How well do climate models simulate precipitation?

Nathalie Schaller, Irina Mahlstein, Jan Cermak and Reto Knutti
Institute for Atmospheric and Climate Science, ETH Zurich
Contact: nathalie.schaller@env.ethz.ch

1. Motivation

- Changes in the hydrological cycle are expected to have severe impacts on societies and ecosystems.
- Model disagreement for precipitation is large and unambiguous statements about future changes are difficult to provide.
- Model evaluation has mostly been performed on a wide range of climate variables and based on statistical measures of biases.
- However, statistical metrics do not correlate strongly with future projections (*Knutti et al., 2010)*.

Aim: define feature-based metrics considering regional changes of precipitation that can be understood physically. a) Rank the models by these metrics and compare with other evaluation methods. b) Investigate the time evolution of the feature-based metrics.

2. Data and Methods

- Observations: GPCP *(Adler et al., 2003)* and ERA-40 *(Uppala et al., 2005)*
- Models: WCRP CMIP3 multimodel ensemble *(Meehl et al., 2007)*

- Compare three different ways of ranking the CMIP3 models:

1. Biases for a range of climate variables and on a global scale *(Reichler and Kim, 2008)* → "RK08 ranking"
2. Biases in precipitation and temperature on a global scale with simple statistical measures → "rmse/corr ranking"
3. Ability of the models to simulate the mean of the regional feature-based metrics for the observational period → "indices ranking"

Table 1: Definition of the precipitation indices. Pr stands for the average precipitation over a season.

Index name	Definition	Domain
African index	$AFI = \overline{Pr}_{\mathrm{JJA}}$	13°-35°S, 14°-42°E
Amazonian index	$AMI = \overline{Pr}_{\mathrm{JJA}}$	24°S-1°N, 31°-59°W, land only
Asian index	$ASI = \overline{Pr}_{\mathrm{JJA}} - \overline{Pr}_{\mathrm{DJF}}$	10°-29°N , 70°-118°W
Australian index	$AUI = \overline{Pr}_{\mathrm{JJA}}$	10°-40°S, 107°-138°E
Central american index	$CAI = \overline{Pr}_{\mathrm{JJA}}$	10°-29°N, 110°-62°W
High latitudes index	$HLI = \overline{Pr}_{\mathrm{DJF}}$	52°-71°N, land only
Mediterranean index	$MEI = \overline{Pr}_{\mathrm{JJA}}$	29°-49°N, 11°W-37°E
Storm tracks index	$STI = \overline{Pr}_{\mathrm{DJFB}} - \overline{Pr}_{\mathrm{DJFA}}$	zone A (35°-46°S)/zone B (49°-60°S)

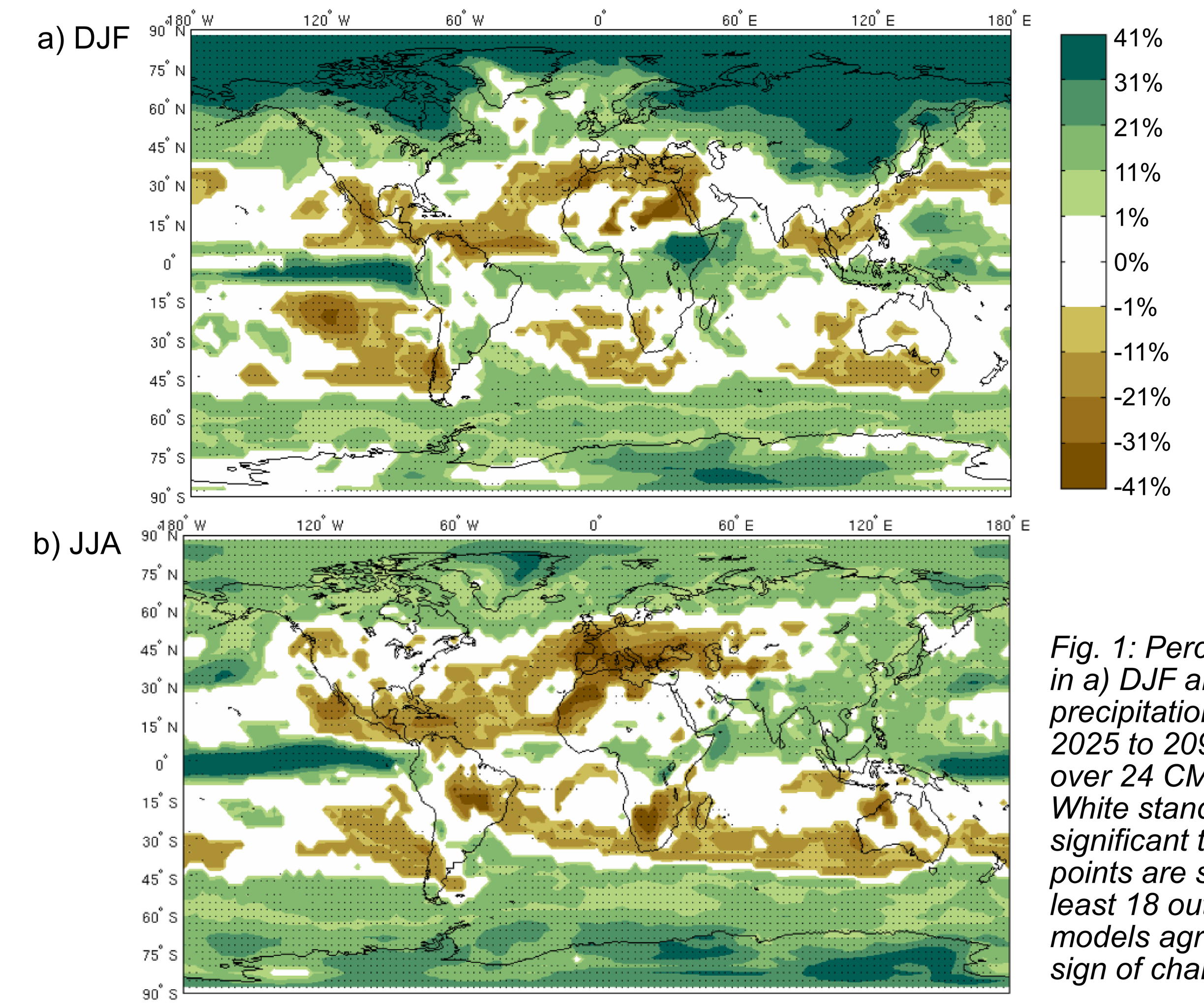

Fig. 1: Percent change in a) DJF and b) JJA precipitation from 2025 to 2099 averaged over 24 CMIP3 models. White stands for non-significant trends. Grid points are stippled if at least 18 out of 24 models agree on the sign of change.

3a. Results: model ranking

- No CMIP3 model appears to consistently outperform the rest.
- Indices ranking: each model can perform above and below average for different regions or variables.
  The multimodel mean performs surprisingly average. Its performance increases the more regions and variables are taken into account since it does not have to compensate for bad performances.
- Rmse/corr ranking: multimodel mean clearly ranks first.
- Comparison with the RK08 ranking: correlation of *R=0.62* with rmse/corr ranking and indices ranking. Including more variables will result in increasing the correlation among the rankings.
- Depending on the purpose, identifying and using the best models in a given region might provide more reliable information than taking the average of all models.

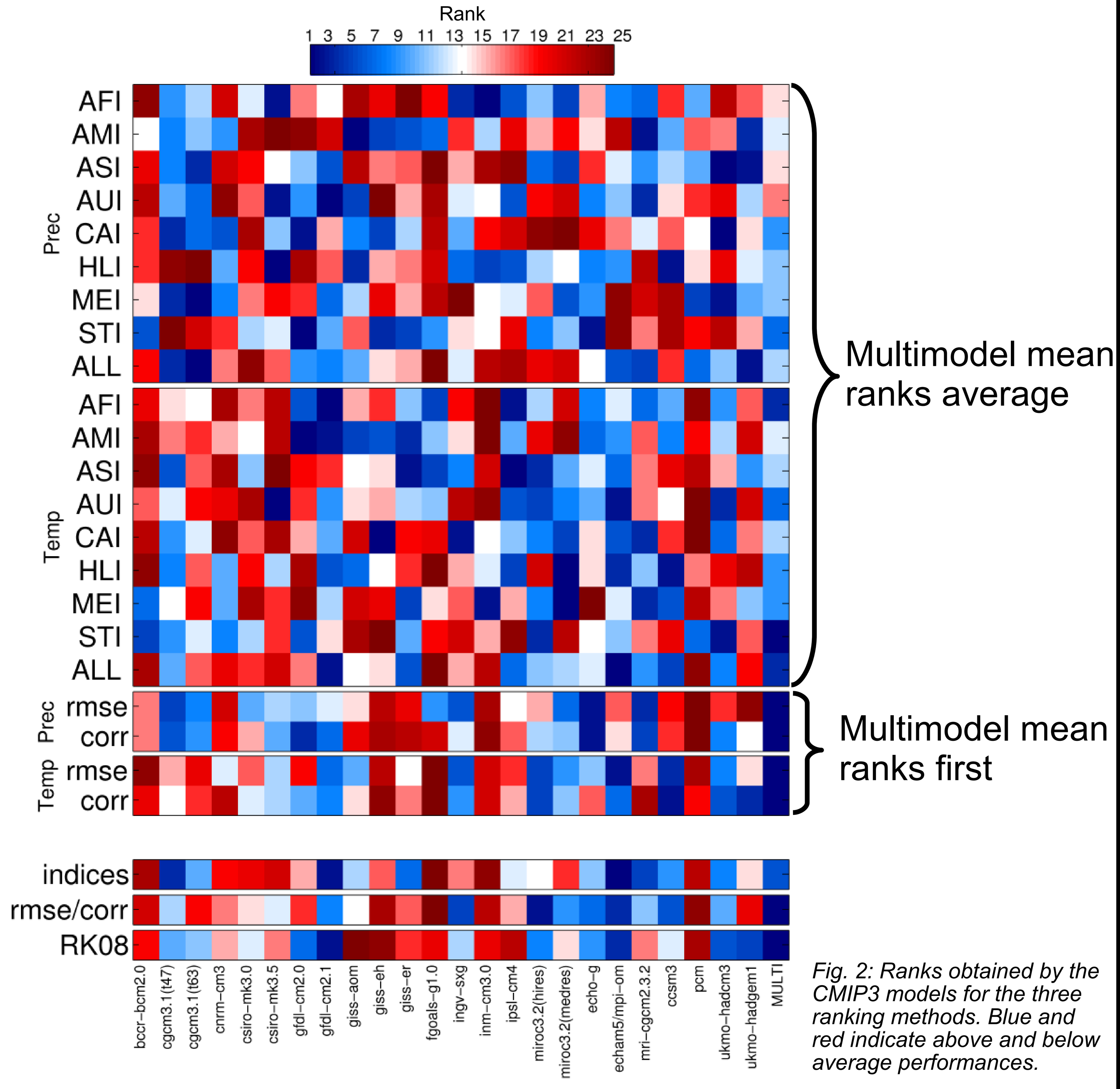

Fig. 2: Ranks obtained by the CMIP3 models for the three ranking methods. Blue and red indicate above and below average performances.

3b. Results: future projections

- The models have similar biases for precipitation. In most cases, the observations (green and orange lines on Fig. 3) are located at one end of the model range. The multimodel mean therefore never performs best, but also never performs poorly by design.
- Considering only the five best models for each index narrows the range of predicted absolute values.
- For future anomalies, model spread is not reduced, supporting the fact that means and trends are generally not well correlated.
- The choice of the reference dataset plays a role for feature-based metrics.

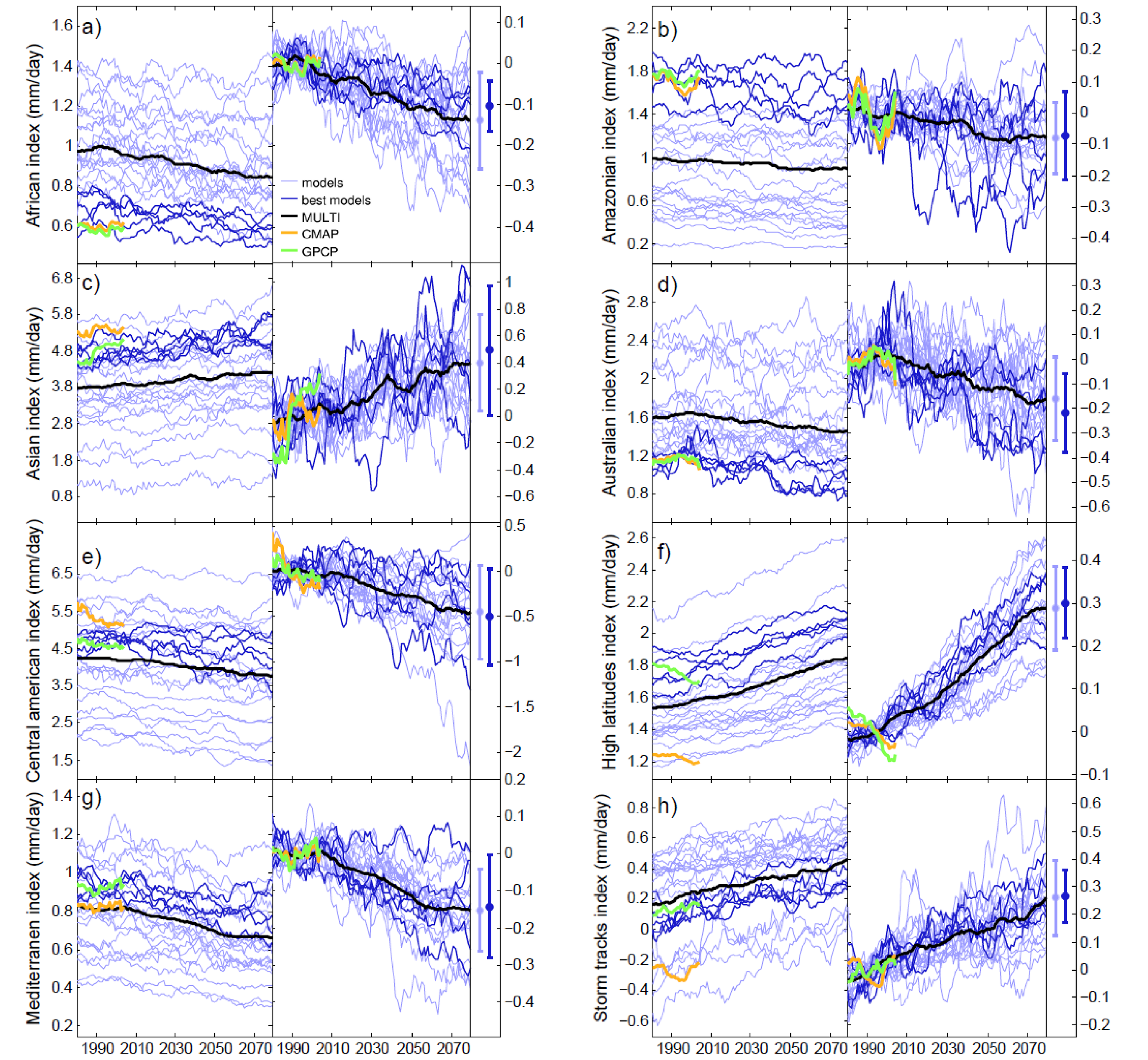

Fig. 3: Time series (left panels of each subfigure) and anomalies relative to the observational period (right panels of each subfigure). An 11-year average is applied to the time series. The mean value and two standard deviations in the year 2079 are shown in light blue (dark blue) for all CMIP3 models (five best models).

4. Conclusion

- Statistic-based evaluations on a global scale favour the multimodel mean only because per definition, it cannot perform poorly.
- Each individual model however can perform above and below average for different regions and variables. The best performances for a region of interest can be identified with feature-based metrics.
- Selecting the best models does not improve the future uncertainty.

References

Adler, R.F. et al., 2003: The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present), *JHyMe*, **4(6)**, 1147-1167
Knutti, R. et al., 2010: Challenges in combining projections from multiple climate model. *Journal of Climate*, in press
Meehl,G.A. et al., 2007: The WCRP CMIP3 multimodel dataset – a new era in climate change research. *Bull. Am. Met. Soc.* **88**, 1383-1394
Reichler,T. & J.Kim, 2008: How well do coupled models simulate today's climate? *BAMS*, **89(3)**, 303-311
Uppala, S.M. et al., 2005: The ERA-40 re-analysis. *Quarterly J. of the Royal Met. Soc.,* ***131***,2961-3012