do not exactly overlap with previous literature do not represent enough novelty to merit publication. Therefore, my recommendation is to reject this manuscript from ESSD. Below, I expand on these points and present potential alternatives for future fruitful publication.

Reply: We thank the reviewer for the useful comments, which helped us to improve the manuscript. Below we have provided a reply to each comment in detail, with replies in blue color.

The two major problems:

Not enough novelty

The authors provide 8 indices. From those, four (Atlantic Multidecadal Oscillation, AMO; North Atlantic Oscillation, NAO; Southern Annular Mode, SAM; Interdecadal Pacific Oscillation, IPO) are publicly available and operationally maintained by the Climate Variability Diagnostics Package (CVDP; Phillips et al., 2014; Maher et al., 2025—please see supplementary material) community effort using the same methodology that the authors provide. Apart from the time period that the authors cover, CVDP also provides these indices for other CMIP versions, as well as for the broader suite of experiments that they encompass beyond the pre-industrial simulations.

The authors also provide two ENSO indices based on Empirical Orthogonal Functions for the Central Pacific and Eastern Pacific flavors of ENSO. Although these differ from the more broadly used ENSO indices provided, for example, in the CVDP package, the manuscript does not convincingly demonstrate that the chosen EOF-based definitions are more robust across models, superior to existing alternatives, or uniquely enabled by the dataset. The authors justify using these indices by stating that EOF-based indices are more robust to model biases, but, under my consideration, a deeper review of the literature and quantitative analyses are needed to justify that statement. Just as an example, I recommend reviewing the recently developed RONI index (L'Hereux et al., 2025). Furthermore, beyond the mere indices, community tools already provide composites, spectra, and regressions for piControl simulations (link). This same point applies to the Indian Ocean Dipole index.

As for the remaining mode, the Subsurface Dipole Mode, this component appears less covered by existing diagnostics frameworks and could represent a meaningful contribution if foregrounded, more thoroughly documented, and clearly positioned as *the* primary data product. If the authors decide to pursue this mode as the primary contribution of the paper, then a more thorough explanation should be provided, and the index should also be computed for the whole suite of CMIP experiments beyond pre-industrial simulations in order to make the paper stand out as a significant contribution useful to the broader community. The first author already published a paper on this mode of variability (CITE), which could serve as a foundation for a more complete data contribution to the community.

In summary, from the eight provided modes, four are fully covered by past literature, three are partially covered and deeper demonstration of superiority with respect to existing tools needs to be provided to make this paper stand out, and one is a novel contribution but not highlighted or explained in enough detail, and not computed for a broader set of experiments in order to have a significant impact.

Reply: While the datasets generated and the analysis undertaken here has many similarities to the Climate Variability Diagnostics Package (CVDP) data repository (Data Repository | Community Earth System Model), they also exhibit important differences. These differences, summarized in the accompanying Table R1 (shown in red), reflect variations in the definitions, filtering methods, and latitude ranges used for each climate mode in the simulations, as well as the introduction of additional newly proposed indices (for SAM, and NAO).

We will introduce the EOF based 500 hPa geopotential height definition for both the North Atlantic Oscillation (NAO) over the North Atlantic and the Southern Annular Mode (SAM) over the extratropical Southern Hemisphere. This approach captures the large scale tropospheric circulation, making it particularly useful for exploring the underlying physical mechanisms and associated climate footprints of these modes. In contrast, definitions based on sea level pressure (SLP) primarily reflect surface climate impacts, and therefore serve a complementary purpose. These different definitions are also valid and follow standard published approaches. All our data sets now give an alternative data set of the common climate modes from which the community can use to assess the level of uncertainty of the scientifically valid approaches taken by CVDP data repository and by us with this collection. There are potentially important differences, and these differences would inform the level of uncertainty associated with different estimates of the strength of these common modes. A key aspect of science and especially IPCC assessments is to test the uncertainty and to understand the level of differences caused by equally valid scientific approaches. This additional collection now allows ensemble approaches as well. It is for these reasons we believe that the publication of this collection has value. Already the collection has been downloaded more than 400 times showing its value.

Index	**Our Definition (ESSD)**	**Reviewer referred Definition (CVDP)**	**Definition Status (ESSD/ CVDP / New Definition)**
AMO (Atlantic Multidecadal Oscillation)	Average of 10 years low pass filtered detrended monthly SST Anomaly average over the North Atlantic (0-60ºN,75ºW-7.5ºW) (Enfield et al. 2001; Wang et al. 2009; Deser et al. 2021)	Monthly index timeseries defined as area-weighted SST* anomalies averaged over the North Atlantic (0:60°N, 80°W:0°E), where SST* denotes that the global (60°S:60°N) mean SST anomaly has been removed at each timestep. Pattern created by regressing SST* anomalies onto the index timeseries and smoothing with a 9-point spatial filter. Low pass-filtered timeseries (black curve superimposed on the monthly timeseries) is based on a 61-month running mean. (Trenberth and Shea 2005)	While both definitions represent the AMO, they differ in methods, our approach uses basin-averaged, linear trend removed, low-pass filtered SST anomalies, whereas the CVDP index removes the global mean SST anomaly to isolate Atlantic relative variability. Because these approaches treat the global warming signal differently, using both helps test robustness and capture complementary variability. The CVDP method suits forced simulations, while our detrending approach is only implemented for piControl runs, where it primarily removes model drift.
IPO (Interdecadal Pacific Oscillation)	1st EOFPC of 10 years low pass filtered detrended monthly SST anomaly averaged over the Pacific Ocean (70ºS-70ºN, 120ºE-80ºW). (Dong and McPhaden 2017; Han et al. 2014; Power et al. 1999)	Monthly index timeseries defined as the leading principal component (PC) of 13yr low pass filtered Pacific (40°S:60°N, 110°E:70°W) area-weighted SST* anomalies, where SST* denotes that the global mean SST anomaly has been removed at each timestep. Pattern created by regressing SST anomalies (in Celsius) at each grid box onto the normalized PC timeseries. At least 40 years of data are required for the IPO to be calculated. (Mehl and Hu 2007)	The main difference lies in the low pass filter and the latitude range. Many literatures use broader latitude bands (60°S or 70°S) to capture a wider range of variability and potential hemispheric contributions to the IPO. While this omission from CVDP is not expected to significantly affect the IPO pattern, different definitions can be useful to quantify structural uncertainty, assess the sensitivity

			of results to methodological choices.
El Niño (EP El Niño and CP El Niño)	1st and 2nd EOFPC of 2-9 years band pass filtered detrended monthly SST anomaly over the tropical Pacific Ocean (120ºE-80ºW, 30ºS-30ºN). EOFPC1 (EOFPC2) represents the EP El Niño (CP El Niño). (Xu et al 2017) Here EP El Niño is denoted as ENSO1 and CP El Niño is denoted as ENSO2.	**Nino3=**Area-averaged SST anomalies (in Celsius) computed over 5°S:5°N, 90:150°W. Red/blue shading denotes positive/negative departures from the best-fit linear trend line. **Nino3.4=**Area-averaged SST anomalies (in Celsius) computed over 5°S:5°N, 120:170°W. Red/blue shading denotes positive/negative departures from the best-fit linear trend line. **Nino4=**Area-averaged SST anomalies (in Celsius) computed over 5°S:5°N, 160°E:150°W. Red/blue shading denotes positive/negative departures from the best-fit linear trend line.	The definition used in this study is EOF based (Qi et al. 2021; Singh et al. 2011; Xu et al. 2017), whereas the CVDP definition relies on an area averaged time series index.
SAM (Southern Annular Mode)	1st EOFPC of detrended monthly sea level pressure anomaly south of 20ºS (Cai and Cowan 2007; Miller et al. 2006)	Seasonal/annual PSL averages are formed, square root of the cosine of latitude weighting is applied, and then the 1st (SAM), 2nd (PSA1) and 3rd (PSA2) EOFs and associated principal component (PC) timeseries are computed over 20:90°S, 0:360°E. Patterns created by regressing global PSL anomalies (in hPa) onto normalized PC timeseries. (Thompson and Wallace 2000)	The definition is almost identical; however, an additional definition is introduced by computing the EOF of the 500 hPa geopotential height field over North Atlantic Ocean, where the first leading mode defines the NAO. (Watcher et al. 2020; Minglin & Xiuzhen 2022)
NAO (North Atlantic Oscillation)	1st EOFPC of detrended monthly sea level pressure anomaly over the North Atlantic Ocean (90ºW-40ºE,20ºN-80ºN) (Hurrel et al. 2003; Hurrel & Deser 2009)	Seasonal/annual PSL averages are formed, square root of the cosine of latitude weighting is applied, and then the leading EOF and associated principal component (PC) timeseries are computed over 20:80°N, 90°W:40°E. Pattern created by regressing global PSL anomalies (in hPa) onto normalized PC timeseries. (Hurrel et al. 2005)	The definition is almost identical; however, an additional definition is introduced by computing the EOF of the 500 hPa geopotential height field over Southern Hemisphere, where the first leading mode defines the SAM. (Davini et al. 2012; Pinto et al. 2012; Simpson et al. 2024)
SDM (Subsurface Dipole Mode)	1st EOFPC of detrended monthly upper 500m OHC anomaly over the tropical Indian Ocean		The SDM index is proposed for the first time in this study and has been appreciated by the reviewer.

	(40ºE-110ºE, 20ºS-25ºN) (Mohapatra & Gnanaseelan, 2021)		
IOD (Indian Ocean Dipole)	2nd EOFPC of detrended monthly SST anomaly over the Indian Ocean (40ºE-110ºE, 20ºS-25ºN) (Krishnamurthy & Kirtman, 2003)	Defined as the difference between area-averaged SST anomalies (in Celsius) computed over 10°S:10°N, 50:70°E and area-averaged SST anomalies computed over 0:10°S, 90:110°E. (Saji et al. 1999)	The definition used in this study is EOF based, whereas the CVDP definition relies on an area-averaged time series index.

Table R1. Comparison of climate mode definitions used in this study (including newly introduced definitions) and those reported in CVDP.

In response to the reviewer's comment, we acknowledge that widely used ENSO indices provided by community tools (e.g., the CVDP package) are also applied for analysing ENSO variability. Our intention in including EOF based indices for Eastern Pacific (EP) and Central Pacific (CP) ENSO is not to claim that these definitions are superior to existing indices, but rather to use them as complementary diagnostics to characterize ENSO diversity across models. We have now revised the text to clarify this point in the manuscript. Several studies have identified EP and CP ENSO based on the first two leading EOF of tropical Pacific SST variability (e.g., Singh et al., 2011; Xu et al., 2017; Qi et al., 2021), with the EOF patterns closely resembling canonical EP and CP ENSO in multi model assessments (Xu et al., 2017).

Additionally, the Indian Ocean Dipole (IOD) has been described as the second leading EOF mode of tropical Indian Ocean SST variability, and the commonly used Dipole Mode Index (DMI) as the IOD index is defined as the SST difference between the western and southeastern equatorial Indian Ocean. It is observed that the regions of highest variance in the EOF pattern coincide with these two regions, highlighting them as the primary centres of variability (Saji et al. 1999). We would like to clarify that the EOF based indices are used here as an additional diagnostic to examine the spatial and temporal structure and representation of modes across CMIP6 piControl simulations, complementing the area averaged indices.

Unfair comparison between models and observations

Although the exploration of the differences between the pre-industrial simulations and observations is interesting, its usefulness to the climate science community is somewhat limited. On one hand, it is methodologically inappropriate to rank model skill by directly comparing spatial patterns from pre-industrial control simulations with the historical observational record. It has been demonstrated that human emissions deeply altered the characteristics of modes of variability (e.g., Klavans et al., 2025), and substantial differences between pre-industrial and historical climates are therefore expected. It is not possible to know with certainty, solely based on the presented analysis, which version of the pre-industrial climate is more or less wrong. On the other hand, a deep and thorough comparison of modes of variability in models vs. observations is already available (e.g., Fasullo et al., 2020; Maher et al., 2025). Furthermore, these papers also take into account observational uncertainty and the fact that we only have one observed realization of internal variability (for the historical forcings - not pre-industrial) vs. the multiple realizations available in the CMIP archives. If assessing the correctness of the representation of modes of variability in Earth system models is the goal, authors should focus on aspects that have not been thoroughly explored before, and in colloquial words, be mindful of comparing apples with apples. Furthermore, if this is the main scope of the paper, I would recommend submitting to a journal with more emphasis on Earth system model evaluation rather than scientific data publication.

Reply: We thank the reviewer for this important and insightful comment. We agree that a direct comparison between piControl simulations and the historical/observational record is not appropriate for a rigorous assessment of model skill. In the revised manuscript, we have clarified that the primary objective of this comparison is not to rank or evaluate model skill, but rather to provide a qualitative assessment of the large scale spatial structures of the derived modes. In this context, the observational datasets are used only as a reference for general pattern realism, and not as a benchmark for quantitative agreement. We also note that, prior to the EOF analysis, long term trends were removed from the observational datasets to minimize the influence of externally forced climate change and to better isolate internal variability. However, we acknowledge that such preprocessing does not fully reconcile the differences between pre-industrial and present day climates.

To avoid potential misinterpretation, we have revised the text to explicitly state these limitations and have removed any language that could be interpreted as a quantitative evaluation of model skill.

Conclusion: Based on the considerations above, I see two viable paths for successful publication of this work:

1. If the focus on a data publication is to be kept, a deeper justification of the methodological decisions (that must be different from the already available community resources) should be provided, showcasing and quantifying robustly why these would be preferred over the “status quo” of publicly available data. The novelty and usefulness of the data beyond what already exists should be explicitly proven.
2. If the authors would like to maintain and strengthen the models vs. observations comparison, this should focus on aspects that have not been covered before, which should be accompanied by a broad and thorough review of the existing literature. Additionally, care should be taken about the validity of the comparison in light of different forcings and the presence of internal variability and observational uncertainty.

Reply: Thank you for the thoughtful comments showing ways for publication of the present paper. We stick to the data description paper along with additional derived datasets for climate modes (SAM, NAO).

References

- Phillips et al., 2014. Evaluating Modes of Variability in Climate Models. Eos.
- Maher et al., 2025. The updated Multi-Model Large Ensemble Archive and the Climate Variability Diagnostics Package: new tools for the study of climate variability and change. Geosci. Model Dev.
- L'Heureux et al., 2025. A relative sea surface temperature index for classifying ENSO events in a changing climate.
- Klavans et al., 2025. Human emissions drive recent trends in North Pacific climate variations.
- Fasullo et al., 2020. Evaluation of Leading Modes of Climate Variability in the CMIP Archives. Journal of Climate.

Reply: The above has been added to the manuscript.

Reference

- Davini, P., Cagnazzo, C., Neale, R., and Tribbia, J.: Coupling between Greenland blocking and the North Atlantic Oscillation pattern, *Geophys. Res. Lett.*, 39, https://doi.org/10.1029/2012GL052315, 2012.
- Pinto, J. G. and Raible, C. C.: Past and recent changes in the North Atlantic Oscillation, *Wiley Interdiscip. Rev. Clim. Change*, 3, 79–90, https://doi.org/10.1002/wcc.150, 2012.
- Simpson, I., Hanna, E., Baker, L., Sun, Y., and Wei, H. L.: North Atlantic atmospheric circulation indices: Links with summer and winter temperature and precipitation in north-west Europe, including persistence and variability, *Int. J. Climatol.*, 44, 902–922, https://doi.org/10.1002/joc.8364, 2024.
- Zheng, M. and Li, X.: Distinct patterns of monthly Southern Annular Mode events, *Atmos. Ocean. Sci. Lett.*, 15, 100206, https://doi.org/10.1016/j.aosl.2022.100206, 2022.
- Wachter, P., Beck, C., Philipp, A., Höppner, K., and Jacobeit, J.: Spatiotemporal variability of the Southern Annular Mode and its influence on Antarctic surface temperatures, *J. Geophys. Res. Atmos.*, https://doi.org/10.1029/2020JD033818, 2020.
- Qi, Q., Duan, W., and Xu, H.: The most sensitive initial error modes modulating intensities of CP- and EP-El Niño events, *Dynamics of Atmospheres and Oceans*, **96**, 101257, https://doi.org/10.1016/j.dynatmoce.2021.101257, 2021.
- Singh, A., Delcroix, T., and Cravatte, S.: Contrasting the flavors of El Niño–Southern Oscillation using sea surface salinity observations, *Journal of Geophysical Research: Oceans*, **116**, C06016, https://doi.org/10.1029/2010JC006862, 2011.

- Xu, K., Tam, C.-Y., Zhu, C., and Liu, B.: CMIP5 projections of two types of El Niño and their related tropical precipitation in the twenty-first century, *Journal of Climate*, **30**, 849–864, https://doi.org/10.1175/JCLI-D-16-0413.1, 2017.
- Saji, N. H., Goswami, B. N., Vinayachandran, P. N., and Yamagata, T.: A dipole mode in the tropical Indian Ocean, *Nature*, **401**, 360–363, https://doi.org/10.1038/43854, 1999.