Supplement of

Variability and extremes: statistical validation of the Alfred Wegener Institute Earth System Model (AWI-ESM)

Justus Contzen et al.

Correspondence to: Justus Contzen (justus.contzen@awi.de)

The copyright of individual parts of the supplement might differ from the article licence.
S1 Evaluation of the EC-Earth3-Veg-LR model

In this section, we present additional figures evaluating the EC-Earth3-Veg-LR model (EC-Earth Consortium, 2020). This model yielded the lowest value for the AWQD out of all 27 CMIP6 models investigated (see Section S2 for a complete list). In analysing the data, we followed the procedure described in the methods section of the main paper. In Fig. S1, empirical means and standard deviations of the EC-Earth3-Veg-LR model and the reanalysis data as well as their anomalies are shown. It can be observed that model and reanalysis data generally have similar mean values. In South America, at the western coast and parts of the eastern coasts, an overestimation of the mean is visible, while mean values are underestimated in the northern part of South America and in Central America. In southern Africa, a slight overestimation of the mean values occurs. In India and the Himalayas, a complex pattern of overestimation and underestimation depending on the location emerges. Similar conclusions can be made for the empirical standard deviation, with a misestimation of it in most cases corresponding also to a misestimation of the empirical mean. A notable exception to this is Australia, where empirical standard deviations are often overestimated, while the mean values are not. QQ-Plots of the empirical statistics (Fig. S2) show that the mean values of climate model and reanalysis data are relatively similarly distributed, while for the empirical standard deviations, a tendency for underestimation especially for higher values can be observed. For both statistics, the highest values that are present in the reanalysis data do not correspond to equally high values in the climate model data. A comparison of the anomalies of empirical mean and empirical standard deviation (Fig. S2 c) shows the presence of heteroscedasticity, similar to what was already observed when analysing the AWI-ESM. In S3, the results of the KS-Tests on GEV distribution are depicted, and the hypothesis of the data being GEV distributed is not rejected for nearly all grid points, with the exception being large parts of the Sahara desert.
S2 Evaluation of CMIP6 models

We have analysed a total of 27 different models participating in CMIP6 (Eyring et al., 2016). The models have been compared with the CRU TS4.04 reanalysis precipitation dataset (University of East Anglia Climatic Research Unit et al., 2020) using the methodology described in the main article. The time span of the analysis is always from 1930 to 2005. The spatial resolution of the models varies, and for each climate model data set, the reanalysis data are interpolated onto the grid of the climate model. In the following table, a list of all models considered can be found and for each model the reference to the data set, the model run used, the spatial resolution of the atmospheric component and the calculated AWQD are documented. We then present for each model key figures of our analysis: The anomaly of the estimated GEV parameters and the 95% GEV distribution quantiles (with shaded areas indicating that the model parameter/quantile lies within the range of the reanalysis parameter/quantile and vice versa), and the clusterings based on dissimilarity measure $D_0$ for both model and reanalysis data. The figures not presented here can be reproduced using the R code that is supplemented to the paper (all data used are publicly available; references to
Figure S2. QQ-Plots comparing the empirical mean values (a) and the empirical standard deviations (b) and of the annually maximised monthly precipitation of the CRU TS4.04 reanalysis data set and of the EC-Earth3-Veg-LR model data set. Deviance of empirical mean and standard deviation plotted against each other (c). Units are mm/month.
Figure S3. P-values of Kolmogorov-Smirnov tests for the hypothesis that the data follow a GEV distribution with parameters estimated using probability-weighted moments. Test results for the EC-Earth3-Veg-LR climate model (a) and for the CRU TS4.04 reanalysis data (b).

Table S1. Analysed CMIP6 models and their spatial resolutions and AWQDs.

| Model name           | Reference to data set | Model run | Resolution | AWQD |
|----------------------|-----------------------|-----------|------------|------|
| ACCESS-CM2           | Dix et al. (2019)     | r5i1p1f1  | 192 × 144  | 76.73|
| ACCESS-ESM1.5        | Ziehn et al. (2019)   | r9i1p1f1  | 192 × 143  | 85.86|
| AWI-CM1.1MR          | Semmler et al. (2018) | r1i1p1f1  | 384 × 192  | 54.70|
| AWI-ESM1.1LR         | Danek et al. (2020)   | r1i1p1f1  | 192 × 96   | 52.98|
| BCC-CSM2-MR          | Wu et al. (2018)      | r1i1p1f1  | 320 × 160  | 83.00|
| CAMS-CSM1.0          | Rong (2019)           | r2i1p1f1  | 320 × 160  | 54.35|
| CanESM5              | Swart et al. (2019)   | r7i1p2f1  | 128 × 64   | 88.72|
| CESM2                | Danabasoglu (2019a)   | r1i1p1f1  | 288 × 90   | 63.27|
| CESM2-FV2            | Danabasoglu (2019b)   | r3i1p1f1  | 144 × 96   | 62.99|
| CESM2-WACCM-FV2      | Danabasoglu (2019c)   | r2i1p1f1  | 144 × 96   | 61.51|
| CNRM-CM6-1           | Voldoire (2018)       | r29i1p1f2 | 256 × 128  | 70.57|
| EC-Earth3-Veg-LR     | EC-Earth Consortium (2020) | r1i1p1f1 | 320 × 160  | 44.71|
| FGOALS-f3-L          | Yu (2019)             | r1i1p1f1  | 288 × 180  | 60.86|
| FGOALS-g3            | Li (2019)             | r6i1p1f1  | 180 × 80   | 83.08|
| GISS-E2.1G           | NASA/GISS (2018)      | r9i1p1f1  | 144 × 90   | 69.92|
| HadGEM3-GC31-LL      | Ridley et al. (2019)  | r3i1p1f3  | 192 × 144  | 75.83|
| INM-CM4.8            | Volodin et al. (2019b)| r1i1p1f1  | 180 × 120  | 67.25|
| INM-CM5.0            | Volodin et al. (2019a)| r6i1p1f1  | 180 × 120  | 74.32|

Continued on next page
| Model name        | Reference to data set | Model run | Resolution | AWQD |
|-------------------|-----------------------|-----------|------------|------|
| IPSL-CM6A-LR      | Boucher et al. (2018) | r26i1p1f1 | 144 × 143  | 78.26|
| MIROC-ES2L        | Hajima et al. (2019)  | r1i1p1f2  | 128 × 64   | 59.95|
| MPI-ESM1.2-HR     | Jungclaus et al. (2019)| r10i1p1f1 | 384 × 192  | 54.16|
| MPI-ESM1.2-HAM    | Neubauer et al. (2019)| r3i1p1f1  | 192 × 96   | 59.87|
| MRI-ESM2.0        | Yukimoto et al. (2019)| r10i1p1f1 | 320 × 160  | 63.92|
| NEM3              | Cao and Wang (2019)   | r1i1p1f1  | 192 × 96   | 62.12|
| NorESM2-LM        | Seland et al. (2019)  | r2i1p1f1  | 144 × 96   | 55.60|
| NorESM2-MM        | Bentsen et al. (2019) | r3i1p1f1  | 288 × 192  | 52.43|
| TaiESM1.0         | Lee and Liang (2020)  | r2i1p1f1  | 288 × 192  | 58.99|
Figure S4. Model ACCESS-CM2: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S5. Model ACCESS-CM2: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 76.73. The numbers of clusters are 267 (model data) and 177 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S6. Model ACCESS-ESM1.5: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S7. Model ACCESS-ESM1.5: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 85.86. The numbers of clusters are 235 (model data) and 192 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S8. Model AWI-CM1.1MR: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S9. Model AWI-CM1.1MR: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 54.70. The numbers of clusters are 252 (model data) and 191 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S10. Model AWI-ESM1.1LR: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S11. Model AWI-ESM1.1LR: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 52.98. The numbers of clusters are 188 (model data) and 177 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
**Figure S12.** Model BCC-CSM2-MR: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S13. Model BCC-CSM2-MR: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is $83.00$. The numbers of clusters are 319 (model data) and 186 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S14. Model CAMS-CSM1.0: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S15. Model CAMS-CSM1.0: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 54.35. The numbers of clusters are 223 (model data) and 186 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S16. Model CanESM5: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S17. Model CanESM5: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 88.72. The numbers of clusters are 196 (model data) and 162 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S18. Model CESM2: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S19. Model CESM2: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 63.27. The numbers of clusters are 201 (model data) and 185 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S20. Model CESM2-FV2: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S21. Model CESM2-FV2: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is $62.99$. The numbers of clusters are 131 (model data) and 177 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S22. Model CESM2-VACCMM-FV2: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S23. Model CESM2-VACCM-FV2: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 61.51. The numbers of clusters are 130 (model data) and 177 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S24. Model CNRM-CM6-1: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S25. Model CNRM-CM6-1: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is $70.57$. The numbers of clusters are 258 (model data) and 182 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S26. Model EC-Earth3-Veg-LR: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S27. Model EC-Earth3-Veg-LR: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 44.71. The numbers of clusters are 224 (model data) and 186 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S28. Model FGOALS-f3-L: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S29. Model FGOALS-f3-L: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is $60.86$. The numbers of clusters are 224 (model data) and 183 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S30. Model FGOALS-g3: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S31. Model FGOALS-g3: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 83.08. The numbers of clusters are 262 (model data) and 170 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S32. Model GISS-E2.1G: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S33. Model GISS-E2.1G: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 69.92. The numbers of clusters are 182 (model data) and 181 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S34. Model HadGEM3-GC31-LL: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S35. Model HadGEM3-GC31-LL: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 75.83. The numbers of clusters are 304 (model data) and 177 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
INM-CM4.8

Figure S36. Model INM-CM4.8: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S37. Model INM-CM4.8: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 67.25. The numbers of clusters are 241 (model data) and 189 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S38. Model INM-CM5.0: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S39. Model INM-CM5.0: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 74.32. The numbers of clusters are 229 (model data) and 189 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S40. Model IPSL-CM6A-LR: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S41. Model IPSL-CM6A-LR: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 78.26. The numbers of clusters are 324 (model data) and 177 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S42. Model MIROC-ES2L: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S43. Model MIROC-ES2L: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 59.95. The numbers of clusters are 130 (model data) and 162 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S44. Model MPI-ESM1.2-HR: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S45. Model MPI-ESM1.2-HR: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 54.16. The numbers of clusters are 248 (model data) and 192 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
**Figure S46. Model MPI-ESM1.2-HAM:** Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S47. Model MPI-ESM1.2-HAM: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 59.87. The numbers of clusters are 194 (model data) and 177 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S48. Model MRI-ESM2.0: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S49. Model MRI-ESM2.0: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 63.92. The numbers of clusters are 230 (model data) and 186 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S50. Model NESM3: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S51. Model NESM3: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 62.12. The numbers of clusters are 154 (model data) and 177 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S52. Model NorESM2-LM: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S53. Model NorESM2-LM: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is $55.60$. The numbers of clusters are 134 (model data) and 177 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S54. Model NorESM2-MM: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S55. Model NorESM2-MM: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is 52.43. The numbers of clusters are 203 (model data) and 185 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S56. Model TaiESM1.0: Estimated GEV parameters (a, c, e) and their anomaly compared to the reanalysis parameters (b, d, f). The GEV parameters are location (a, b), scale (c, d) and shape (e, f). Values exceeding the scale limits are truncated. Units are mm/month.
Figure S57. Model TaiESM1.0: Differences of the estimated 95% GEV quantiles between model and reanalysis data (a). Clusterings using dissimilarity measure $D_0$ and threshold $h = 0.825$ for model data (b) and reanalysis data (c). The average weighted quantile difference for this model is **58.99**. The numbers of clusters are 214 (model data) and 185 (reanalysis data). Values exceeding the scale limits are truncated. Units are mm/month.
References

Bentsen, M., Oliviè, D. J. L., Seland, O., Toniazzo, T., et al.: NCC NorESM2-MM model output prepared for CMIP6 CMIP historical. Version 20200702, https://doi.org/10.22033/ESGF/CMIP6.8040, 2019.

Boucher, O., Denvil, S., Levasseuer, G., Cozic, A., et al.: IPSL IPSL-CM6A-LR model output prepared for CMIP6 CMIP historical. Version 20180803, https://doi.org/10.22033/ESGF/CMIP6.5195, 2018.

Cao, J. and Wang, B.: NUIST NESMv3 model output prepared for CMIP6 CMIP historical. Version 20190701, https://doi.org/10.22033/ESGF/CMIP6.8769, 2019.

Danabasoglu, G.: NCAR CESM2 model output prepared for CMIP6 CMIP historical. Version 20190514, https://doi.org/10.22033/ESGF/CMIP6.7627, 2019.

Danabasoglu, G.: NCAR CESM2-FV2 model output prepared for CMIP6 CMIP historical. Version 20200226, https://doi.org/10.22033/ESGF/CMIP6.11297, 2019.

Danabasoglu, G.: NCAR CESM2-W ACCM-FV2 model output prepared for CMIP6 CMIP historical. Version 20200226, https://doi.org/10.22033/ESGF/CMIP6.11298, 2019.

Danek, C., Shi, X., Stepanek, C., Yang, H., Barbi, D., Hegewald, J., and Lohmann, G.: AWI AWI-ESM1.1LR model output prepared for CMIP6 CMIP historical. Version 20200212, https://doi.org/10.22033/ESGF/CMIP6.9328, 2020.

Dix, M., Bi, D., Dobrohotoff, P., Fiedler, R., et al.: CSIRO-ARCCSS ACCESS-CM2 model output prepared for CMIP6 CMIP historical. Version 20210607, https://doi.org/10.22033/ESGF/CMIP6.4271, 2019.

EC-Earth Consortium: EC-Earth-Consortium EC-Earth3-Veg-LR model output prepared for CMIP6 CMIP historical. Version 20200217, https://doi.org/10.22033/ESGF/CMIP6.4707, 2020.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci Model Dev, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Hajima, T., Abe, M., Arakawa, O., Suzuki, T., et al.: MIROC MIROC-ES2L model output prepared for CMIP6 CMIP historical. Version 20191118, https://doi.org/10.22033/ESGF/CMIP6.5602, 2019.

Jungclaus, J., Bittner, M., Wieners, K.-H., Wachsmann, F., et al.: MPI-M MPI-ESM1.2-HR model output prepared for CMIP6 CMIP historical. Version 20190710, https://doi.org/10.22033/ESGF/CMIP6.6594, 2019.

Lee, W.-L. and Liang, H.-C.: AS-RCEC TaiESM1.0 model output prepared for CMIP6 CMIP historical. Version 20210416, https://doi.org/10.22033/ESGF/CMIP6.9755, 2020.

Li, L.: CAS FGOALS-g3 model output prepared for CMIP6 CMIP historical. Version 20190818, https://doi.org/10.22033/ESGF/CMIP6.3356, 2019.

NASA/GISS: NASA-GISS GISS-E2.1G model output prepared for CMIP6 CMIP historical. Version 20180830, https://doi.org/10.22033/ESGF/CMIP6.7127, 2018.

Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Stoll, J., et al.: HAMMOZ-Consortium MPI-ESM1.2-HAM model output prepared for CMIP6 CMIP historical. Version 20191218, https://doi.org/10.22033/ESGF/CMIP6.5016, 2019.

Rong, X.: CAMS CAMS_CSM1.0 model output prepared for CMIP6 CMIP historical. Version 20190829, https://doi.org/10.22033/ESGF/CMIP6.9754, 2019.

Seland, O., Bentsen, M., Oliviè, D. J. L., Toniazzo, T., et al.: NCC NorESM2-LM model output prepared for CMIP6 CMIP historical. Version 20190920, https://doi.org/10.22033/ESGF/CMIP6.8036, 2019.

Semmler, T., Danilov, S., Rackow, T., Sidorenko, D., et al.: AWI AWI-CM1.1MR model output prepared for CMIP6 CMIP historical. Version 20200511, https://doi.org/10.22033/ESGF/CMIP6.2686, 2018.

Voldoire, A.: CNRM-CERFACS CNRM-CM6-1 model output prepared for CMIP6 CMIP historical. Version 20200529, https://doi.org/10.22033/ESGF/CMIP6.1375, 2018.

Volodin, E., Mortikov, E., Gritsun, A., Lykossov, V., et al.: INM INM-CM5-0 model output prepared for CMIP6 CMIP historical. Version 20190709, https://doi.org/10.22033/ESGF/CMIP6.5070, 2019.
Volodin, E., Mortikov, E., Gritsun, A., Lykossov, V., et al.: INM INM-CM4-8 model output prepared for CMIP6 CMIP historical. Version 20190530, https://doi.org/10.22033/ESGF/CMIP6.1422, 2019b.

Wu, T., Chu, M., Dong, M., and Fang, Y.: BCC BCC-CSM2MR model output prepared for CMIP6 CMIP historical. Version 20181126, https://doi.org/10.22033/ESGF/CMIP6.2948, 2018.

Yu, Y.: CAS FGOALS-f3-L model output prepared for CMIP6 CMIP historical. Version 20190927, https://doi.org/10.22033/ESGF/CMIP6.3355, 2019.

Yukimoto, S., Koshiro, T., Kawai, H., Oshima, N., et al.: MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP historical. Version 20210813, https://doi.org/10.22033/ESGF/CMIP6.6842, 2019.

Ziehn, T., Chamberlain, M., Lenton, A., Law, R., et al.: CSIRO ACCESS-ESM1.5 model output prepared for CMIP6 CMIP historical. Version 20200529, https://doi.org/10.22033/ESGF/CMIP6.4272, 2019.