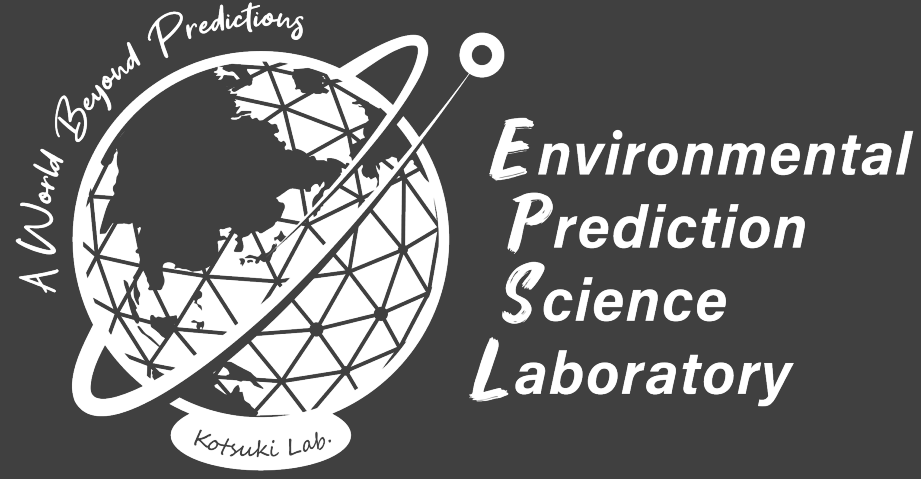


A Comprehensive Benchmarking of Paleoclimate Reanalysis Products Against Major Observational Datasets



Muhammad Abid Khan¹, Atsushi Okazaki^{1,2}

¹ CEReS, Chiba University, ² IAAR, Chiba University
(rabid609@gmail.com)



Background & Motivation

- Understanding past climate is crucial for contextualizing modern climate change and reducing uncertainties in future projections.
- Numerous paleoclimate reanalysis products have been developed, but a systematic, global-scale assessment of their performance is lacking.
- This study fills this gap by providing a comprehensive benchmark to guide data selection in climate research and inform policy

Study Purpose & Scope

Primary Objective: To conduct a rigorous, global evaluation of multiple paleoclimate reanalysis products against leading instrumental reference datasets.

Key Aim: Identify which reconstruction most accurately captures historical climate variability, providing a reliable benchmark for the community.

Data and Method

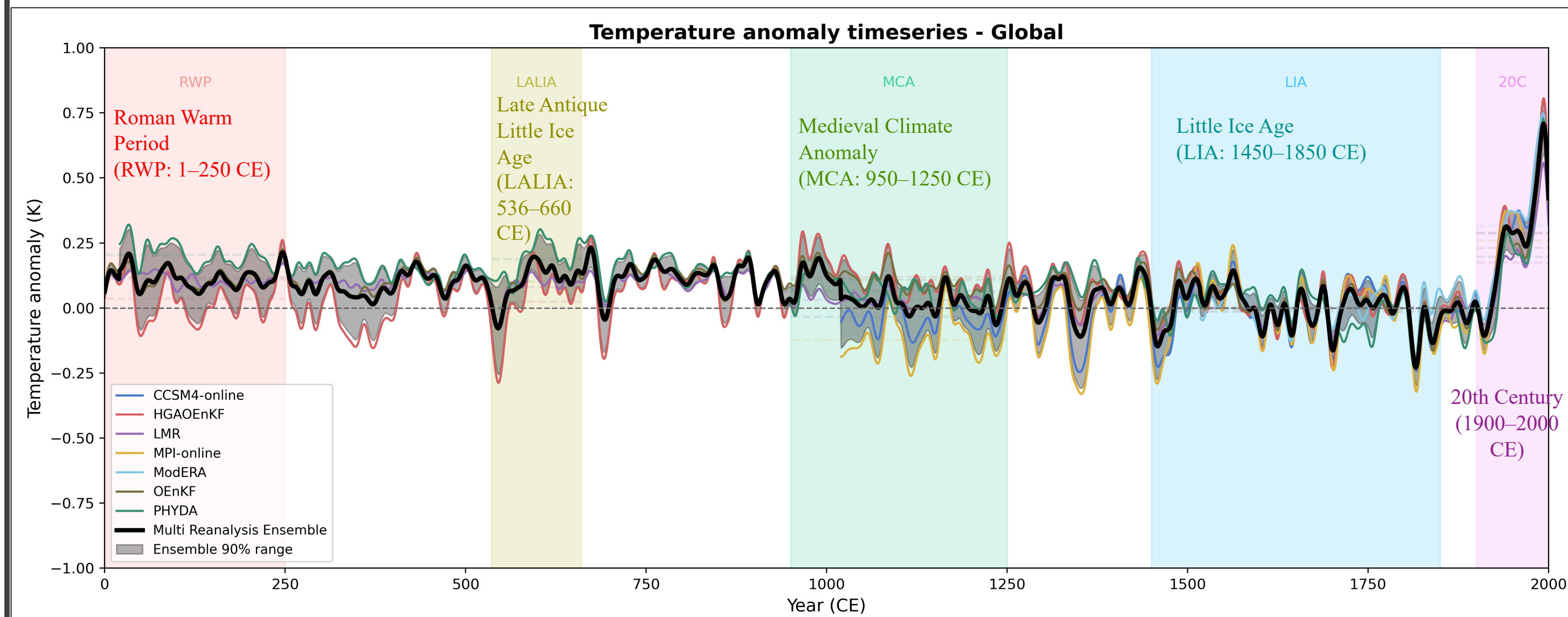
- Paleoclimate Reanalysis:** CCSM4-online, MPI-online, HGAOEnKF, LMR, ModE-RA, OEnKF, PHYDA
- Instrumental Datasets:** HadCRUT5, GISTEMP, 20CRv3, NOAAGlobalTemp, Berkeley Earth
- Data Harmonization Framework:** (a) Regridding, (b) Temporal Alignment, (c) Format Standardization, (d) Harmonized Datasets, (e) Anomaly Calculation

Analysis Framework --> **Applied a multi-faceted statistical approach to assess:**

- Spatial & Temporal Agreement:** Pearson Correlation Coefficient (R)
- Overall, Skill & Error:** Root Mean Square Error (RMSE), Coefficient of Efficiency (CE)
- Variability & Spread:** Standard Deviation, Probability Density Functions (PDFs)
- Ensemble Analysis:** Construction & evaluation of the Multi-Reanalysis Ensemble (MRE) --> $MRE_{t,x} = \frac{1}{N} \sum_{i=1}^N X_{i,t,x}$
- Multi-Metric Comparison:** Taylor Diagrams (synthesizing R, Std. Dev., and RMSE)

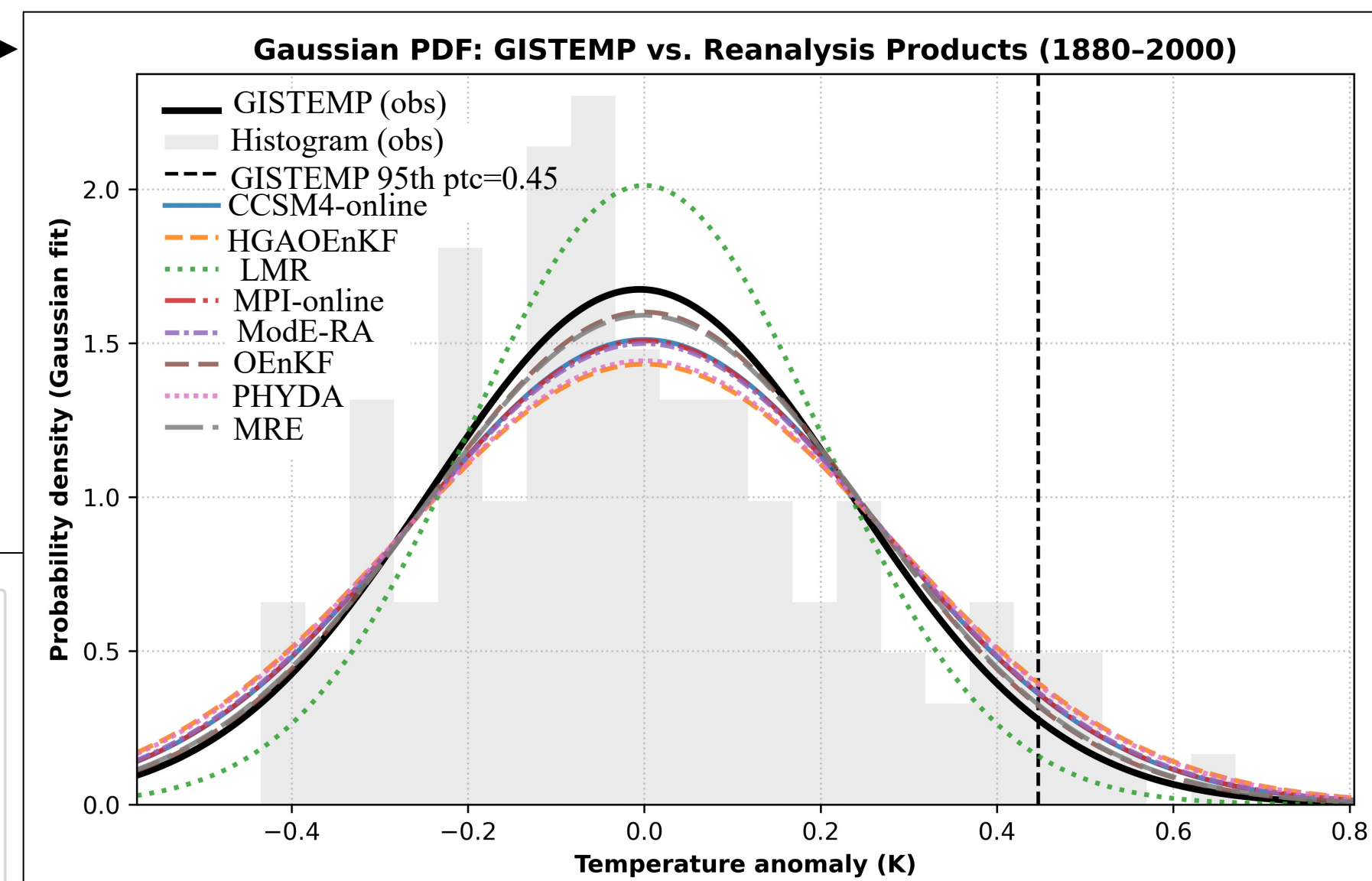
Results & Discussion

Global changes over the past two millennia from multi-reanalysis paleoclimate datasets



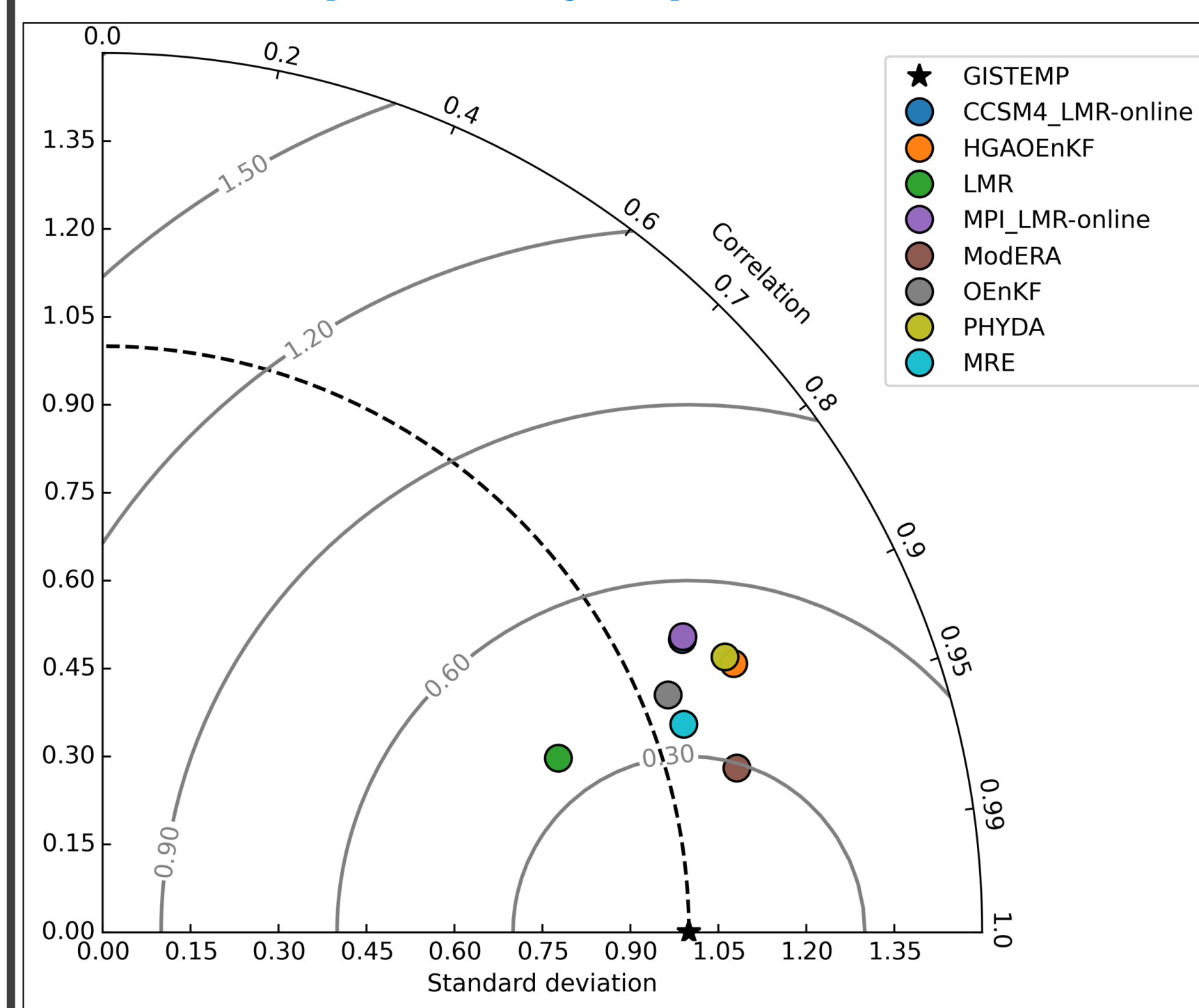
Gaussian PDF comparison

Gaussian probability density functions (PDFs) of annual temperature anomalies for 1880–2000 from reference (GISTEMP) compared with multiple reanalysis products.

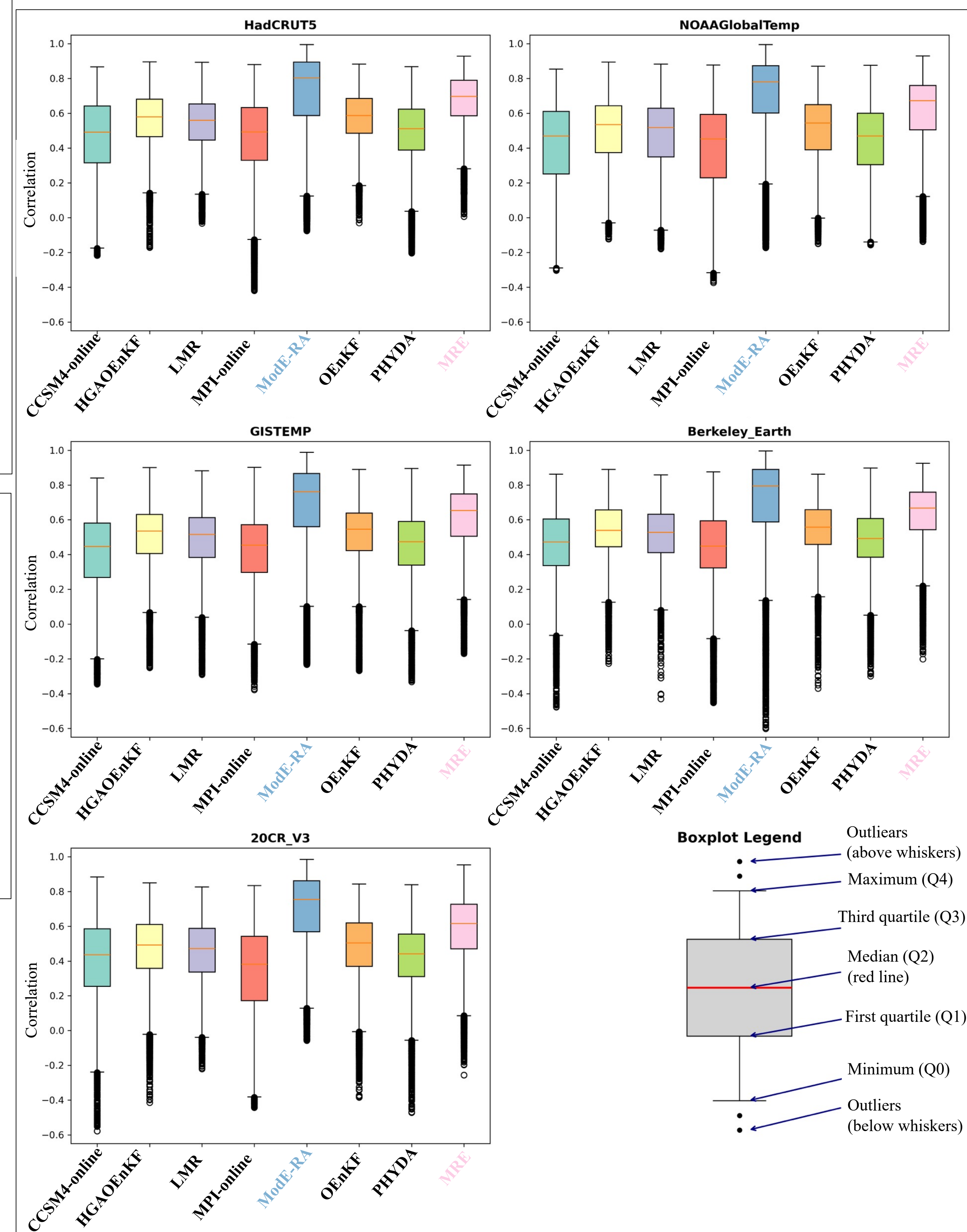


Taylor diagram

Taylor diagram comparing the performance of seven climate reanalysis datasets in reproducing the global mean temperature anomaly relative to the reference (GISTEMP) over 1880 – 2000.



Box plots the spatial distribution of R



Conclusion

- ModE-RA is the top-performing global dataset.
- Show the highest global correlation with observations.
- Most accurately reconstructs historical climate variability.
- Recommended benchmark for paleoclimate applications.