Supplement of Geosci. Model Dev., 19, 5805–5825, 2026
https://doi.org/10.5194/gmd-19-5805-2026-supplement

Geoscientific Model Development
Open Access
EGU

Supplement of

EXSoDOS 1.0: downscaling of weather extremes shifts for ensemble climate projections using ground-based measurements, reanalysis and stochastic modelling

Hendrik Wouters et al.

Correspondence to: Hendrik Wouters (hendrik.wouters@vito.be)

S1. Correlation stability and stationarity of the statistical relationship under climate change and its effect on model results

In the manuscript, correlations and statistical relationships are assumed to be static under climate change. We test correlation stability and static statistical relationships across historical periods by calibrating the model for different periods in case of daily precipitation for Sikasso. The first one is the full 63-year period (1961-2023) as already done, and then two additional model calibrations for the non-overlapping 30-year periods 1961-–1990 and 1991–2020. These 3 models are calibrated on odd years and then applied to generate time series for even years ('even models'). Conversely, 3 additional models are calibrated on even years to generate odd years ('uneven models').

In Tab. S2, predictor-–predictand correlations for the lowest precipitation category are reported for the different calibration sets. Correlations appear stable among the different calibrations for the different months, hence suggests invariance under climate change. However, these changes may still lead to different results in the extremes. In addition, differences also result from differences in cumulative distribution functions of the predictand used for the calibration. To test the sensitivity to the changes in the correlations and predictand CDFs, we generated time series for each of these model sets. The results in distribution and return levels can be found in the respective figures below.

We find that variations in distributions/return levels of model output (even/odd; different calibration sets) are in accordance to variations in the distributions/return levels of observation samples (even vs. odd; 1991-2020 vs 1961-1990). Differences are most pronounced in the tails—and the different calibration sets lead to a variability of comparable magnitude. These results show that the statistical relationships are robust under climate change. Nevertheless, such sensitivity to calibration sets needs to be taken into account in climate model assessments.

S2: comparison between quantile maping to correlated sampling; illustration of inflation

EXSoDOS employs correlated sampling instead of quantile mapping. One of the key advantages to use correlated sampling compared to quantile mapping is that the latter give rise to unrealistic inflation of trends in extremes (see Maraun, 2013) whereas it is avoided by the correlated sampling. We illustrate this with additional results for quantile mapping as shown in Fig. S4 and compare them with the results with correlated sampling shown in Fig. S3 (as discussed earlier in Text S1).

Using Quantile mapping (Fig. S4), we find an increase in return levels in the tails (return period >= 1 year) of the predictor (ERA5 even and uneven) between the different time frames 1961-1990 and 1991-2020 (black lines). The increase in predictor return levels leads to an increase in modelled predictand return levels when using the quantile mapping approach (green and orange lines and spreads). However, return levels in the tails (return period >= 1 year) remain stable according to local observations for even years, and a decrease was found for uneven years. This discrepancy

suggests inflation of outcomes by the quantile mapping approach, and such inflation does not occur when applying correlated sampling (Fig. S3).

Quantile mapping versus correlated sampling also affects the results on the climate projections. To demonstrate this, we evaluate the effect of standard quantile mapping versus correlated downscaling on future climate projections. See Fig. S5 and table S3 below. It is found that climate change signals towards the end of the century are more pronounced with the quantile-mapping approach, especially in the tails, which may be due to inflation of trends. Hence, the method of statistical downscaling largely affects the outcomes, and needs to be accounted for in climate change assessments.

Assumption	Description / rationale	Implications / how we check it
A1. Perfect-prognosis (PP) assumption	Predictors (after bias-adjustment) are assumed to be physically meaningful and transferable between reanalysis and GCMs (Maraun, 2016).	We mitigate PP violations by (i) bias-adjusting predictors with quantile-delta mapping (Cannon et al., 2015) per month and (ii) validating the full predictor→predictand chain against station observations.
A2. Stationary predictor–predictand dependence	The month- and category-dependent correlation ρ between normalized predictor and predictand is assumed approximately invariant under climate change (Sect. 2.2.3).	We test stability by calibrating on independent periods (1961–1990 vs 1991–2020); correlations and resulting tails remain within sampling uncertainty.
A3. Stochastic residual variability	Unexplained sub-grid variability is treated as a stochastic residual r drawn from a standard normal and mapped back through the empirical CDF of observations (Eq. 10–11).	This implies that tail behavior beyond the observational record cannot be guaranteed; we therefore restrict interpretation to return periods supported by record length and report ensemble/sampling uncertainty.
A4. Distributional representativeness of station record	Station observations are assumed long and complete enough to represent climatological distributions and extremes over 30-year windows (WMO, 2017).	We specify minimum record-length and completeness criteria, and we limit validation plots to return periods supported by sample size (Sect. 2.4).
A5. QDM bias-adjustment stationarity	Quantile-delta mapping assumes percentile-dependent model bias is approximately stationary while climate change signals in quantiles are preserved (Cannon et al., 2015).	We discuss known limitations (e.g., trend inflation when mixing scales) and keep bias adjustment on the model grid before station downscaling (Maraun, 2013).
A6. Single-site / single-variable application	EXSoDOS is applied independently per station and variable and therefore does not enforce spatial or cross-variable coherence.	We explicitly state this limitation and provide guidance on when multi-site or multivariate methods (e.g., copulas / correlation-preserving transforms; Switanek et al., 2022) are required (Sect. 4).

Table S1: Key assumptions of EXSoDOS stochastic downscaling

	may	jun	jul	aug	sep	oct
1961-1990 even model	0.28	0.2	0.2	0.2	0.32	0.39
1961-1990 uneven model	0.29	0.22	0.2	0.23	0.39	0.4
1991-2020 even model	0.28	0.21	0.21	0.17	0.33	0.45
1991-2020 uneven model	0.32	0.25	0.17	0.17	0.32	0.44
1961-2023 even model	0.31	0.17	0.2	0.18	0.33	0.43
1961-2023 uneven years (reference)	0.3	0.22	0.21	0.2	0.34	0.42

Table S2: predictor-–predictand correlations for the lowest precipitation category in different months in the rainy months of Sikasso.

Dataset	Window	Average	Std	P95	1y return
observations	1961–1990	3.20	8.99	20.20	67.60
observations	1991–2020	3.40	9.62	22.49	69.70
CMIP6	1961–1990	2.18 (1.78–3.06)	5.14 (4.43–8.61)	9.94 (6.86–13.67)	42.39 (28.90–82.14)
CMIP6	1991–2020	2.35 (1.85–3.48)	5.72 (4.57–9.85)	10.86 (7.21–15.53)	47.24 (31.40–93.95)
CMIP6	2071–2100	2.52 (2.46–3.88)	7.89 (5.90–12.65)	11.25 (9.31–18.21)	69.75 (46.81–114.54)
CMIP6 BC DS	1961–1990	3.20 (3.12–3.30)	9.24 (8.82–9.30)	21.10 (20.56–22.09)	64.36 (61.90–67.14)
CMIP6 BC DS	1991–2020	3.40 (3.36–3.50)	9.72 (9.54–10.30)	22.54 (22.04–23.16)	69.43 (64.91–73.14)
CMIP6 BC DS	2071–2100	3.92 (3.28–4.26)	11.68 (10.04–12.87)	25.67 (21.93–28.59)	82.20 (70.97–90.45)
1961–1990	CMIP6 BC QM	2.88 (2.82–3.11)	8.19 (7.95–8.54)	18.90 (18.18–20.53)	58.56 (56.41–61.16)
1991–2020	CMIP6 BC QM	3.17 (3.08–3.25)	9.02 (8.51–9.94)	20.68 (19.95–20.90)	66.73 (60.75–75.50)
2071–2100	CMIP6 BC QM	3.69 (3.11–4.09)	12.60 (9.64–13.41)	22.52 (19.90–26.95)	96.59 (75.31–102.14)

Table S3: Idem as Tab. 3 in main text (only daily precipitation for Sikasso), but appending results for quantile mapping (CMIP6 BC QM) to the results with correlated sampling (CMIP6 BC DS).

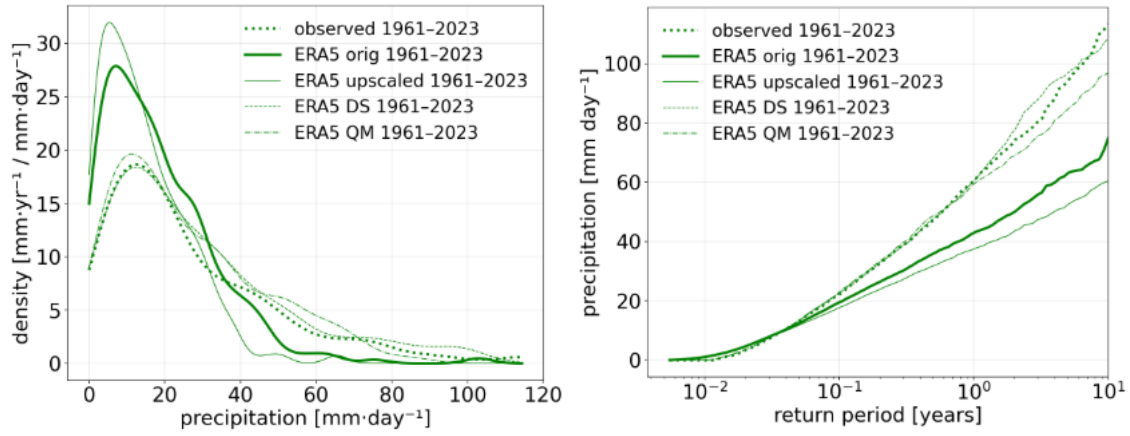

Figure S1. Idem as Fig. 4 and 5, but for combined results of 7 stations in USA, namely USC00141593 (lat=39.5722, lon=-97.2836), USC00445050 (lat=38.0422, lon=-78.0061), USC00021664 (lat=32.0061, lon=-109.357, USC00130157 (lat=42.7536, lon=-92.8022), USC00250640 (lat=40.1306, lon=-99.8278), USC00410404 (32.1633, -95.83), USC00475808 (44.5378, -90.535).

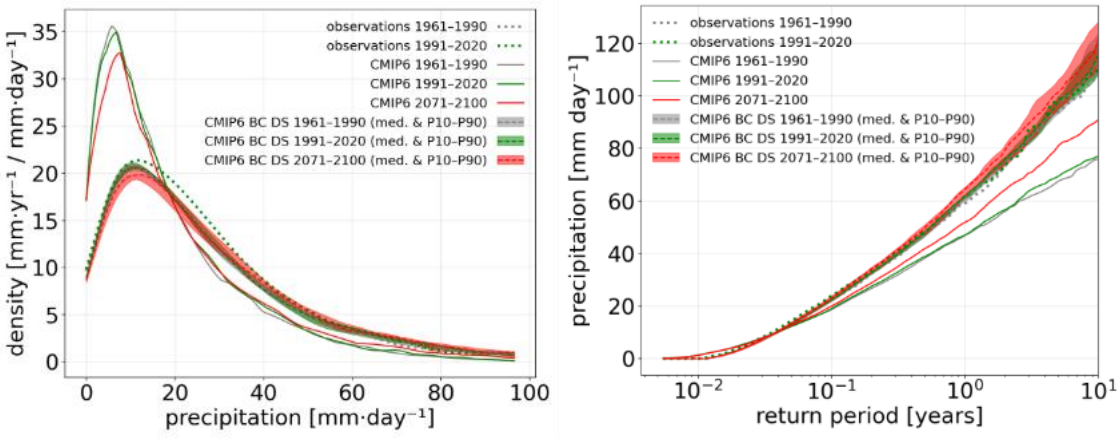

Figure S2. Idem as Fig. 6 and 7, but for combined results of 8 stations in USA.

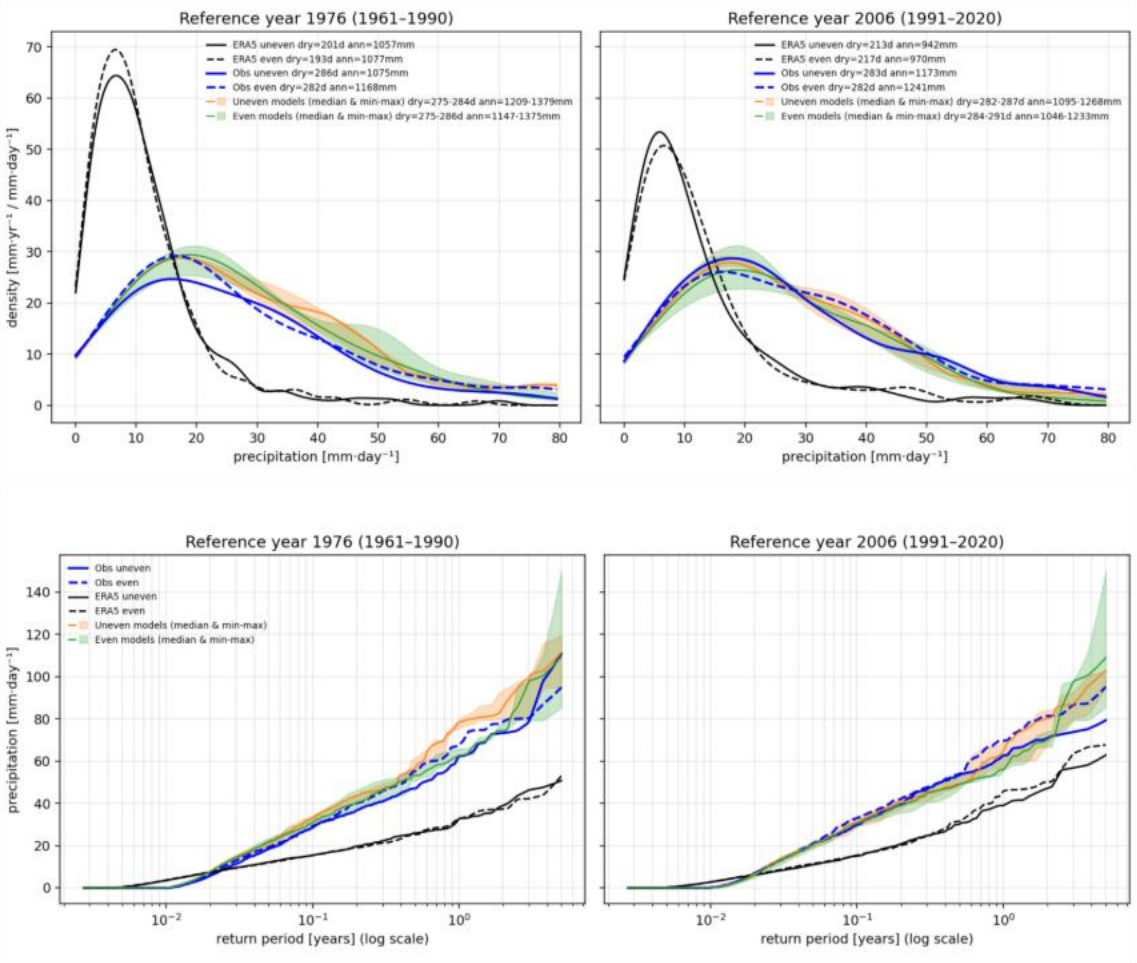

Figure S3: distributions (upper panels) and return levels (lower panels) for observations (Obs), ERA5 (ERA5), and downscaled results from models with the different calibration sets (models). We show results for even (even) and odd (uneven) years. Left panels show results for 1961-1990 and right panels show results for 1991-2020.

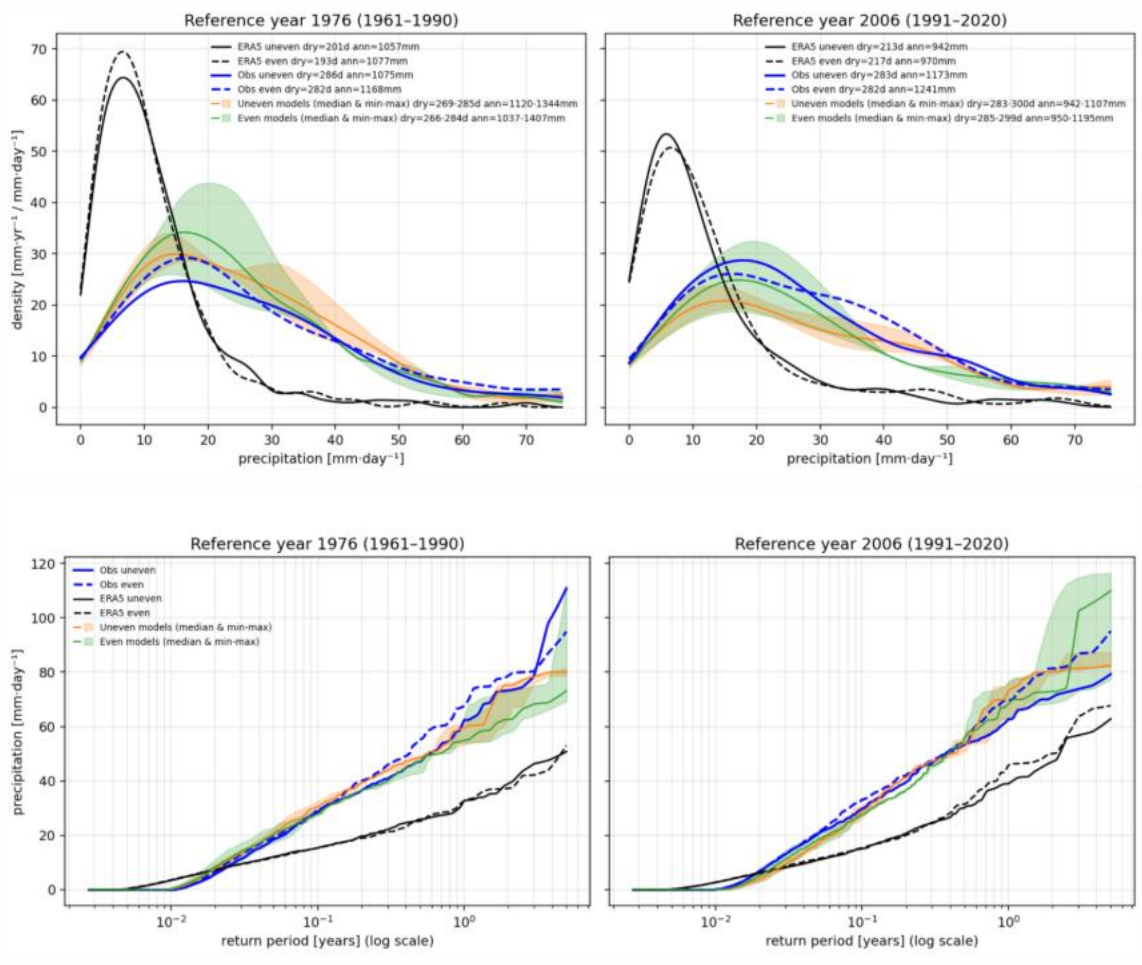

Figure S4: idem as Fig. S3, but with quantile mapping instead of correlated sampling.

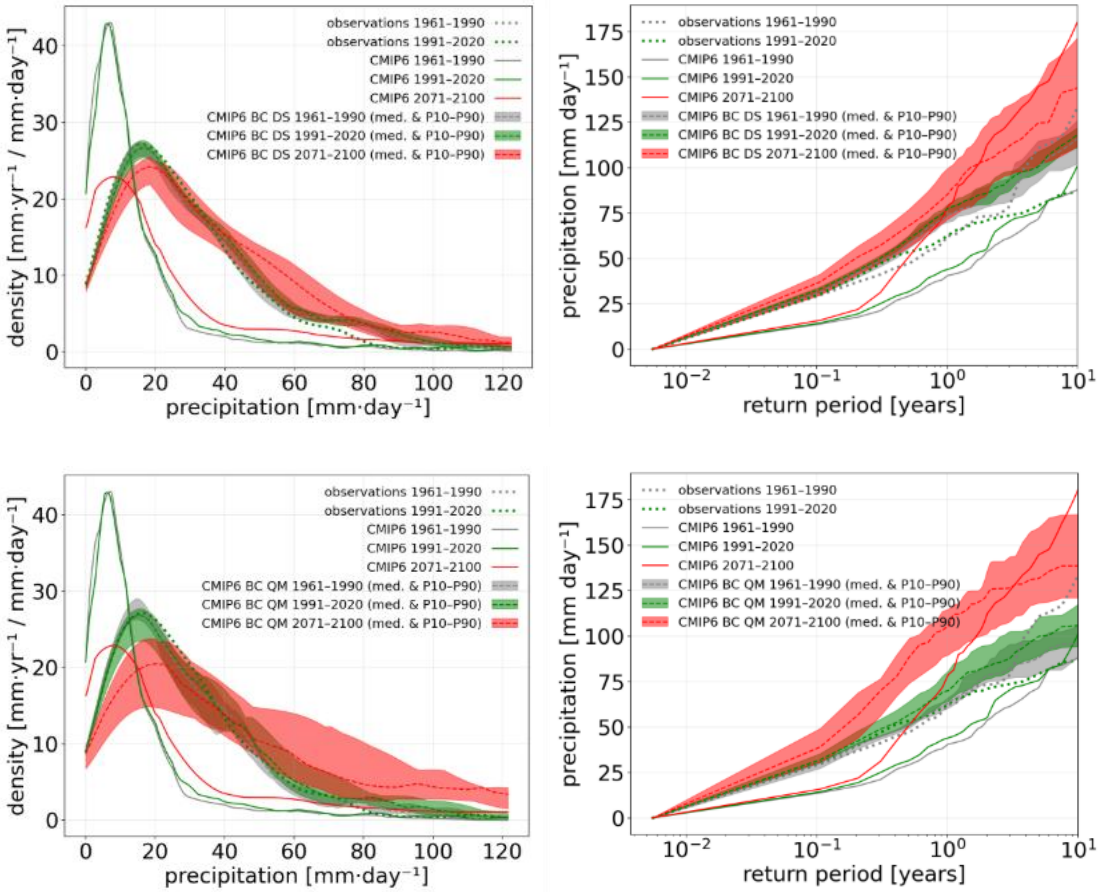

Figure S5: upper panels are the lower right panels of resp Fig. 6 and 7 (precipitation for Sikasso). The lower panels are the same but using quantile mapping for downscaling towards station level (QM).