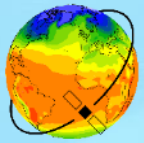
Axel Lauer

Model benchmarking and monitoring

with ESMValTool v2.12.0



ESMValTool
Earth System Model Evaluation Tool

ESMValTool community workshop
Oberpfaffenhofen, Germany,
27-29 May 2024

DLR

New preprocessor

distance_metric

Calculates a distance metric with respect to a given reference dataset over dimensions specified.

metric	unweighted	weighted
Root mean square error (RMSE)	$RMSE_{unweighted} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - R_i)^2}$	$RMSE_{weighted} = \sqrt{\sum_{i=1}^{N} w_i(X_i - R_i)^2}$
Pearson‘s correlation coefficient	$r_{unweighted} = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(R_i - \bar{R})}{\sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{N}(R_i - \bar{R})^2}}$	$r_{weighted} = \frac{\sum_{i=1}^{N}[w_i(X_i - \bar{X})(R_i - \bar{R})]}{\sqrt{\sum_{i=1}^{N}(w_i(X_i - \bar{X})^2)}\sqrt{\sum_{i=1}^{N}(w_i(R_i - \bar{R})^2)}}$
Earth mover‘s distance (EMD)	$EMD = \min_{\gamma \in \mathbb{R}_+^{n \times n}} \sum_{i,j}^{n} \gamma_{ij}\lvert x_i - r_j\rvert$ with $\sum_j^n \gamma_{ij} = \mathrm{p_x(x_i)}$; $\sum_i^n \gamma_{ij} = \mathrm{p_r(r_j)}$	

Extension of monitoring diagnostics

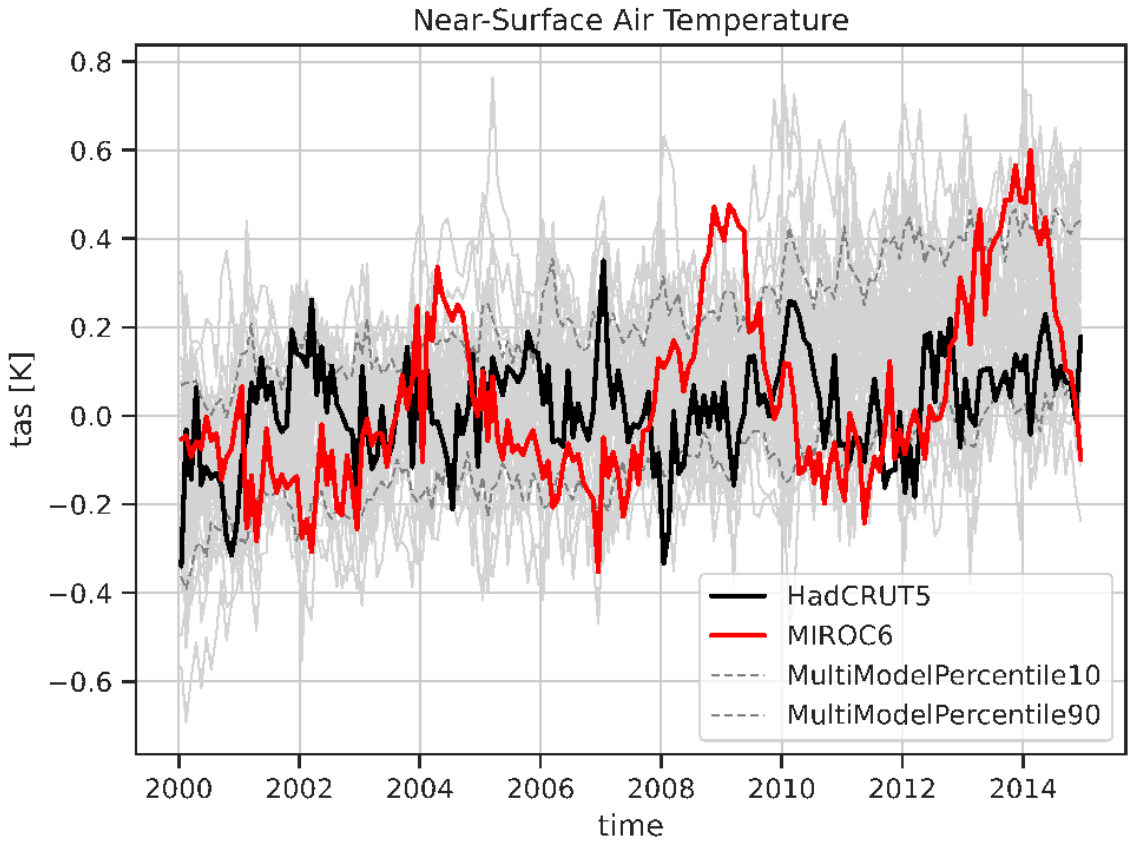
New plot types (multi_datasets.py)

- annual cycles (“benchmarking_annual_cycle”)
- box plots (“benchmarking_boxplot”)
- diurnal cycles (“benchmarking_annual_cycle” and “diurnal_cycle”)
- maps (“benchmarking_map”)
- time series (“benchmarking_timeseries”)
- zonal mean profiles (“benchmarking_zonal”)

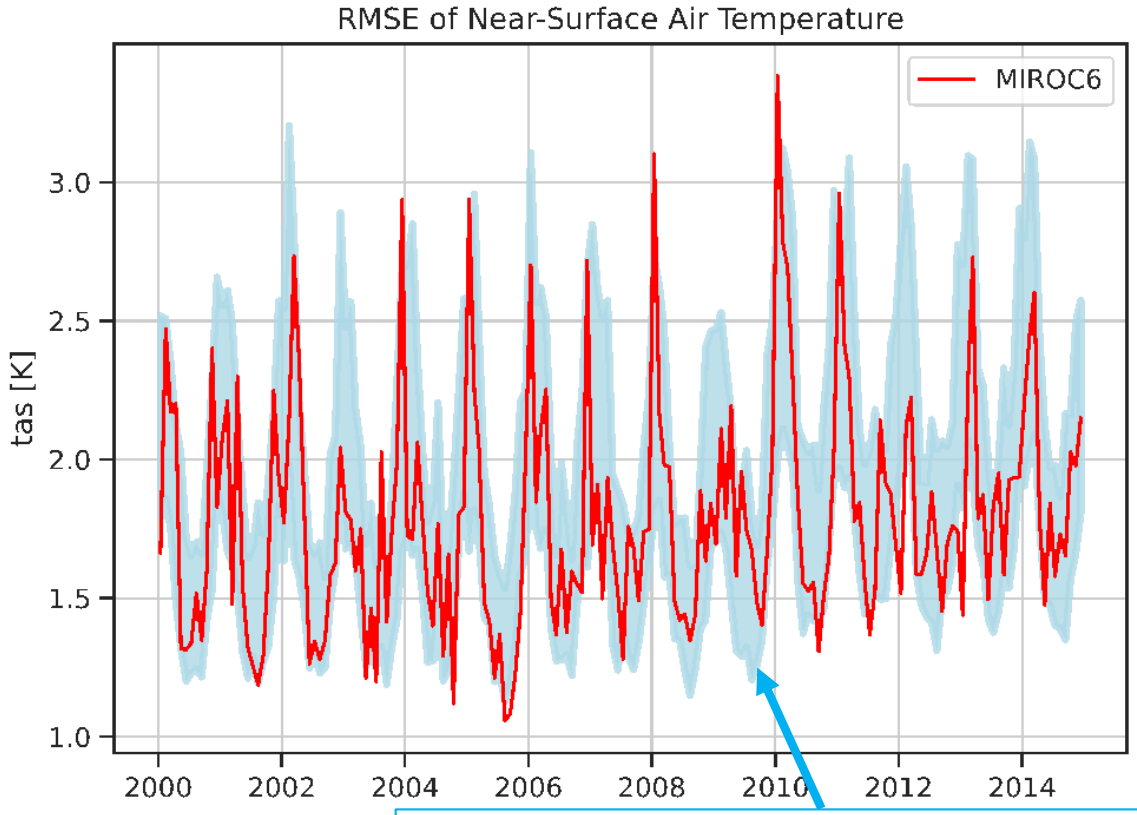
→ **Metrics calculated for an ensemble of models (e.g. CMIP6) can be used for comparison with the results from a select simulation.**

Extension of monitoring diagnostics

Time series (global average anomalies)



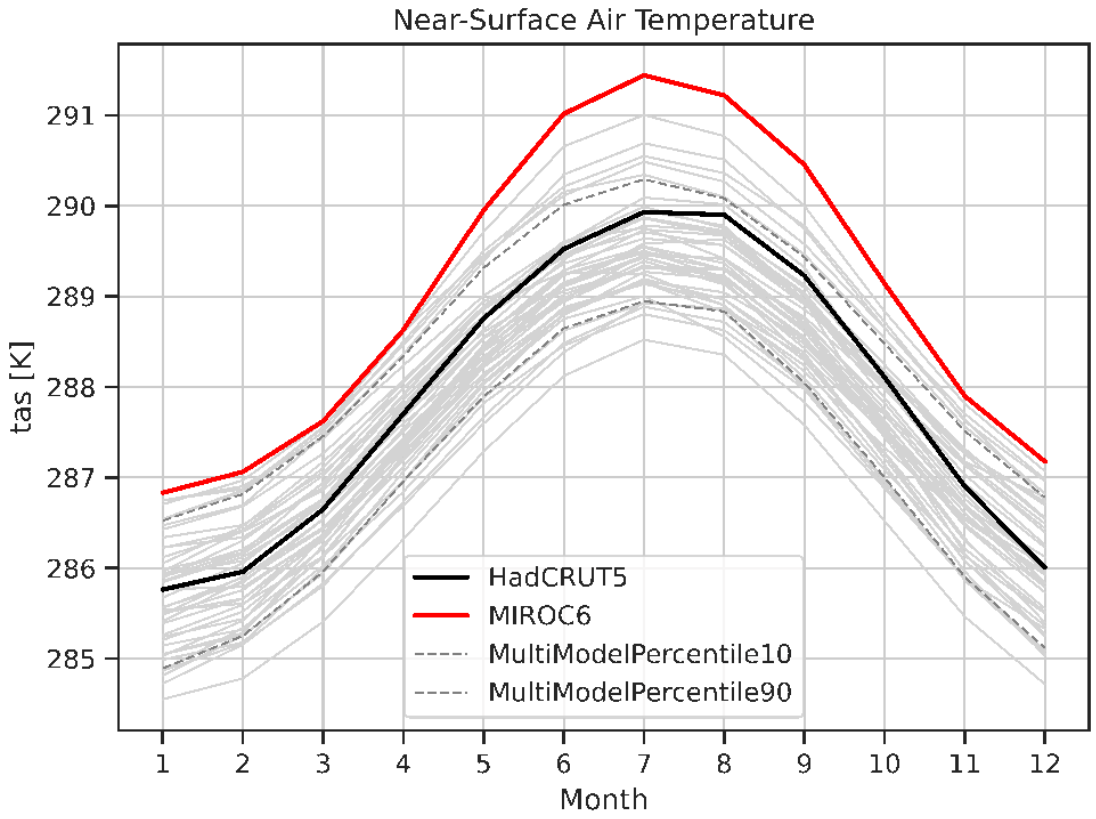
Average RMSE of the monthly mean values at each grid box



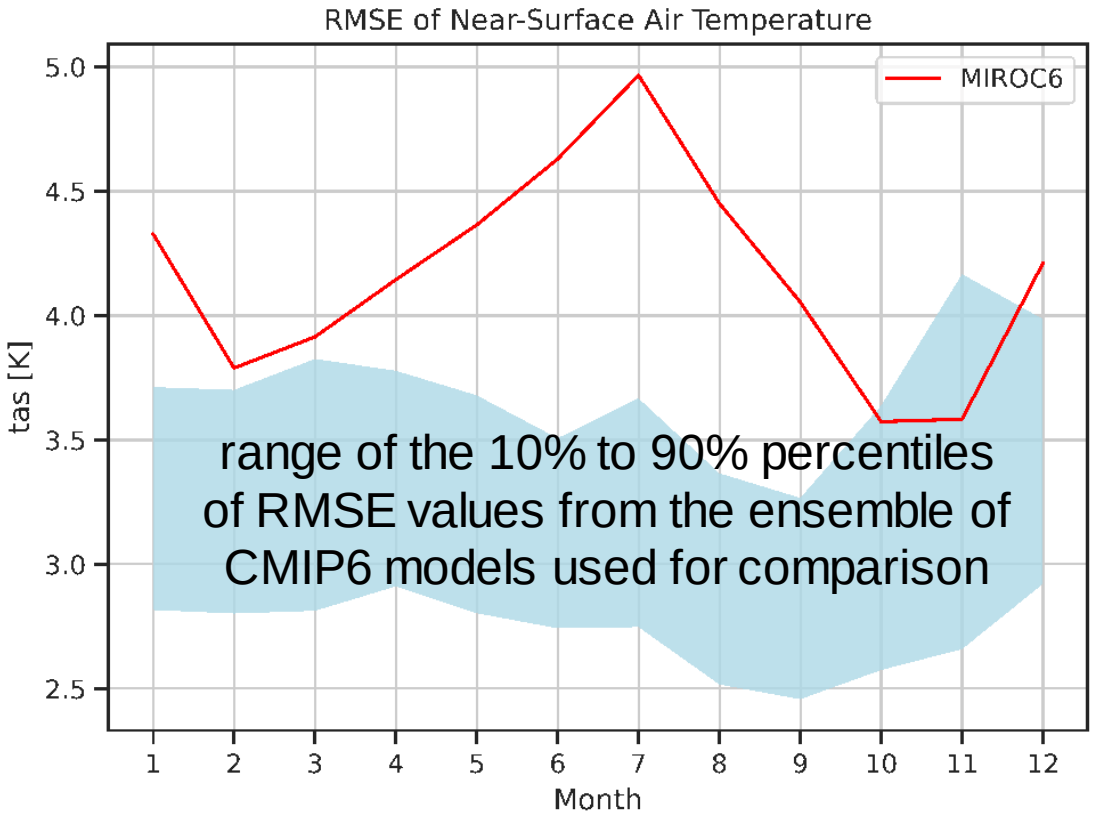
range of the 10% to 90% percentiles of RMSE values from the ensemble of CMIP6 models used for comparison

Extension of monitoring diagnostics

Annual cycle (global average)

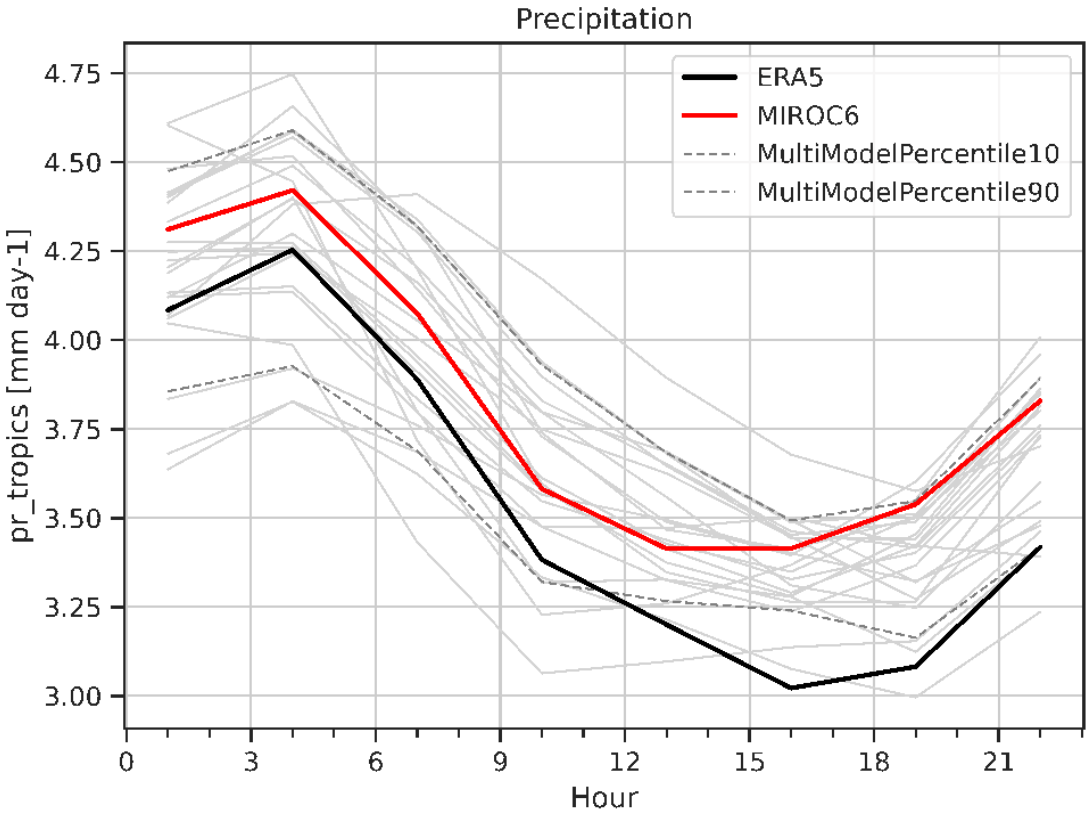


Average RMSE of the annual cycle at each grid box

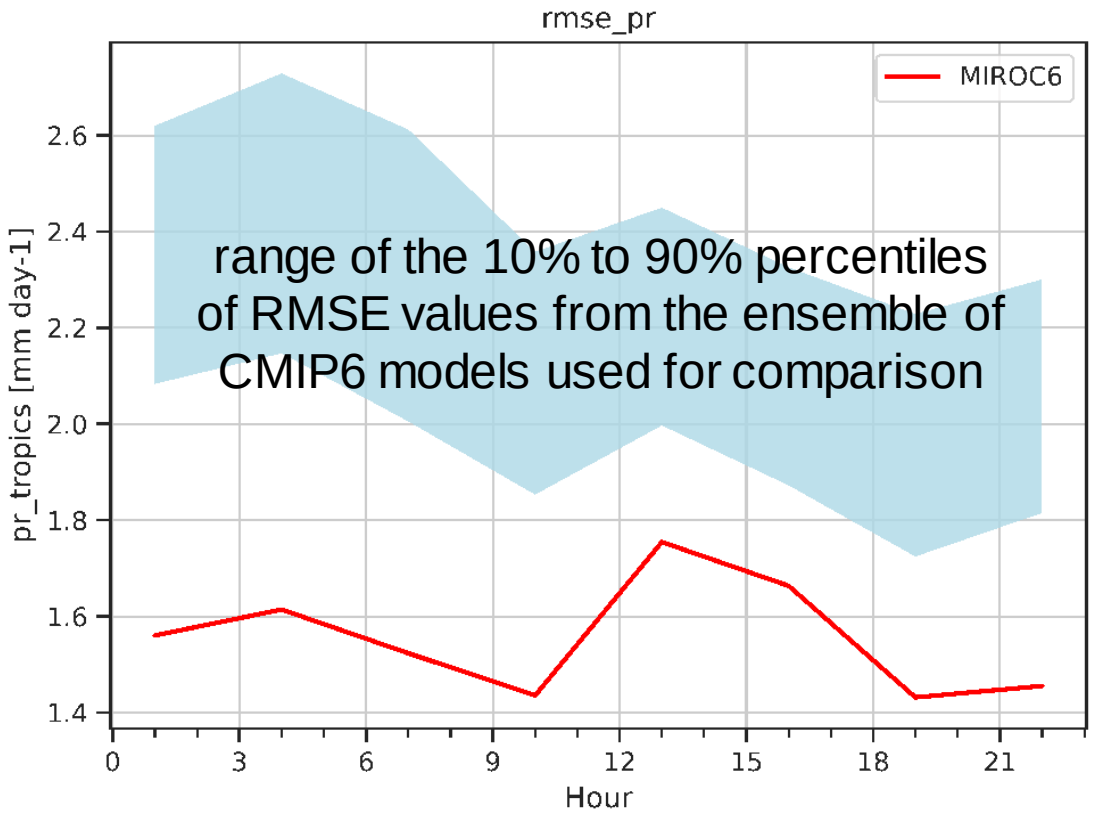


Extension of monitoring diagnostics

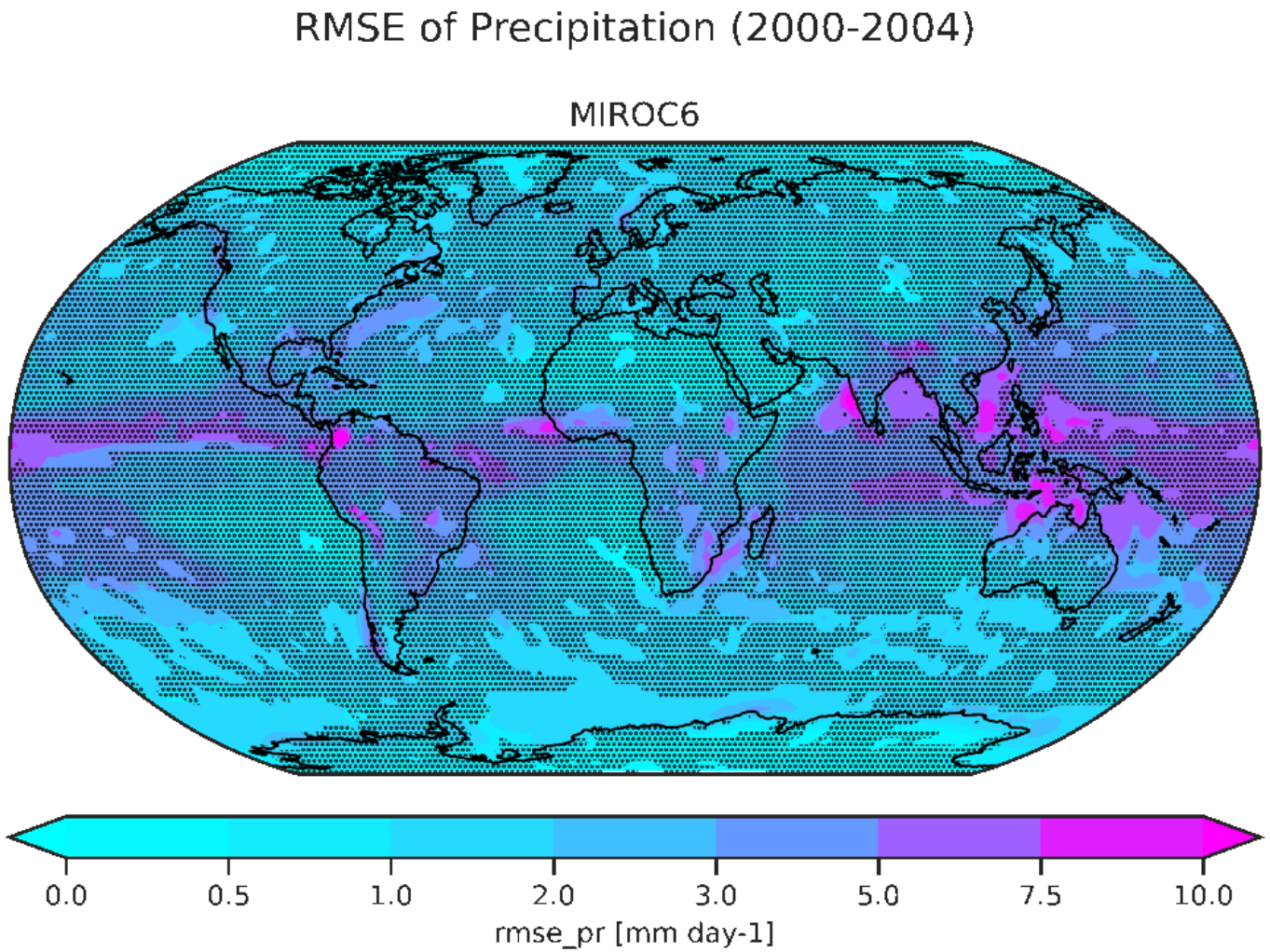
Diurnal cycle (average tropical ocean)



Average RMSE of the diurnal cycle at each grid box

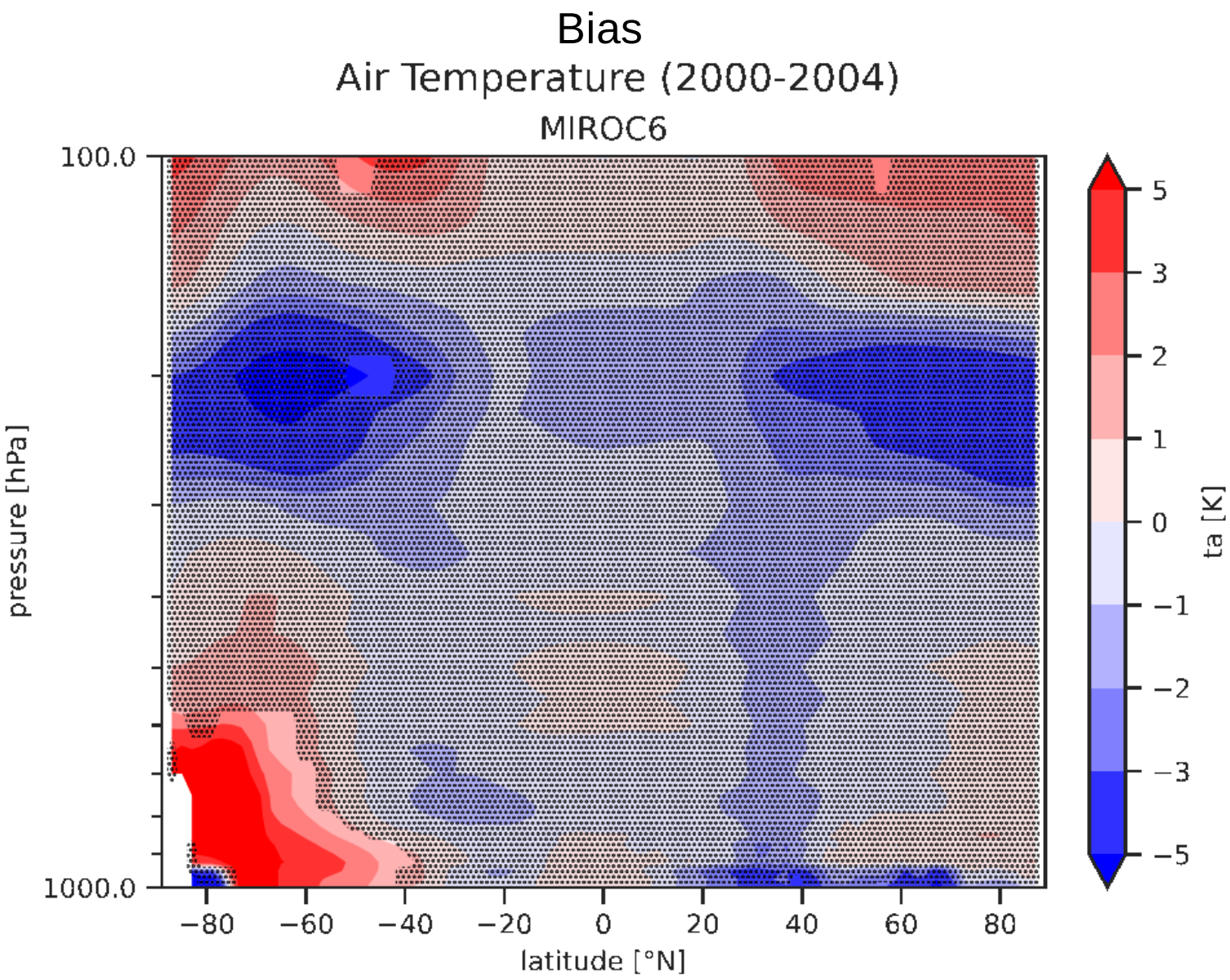


Extension of monitoring diagnostics



Stippled areas mask grid cells where the RMSE is smaller than the 90% percentile of RMSE values from an ensemble of CMIP6 models.

Extension of monitoring diagnostics

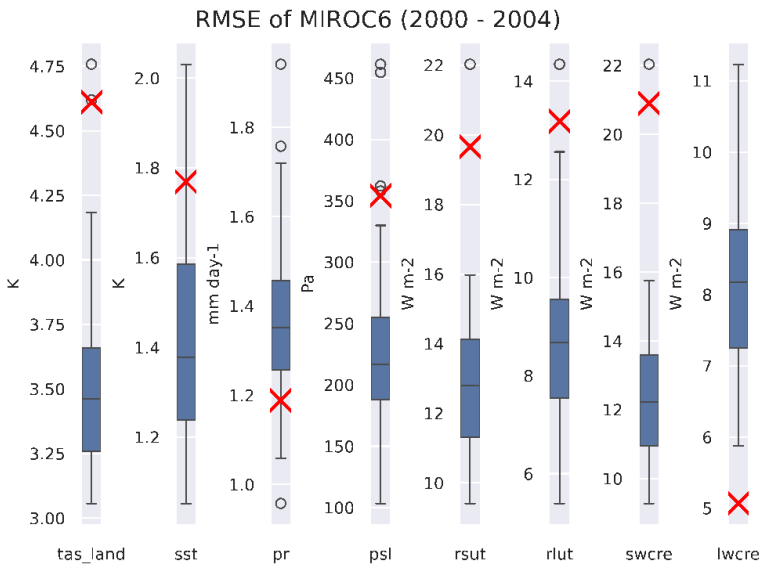


Stippled areas mask grid cells where the absolute BIAS (|BIAS|) is smaller than the maximum of the absolute 10% (|p10|) and the absolute 90% (|p90|) percentiles from an ensemble of CMIP6 models, i.e. **|BIAS|≤max(|p10|,|p90|)**.

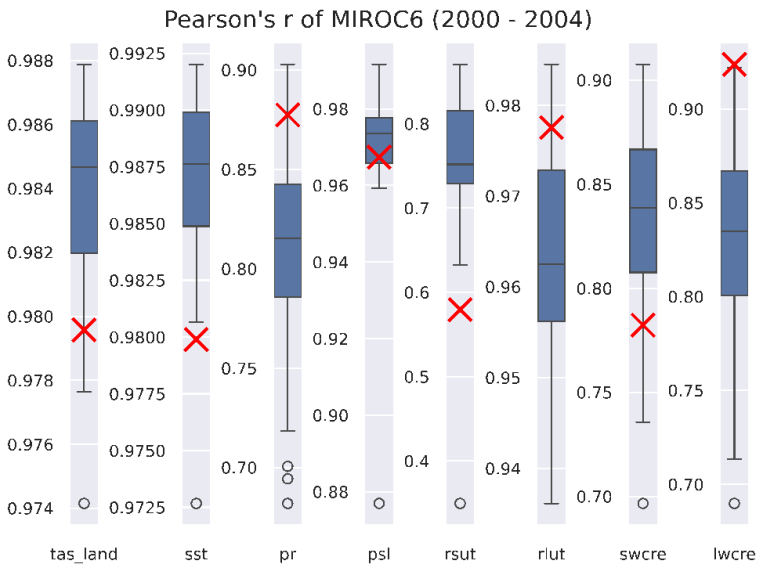
Extension of monitoring diagnostics

Comparison of the geographical pattern of 5-year means of different variables from a simulation of MIROC6 (red cross) in comparison to the CMIP6 ensemble (boxplot)

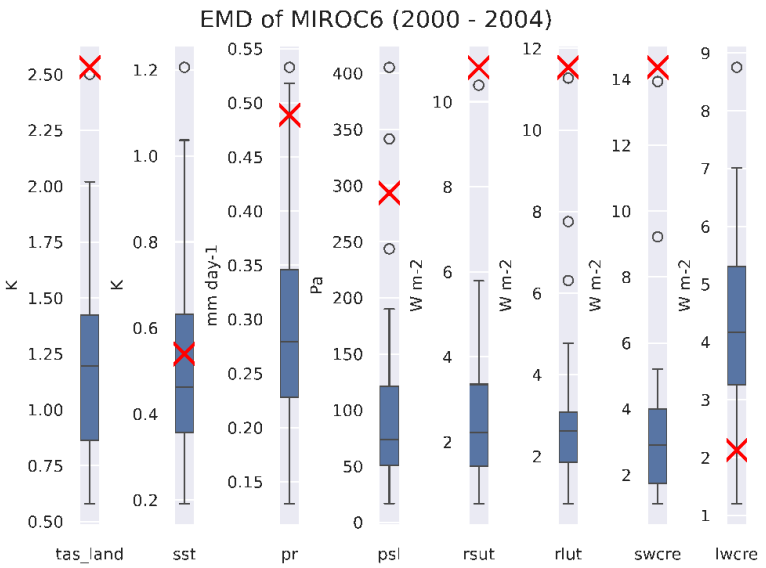
RMSE



Pearson‘s correlation coeff.



EMD



Pull request

Benchmarking recipes (Lauer et al.) #3598