Universität Trier
Physische Geographie
AG Modellbildung und Simulation

Effect of sample size on bias correction performance

Philipp Reiter[1,2], Oliver Gutjahr[3], Lukas Schefczyk[3], Günther Heinemann[3], Markus Casper[2]

RheinlandPfalz
MINISTERIUM FÜR WIRTSCHAFT, KLIMASCHUTZ, ENERGIE UND LANDESPLANUNG

Background

- a bias in climate model output necessitates a pre-processing before using it e.g. as climate forcing
- the common bias correction method of quantile matching (QM) uses transfer functions
- the performance of QM especially in case of precipitation is expected to strongly depend on the sample size used for the calibration of the transfer function
- in this study we investigate critical sample sizes

Domain & Data sets

Domain: Germany **Variable:** daily precipitation

Observational data set: E-OBS, 1961-2000, 25 km

Regional climate model (RCM) reanalysis runs:
EU-ENSEMBLES, ERA-40, 10 RCMs, 1961-2000, 25 km

Method

Bias correction

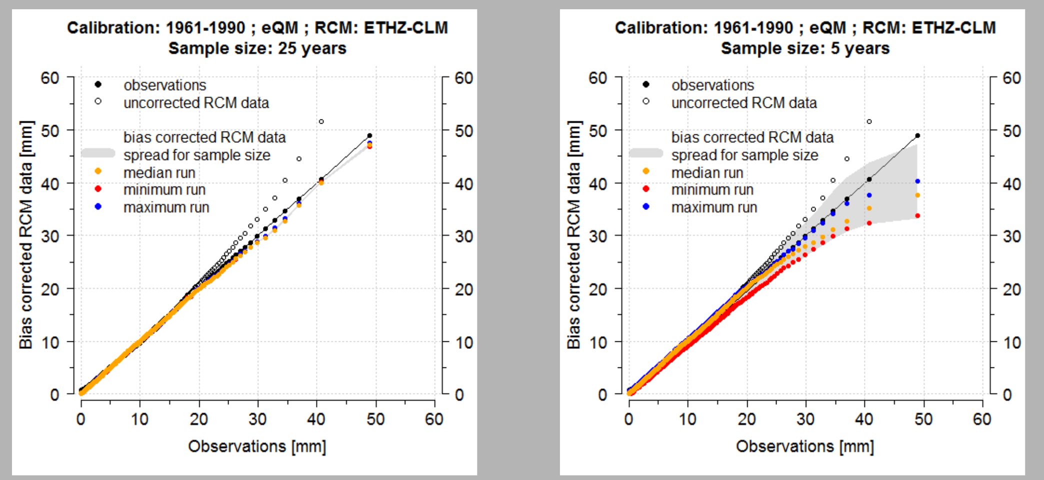
1. the 40-year period (1961-2000) is split into a 30-year calibration and a 10-year validation period
2. cell-by-cell a bias correction is done for the ‚best case' by using the complete calibration period for calibration of the transfer function
3. step 2 is redone for reduced sample sizes (29 years down to 1 year), using all possible combinations of consecutive years

Alterations:
the methodology is repeated for

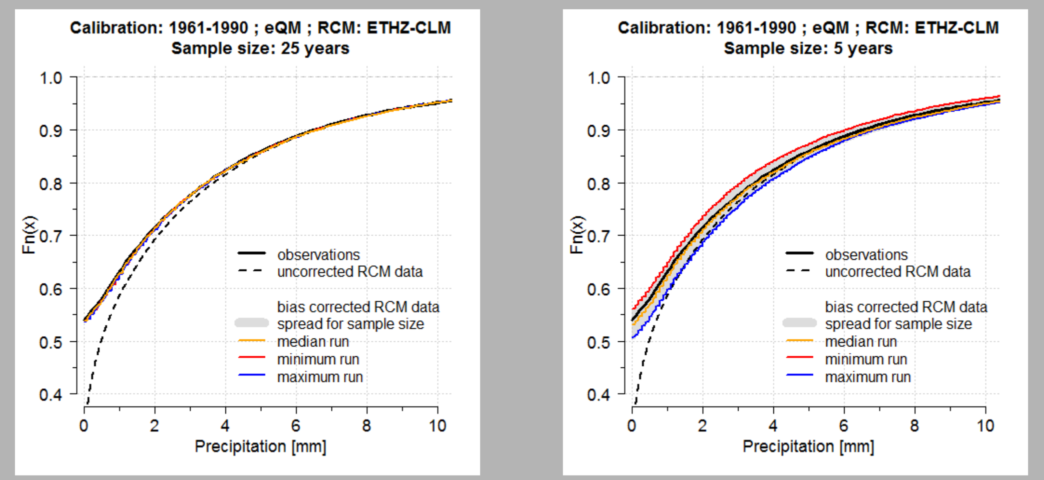
- all 10 RCMs
- 4 QM approaches (eQM[a], gQM[a], GQM[a], PTF[b])
- 3 different splittings of the 40-year period

Bias correction results

Q-Q-Plots



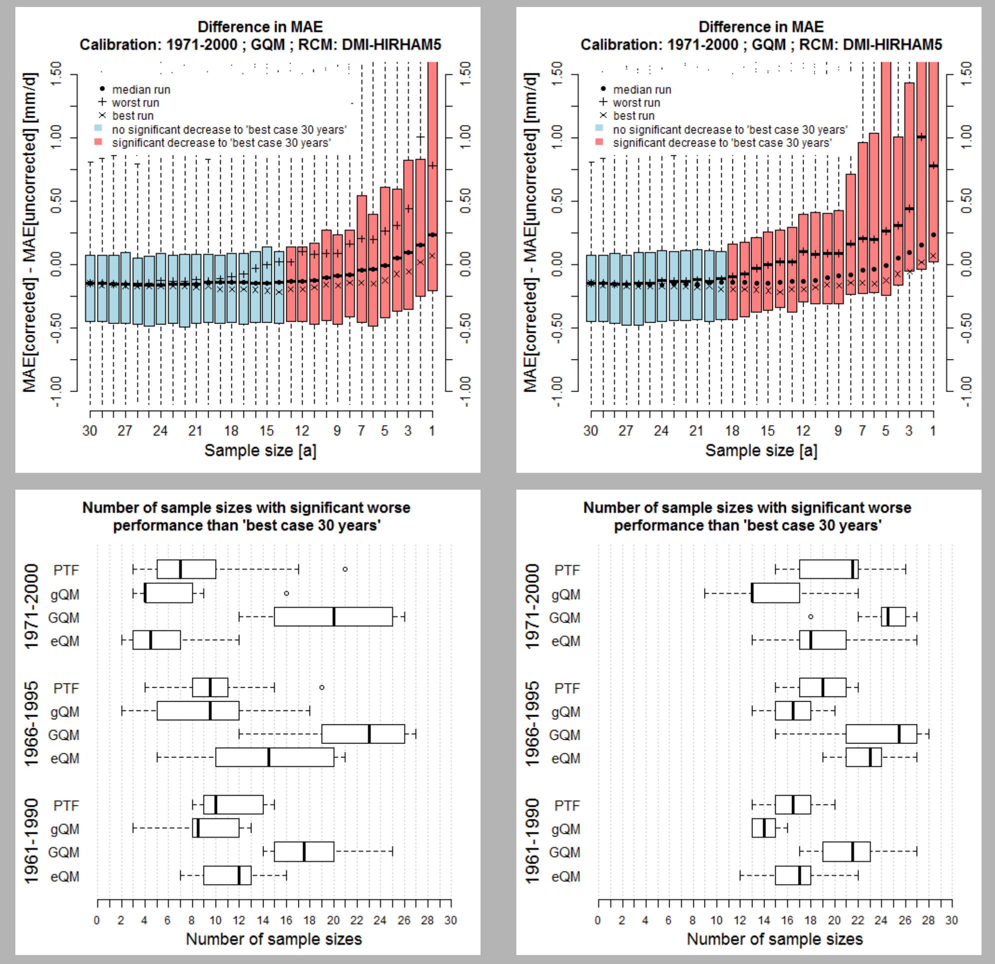
ecdf-Plots



MAE - mean absolute error

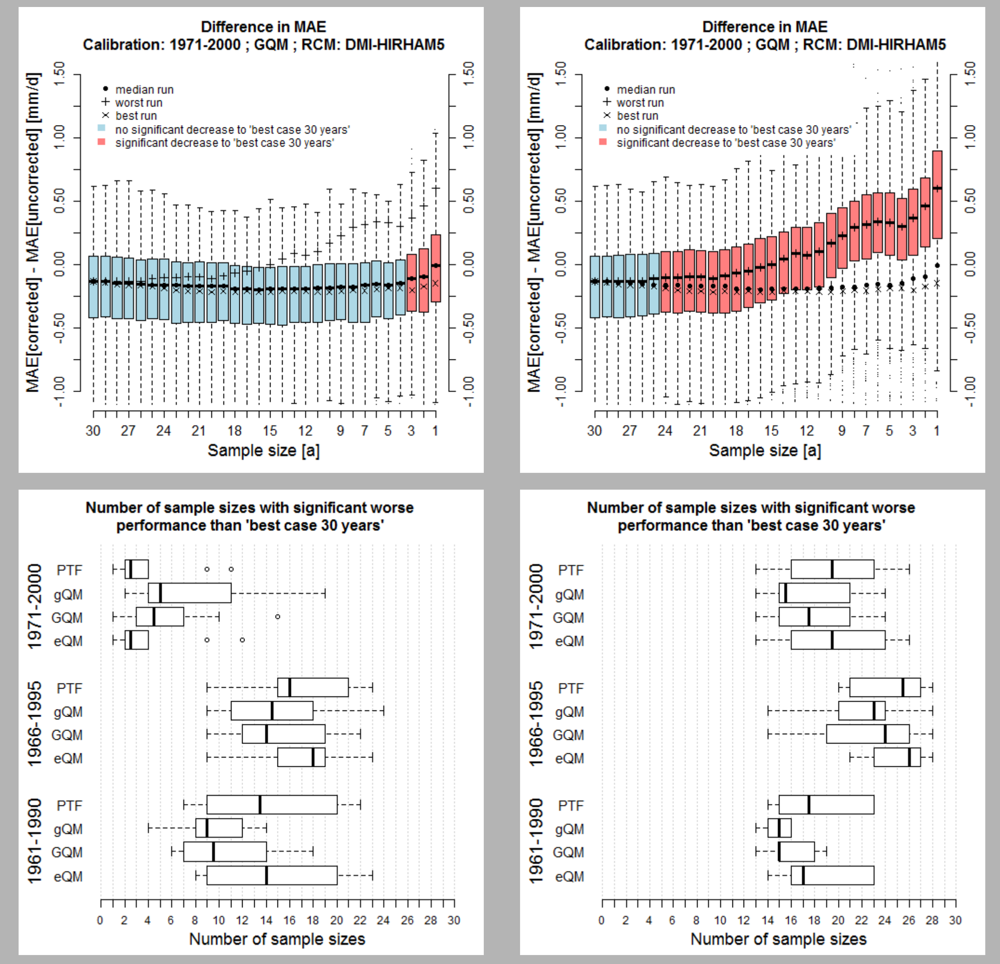
- for the validation period the mean absolute error[c] (MAE) in reference to the observations is calculated cell-by-cell for 10 quantiles (qstep = 0.1) of the cumulative distribution function (CDF)
- the difference between MAEs of the corrected and uncorrected data is calculated
- the distribution of the differences is statistically tested (Mann-Whitney U ; $\alpha = 0.05$) against the distribution of the ‚best case' of 30 years
- finally the test results are summarized for each quantile and their mean (MAEx)

MAEx



median run | worst run

MAE0.8



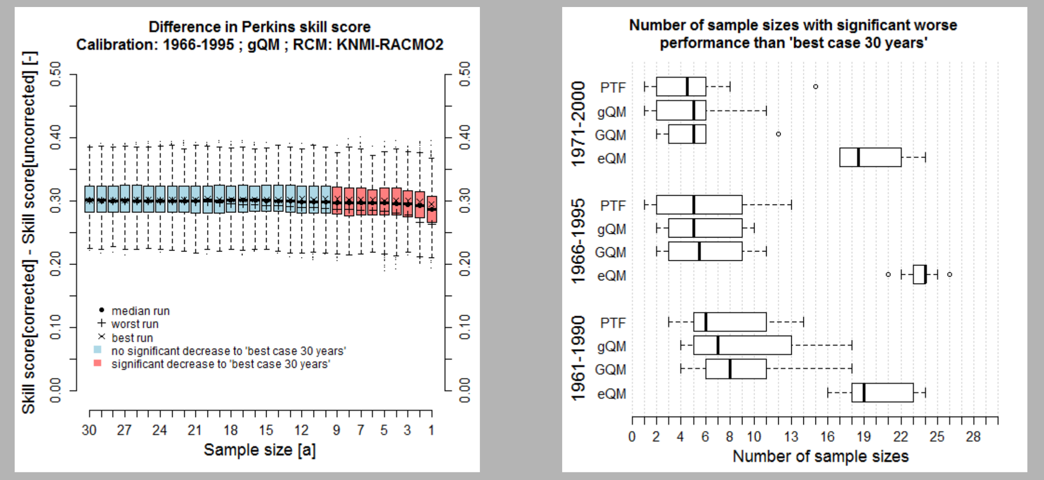
median run | worst run

Perkins skill score

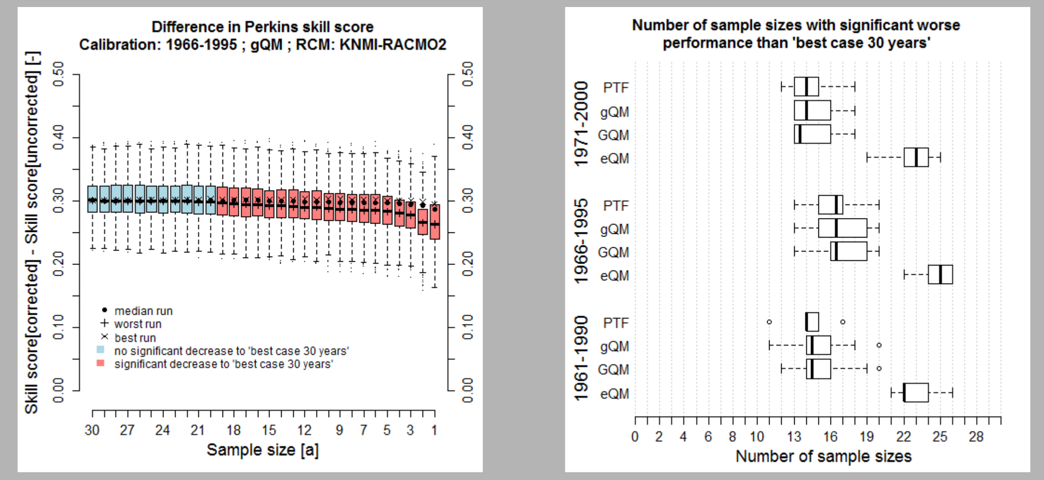
- the Perkins skill score[d] analyses the differences in the probability density function (PDF) of two time series t_1 and t_2

$$S_{score} = \Sigma \min(Z_1, Z_2) \quad \text{where } Z = \text{probability of values in the specific bin}$$

- the skill score is applied cell-by-cell to the bias corrected and uncorrected RCM data of the validation period in reference to the observational data
- as for MAE the differences in the skill score values of the corrected and uncorrected data are calculated and the distributions of the differences are statistically tested against that of the ‚best case' of 30 years and summarized



median run



worst run

First conclusions

- reduction of sample size leads to a decrease in bias correction performance
- the decrease in performance occurs much faster for the worst runs than for the median runs, but overall there is a large spread of the critical sample size
- depending on the scientific question and its related skill score, different ranges of critical sample size can be determined
- with decreasing sample size the correction of extreme values (and also of the lower quantiles of the PDF) becomes unstable
- to determine more accurate critical sample sizes for a combination of calibration period and QM approach, the results need to be combined with the absolute skill score values

[a] Gutjahr, O. & Heinemann, G. (2013): Comparing precipitation bias correction methods for high-resolution regional climate simulations using COSMO-CLM. doi: 10.1007/s00704-013-0834-z
[b] Piani, C. et al. (2010): Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. doi: 10.1016/j.jhydrol.2010.10.024 (Eq. (1.c))
[c] Gudmundsson, L. et al. (2012): Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations— a comparison of methods. doi: 10.5194/hess-16-3383-2012
[d] Perkins, S.E. et al. (2007): Evaluation of the AR4 Climate Models' Simulated Daily Maximum Temperature, Minimum Temperature, and Precipitation over Australia Using Probability Density Functions. doi: 10.1175/jcli4253.1

[1]Rhineland-Palatinate Competence Centre for Climate Change Impacts, Trippstadt, Germany | [2]University of Trier, Physical Geography, Trier, Germany | [3]University of Trier, Environmental Meteorology, Trier, Germany | Email: philipp.reiter@klimawandel-rlp.de

Poster: