Supplementary Material for "Extreme Precipitation Return Levels for Multiple Durations on a Global Scale"

August 25, 2021

Contents

1 Calendar and Hydrological Year

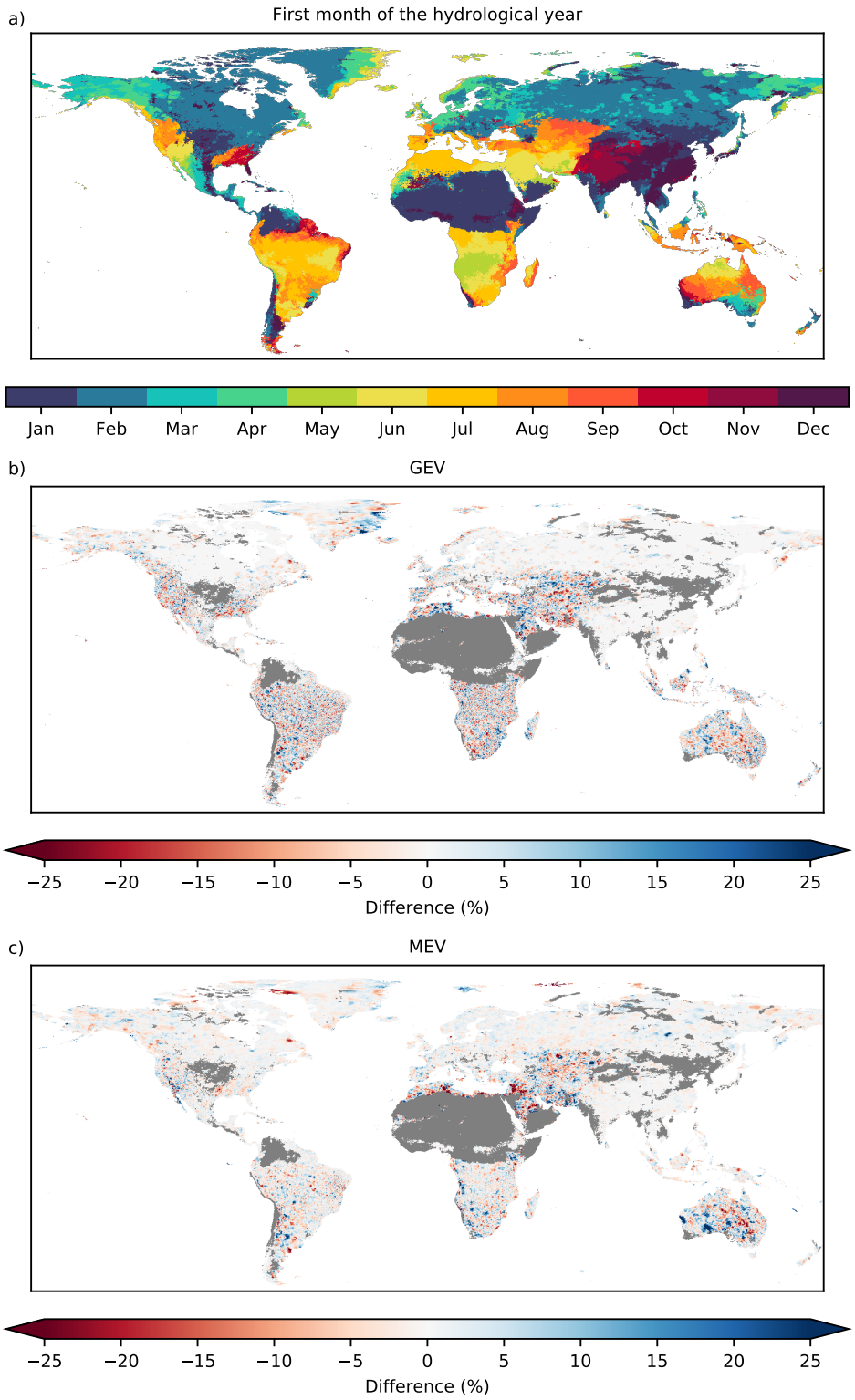

Figure S1: (a) The month indicating the start of the hydrological year. (b) The percentage difference of the daily 1000-year return levels of calendar and hydrological years for GEV. (c) The percentage difference of the daily 1000-year return levels of calendar and hydrological years for MEV. A negative difference indicates that the T1000 estimate is larger using hydrological years, whereas a positive difference indicates that the T1000 estimate is larger using calendar years.

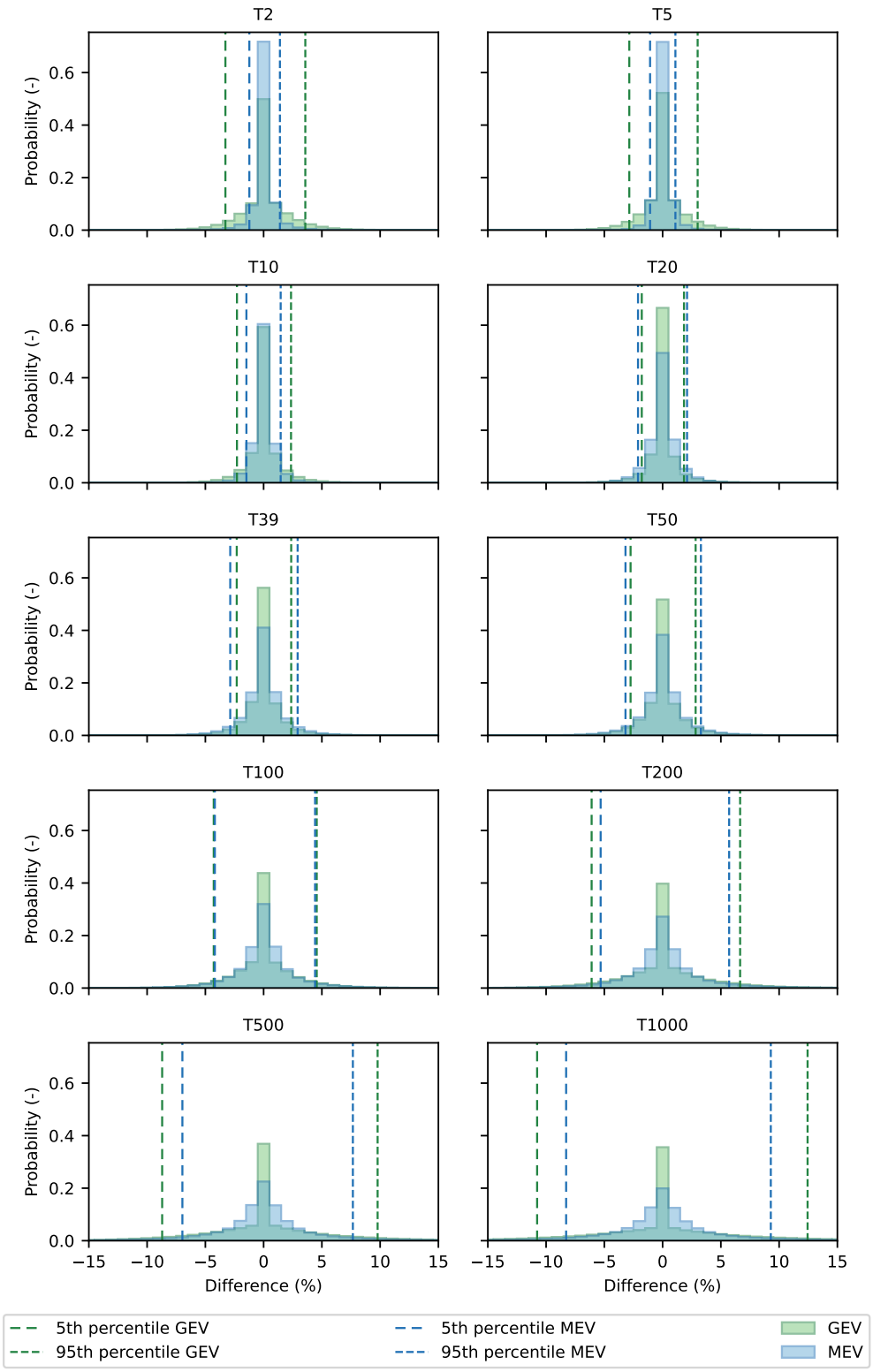

Figure S2: Weighted histograms showing the percentage difference in the values of several return periods (T, see titles) calculated using calendar years and hydrological years. Included in the figure are all cells where the start of the hydrological year is different from the calendar year (i.e., the hydrological year does not start in January, see Figure S1a). A negative difference indicates that the return period estimate is larger using hydrological years, whereas a positive difference indicates that the return period estimate is larger using calendar years.

2 Threshold Analysis for POT

There are many different ways to select an appropriate threshold for the Peak Over Threshold (POT) analysis, such as a value over a specific threshold, a percentage, or an average number of events per year. We refer to Caeiro and Gomes (2016); Langousis, Mamalakis, Puliga, and Deidda (2016) for recent overviews of such threshold selection. Which method is the most effective remains to be an unanswered question. For our global-scale application, we have analyzed to take on average 1 to 5 events per year. This fixed number of events, as opposed to the events over a predefined threshold, ensures that the number of events per year remains constant for all durations, and allows for analyses on the global domain.

The 100-year return levels for all durations and thresholds are shown in Figure S3a. The boxplot shows minimal difference in the $T100$ estimates for the different thresholds for the short durations, and a minor difference for the longer durations, namely that the estimated $T100$ are slightly larger for inclusion of less events per year. Overall the figure shows only slight differences, so more information has to be considered for accurate threshold selection.

A comparison of the three parameters for all durations and thresholds is shown in Figure S3b-d. The location and scale parameters in S3b and c show very similar results for the different thresholds. The shape parameter in S3d, however, has more variation for the different POT-thresholds. For all durations, the variability of the shape parameter decreases for the inclusion of more events per year, so more events to fit the distribution to. For durations between 3 and 48 hours, the shape slightly increases for the inclusion of more events per year. In other words, the shape is lower for on average 1 event per year, and higher for on average 5 events per year for durations between 3 and 48 hours. For 72 hours, the median shape parameter of the different thresholds is constant, though the variability decreases for the inclusion of more events. For the 5 and 10-day durations (120 and 240 hours, respectively), the shape parameter slightly decreases for the inclusion of more (non-extreme) events. The general pattern of the GP shape parameter is similar to that of the GEV, as they both show a decreasing shape parameter for increasing durations.

Previous studies that have looked into the GP shape parameter on the global domain focused on daily durations. Papalexiou and Koutsoyiannis (2013) estimated the mode of the shape for daily precipitation as 0.134, but displaying a large spread. Serinaldi and Kilsby (2014) estimated the shape for four different seasons based on 1898 stations with more than 100 years of data. They found the shape depends on the season and varies between 0.061 and 0.097, also displaying a large spread. Furthermore, they found that if you lower the threshold and include more non-extreme events, the shape parameter is higher. Our results are similar to that of Serinaldi and Kilsby (2014), as the median for daily precipitation for the different thresholds varies between 0.711 and 0.918, and we also found that the inclusion of more non-extreme events leads to a higher shape parameter for daily precipitation.

In short, the above analysis shows that the threshold selection is of minor influence on the estimation of the 100-year return level. Of the three underlying parameters, the threshold selection has minimal influence on the location and scale parameter, though a much greater influence on the shape parameter. The variability of the shape parameter is the highest for the threshold of on average 1 event per year, and decreases when more non-extreme events are used to fit the distribution to. Furthermore, the shape parameter increases for the inclusion of more events for the shorter durations (3-hours to 2-days), remains constant for a 3-day duration, and decreases for the 5 and 10-day durations. Based on these analyses, we chose to show the threshold of on average 3 events per year in the main manuscript, as that threshold has the right balance of a low variability and not as much of an underestimation of the 100-year return levels for the longest durations.

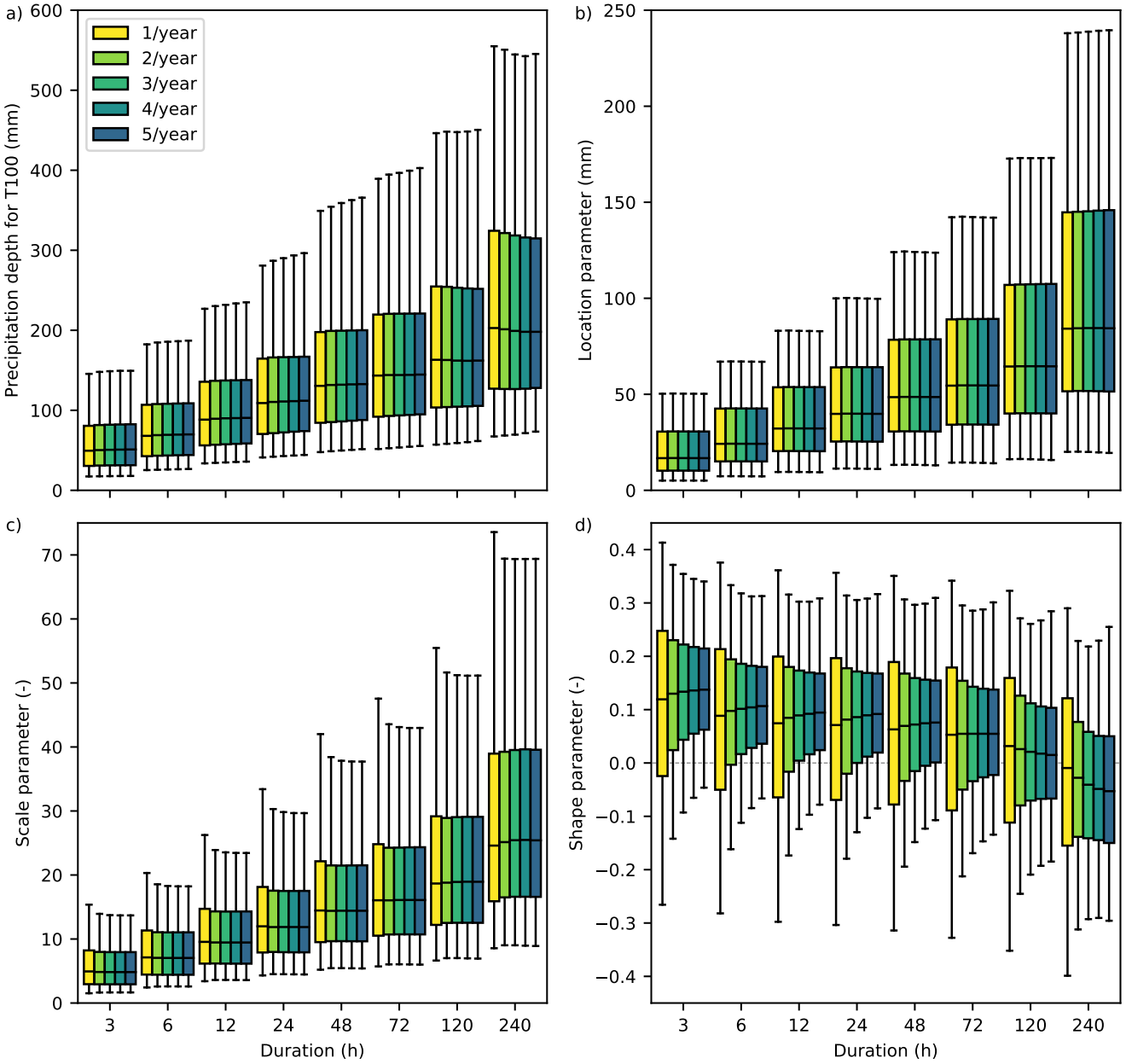

Figure S3: Weighted boxplot showing the distribution of (a) the 100-year precipitation return levels, and the (b) location, (c) scale, and (d) shape parameters for several durations and thresholds for the Peak Over Threshold method. Thresholds range from on average one to five events per year. The top and bottom of the boxes represent the 75th and 25th percentiles, respectively. The whiskers denote the 1st and 99th percentiles. The dashed gray horizontal line in (d) indicates exponential tails. A shape parameter smaller than zero indicates a thin tail with an upper limit, a shape parameter larger than zero indicates a heavy power-law tail.

3 Shape Parameter

The GEV and MEV distributions are both flexible and able to describe different tail behaviors. The tail behavior the two distributions varies, see Section 2.2.2 in the main manuscript for an overview and Figure S4 for different combinations of scale and shape parameters.

Maps of the shape parameter for the three distributions for daily duration are shown in Figure S5. The spatial patterns of the shape parameters are largely similar to those of the heaviness $h_{\mathrm{T10-T100-T1000}}$, described in Section 3.3 and Figures 5 and 6 of the main manuscript. There are, however, some notable differences. For MEV, the mountainous areas are more defined, indicated by higher shape values (less heavy tail behavior). Looking at the shape parameter, we find, however, that the relationship of the Weibull shape parameter with elevation is more complicated. Lower shapes (heavier tails) are generally observed on the leeward side of large mountain ranges, and higher shapes (though still indicative of heavy-tail behavior) on the windward side that is dominated by orographically enhanced frontal precipitation. This is for example visible in the Rocky Mountains, Indonesia, and Norway, and corresponds to findings of Cavanaugh and Gershunov (2015, their Figure 5), who showed that exponential tails are observed in regions where extreme precipitation is predominantly generated by one type of system.

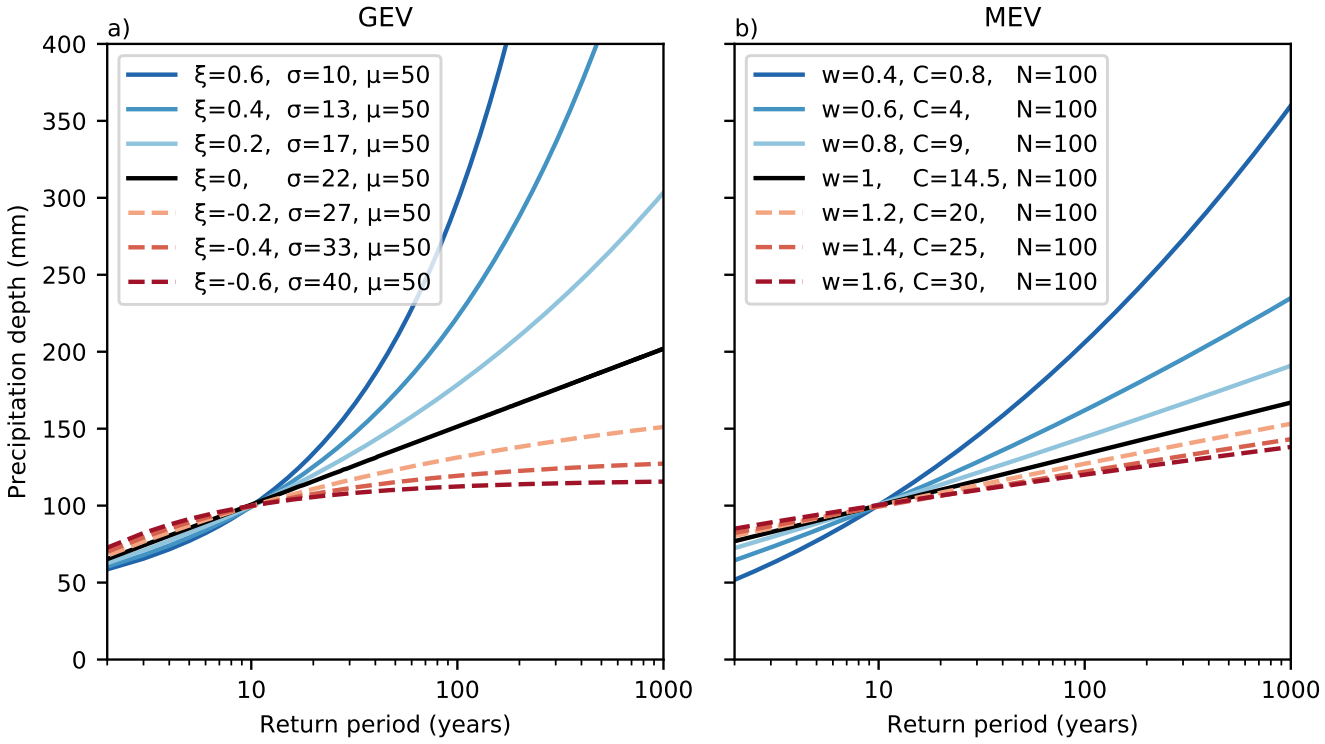

Figure S4: (a) Behavior of the shape (ξ) and scale (σ) parameters of the GEV distribution, with a constant location parameter (μ), and (b) behavior of the shape (w) and scale (C) parameters of the MEV-Weibull distribution, with a constant number of events (N). The results for MEV have been obtained with constant w, C and N parameters for each year. The values of the shape and scale parameter pairs have been chosen such that they all have a precipitation depth of approximately 100 mm for a 10-year return period.

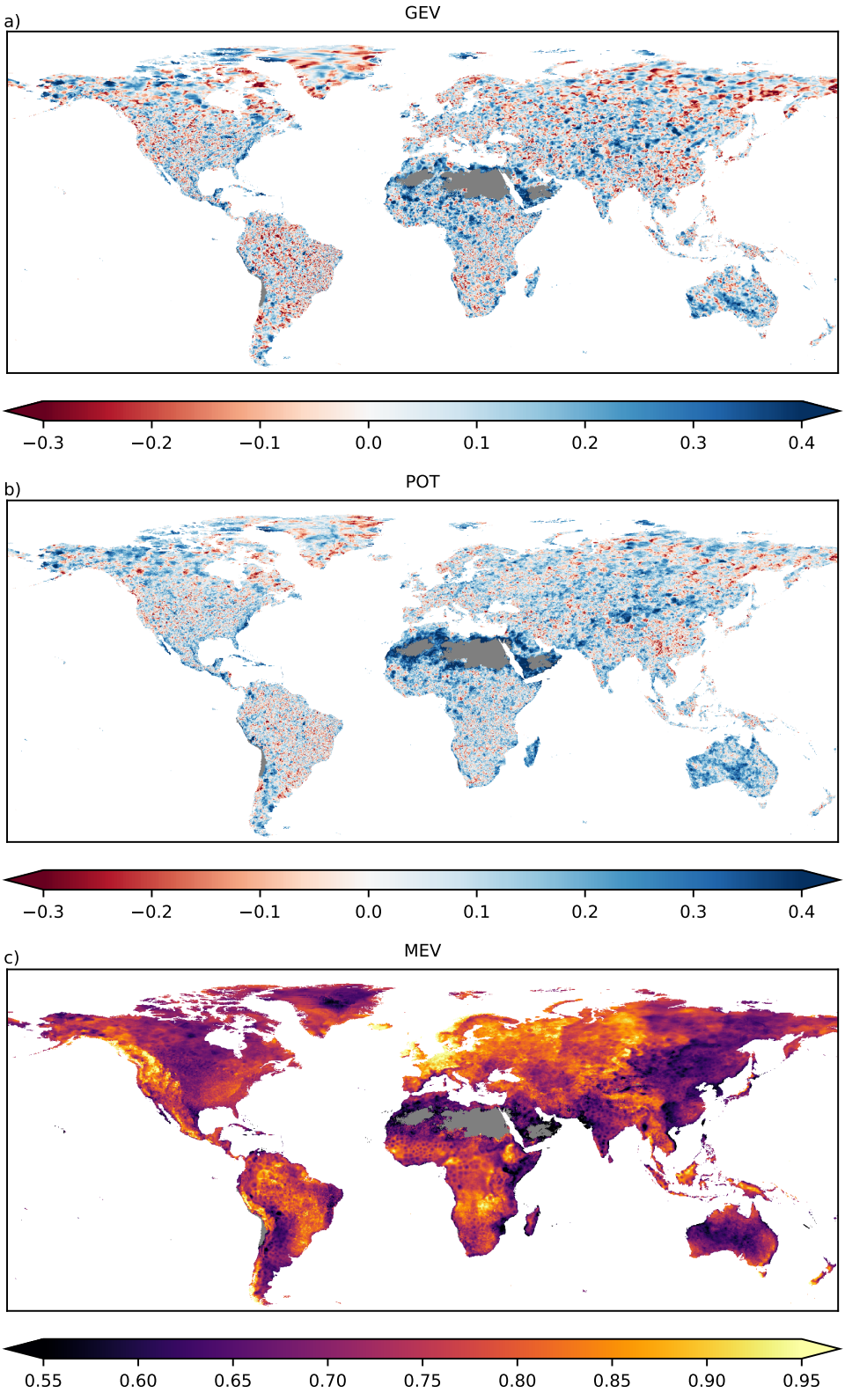

Figure S5: The shape parameter (−) for daily precipitation calculated for different extreme value methods: (a) GEV, equation 1 — ξ_{GEV}, (b) POT, equation 2 — ξ_{GP} and (c) MEV, equation 3 — w. For MEV, the mean shape parameter of all yearly Weibull distributions is displayed. The colorbar min and max are based on the 1st and 99th percentile. For GEV and POT, red indicates a thin shape with an upper limit, white an exponential shape, and blue a heavy power-law shape. For MEV all median shapes are indicative of sub-exponential heavy shapes, though darker colors are heavier than lighter colors.

4 Tail Behavior MEV for Different Return Level Combinations

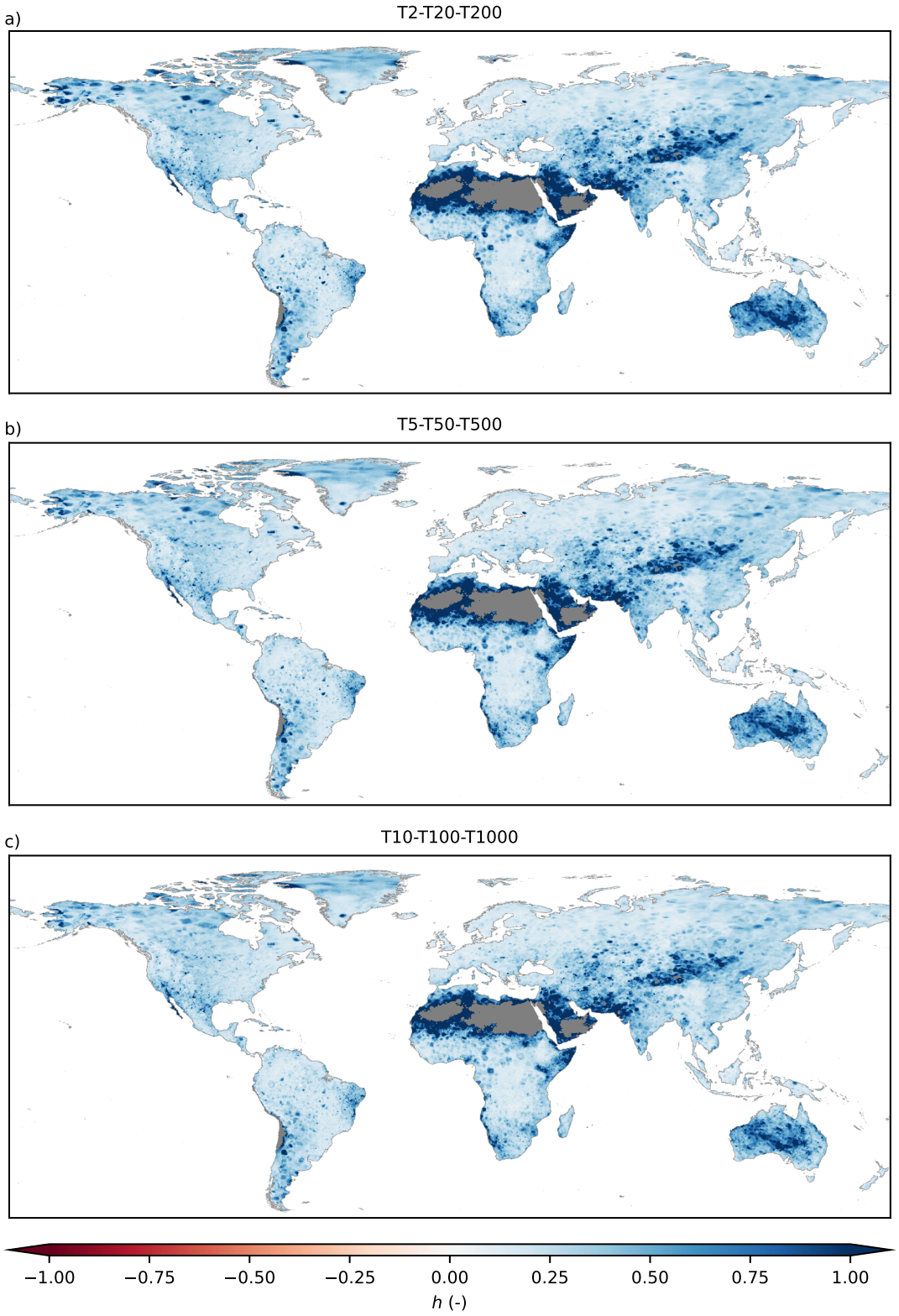

Figure S6: The heaviness amplification factor h (Eq. 5) for daily precipitation calculated for different combinations of return levels for the MEV distribution: (a) $h_{\mathrm{T2-T20-T200}}$, (b) $h_{\mathrm{T5-T50-T500}}$, (c) $h_{\mathrm{T10-T100-T1000}}$. Red indicates a thin tail, white an exponential tail, and blue a heavy tail. See section 2.2.2 for more information on the heaviness metric.

5 Tail Behavior for Multiple Durations

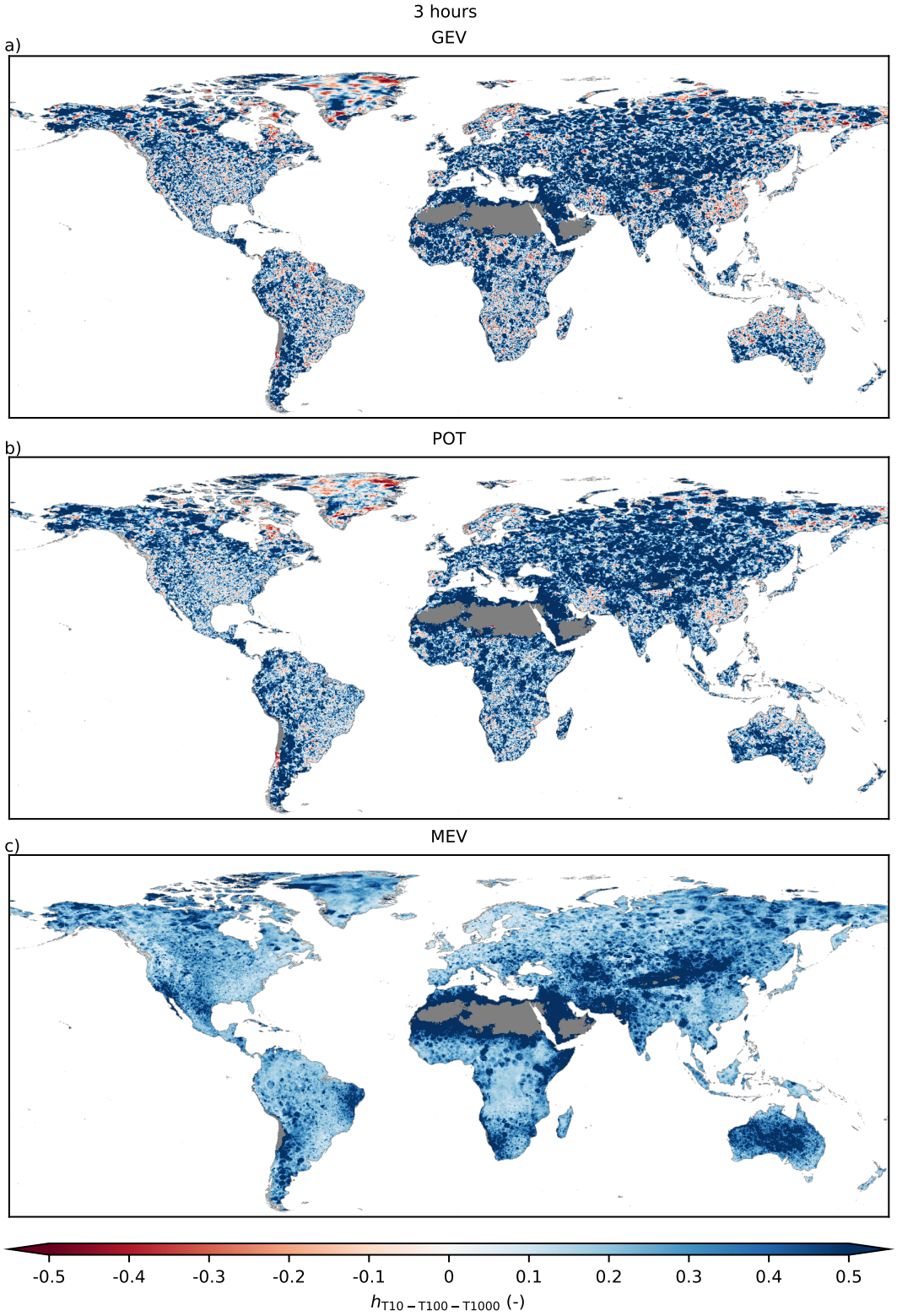

Figure S7: The heaviness amplification factor $h_{\mathrm{T10–T100–T1000}}$ (Eq. 5) for 3-hourly precipitation calculated for different extreme value methods: (a) GEV, (b) POT, (c) MEV. Red indicates a thin tail, white an exponential tail, and blue a heavy tail. See section 2.2.2 for more information on the heaviness metric.

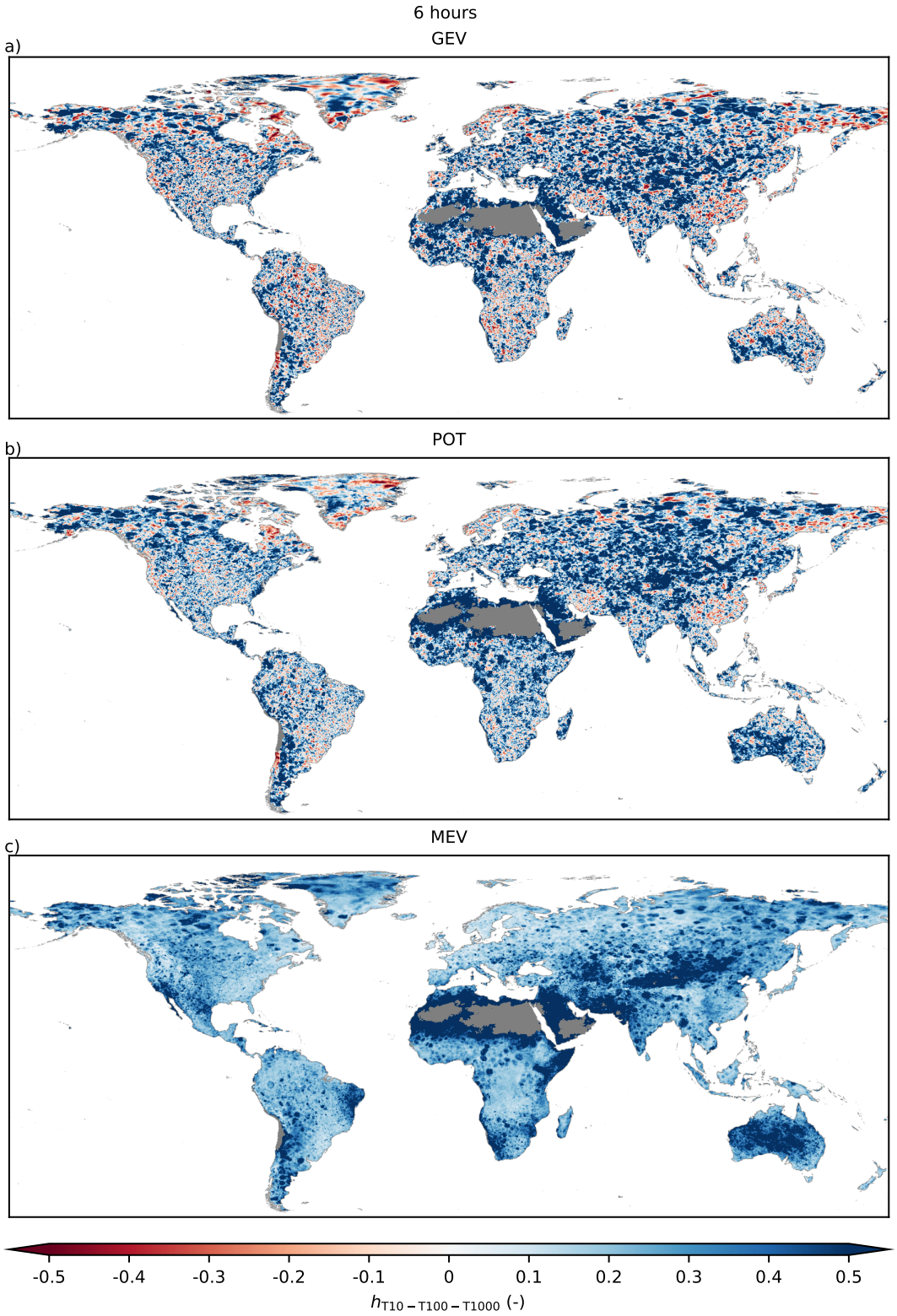

Figure S8: The heaviness amplification factor $h_{\mathrm{T10-T100-T1000}}$ (Eq. 5) for 6-hourly precipitation calculated for different extreme value methods: (a) GEV, (b) POT, (c) MEV. Red indicates a thin tail, white an exponential tail, and blue a heavy tail. See section 2.2.2 for more information on the heaviness metric.

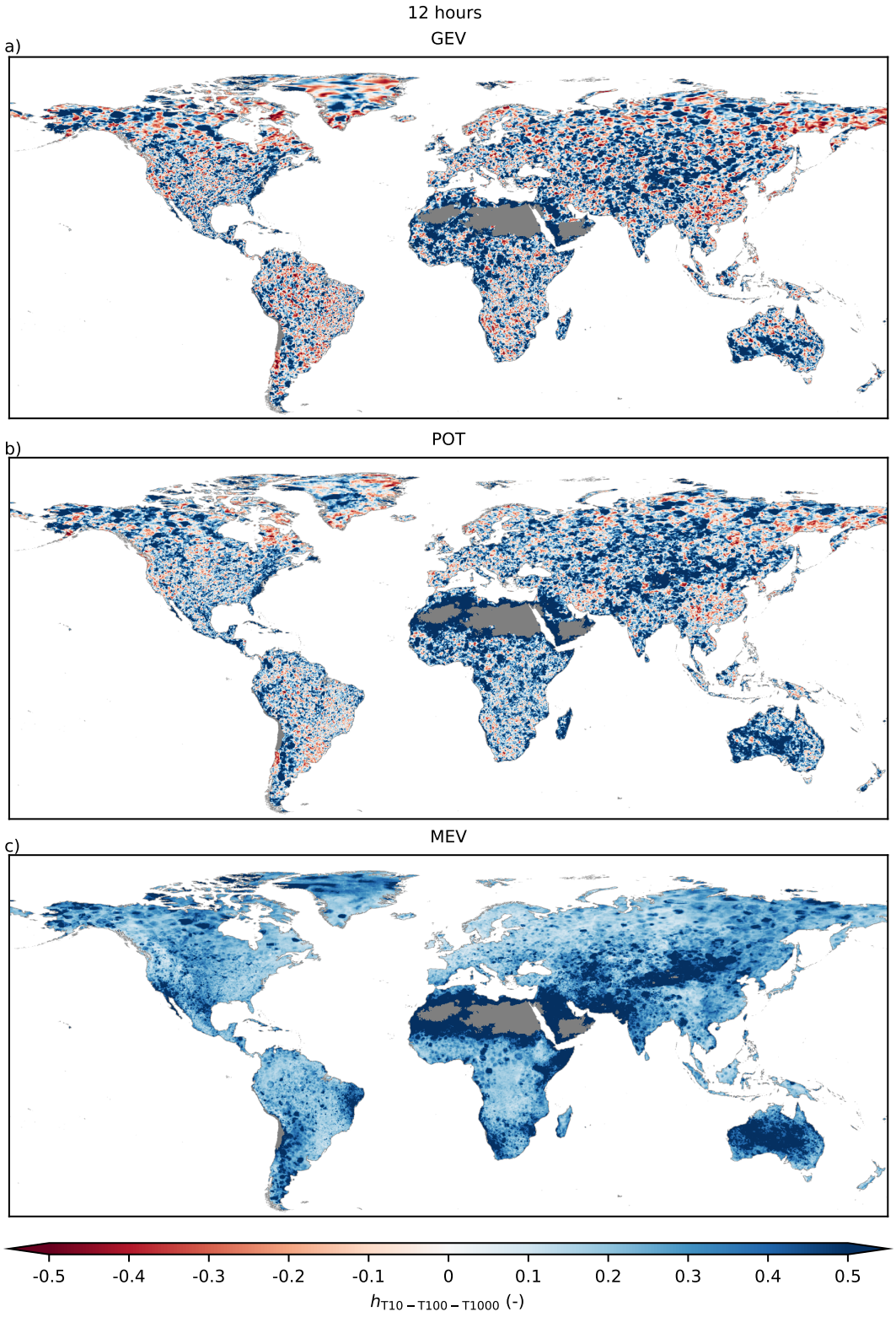

Figure S9: The heaviness amplification factor $h_{\mathrm{T10-T100-T1000}}$ (Eq. 5) for 12-hourly precipitation calculated for different extreme value methods: (a) GEV, (b) POT, (c) MEV. Red indicates a thin tail, white an exponential tail, and blue a heavy tail. See section 2.2.2 for more information on the heaviness metric.

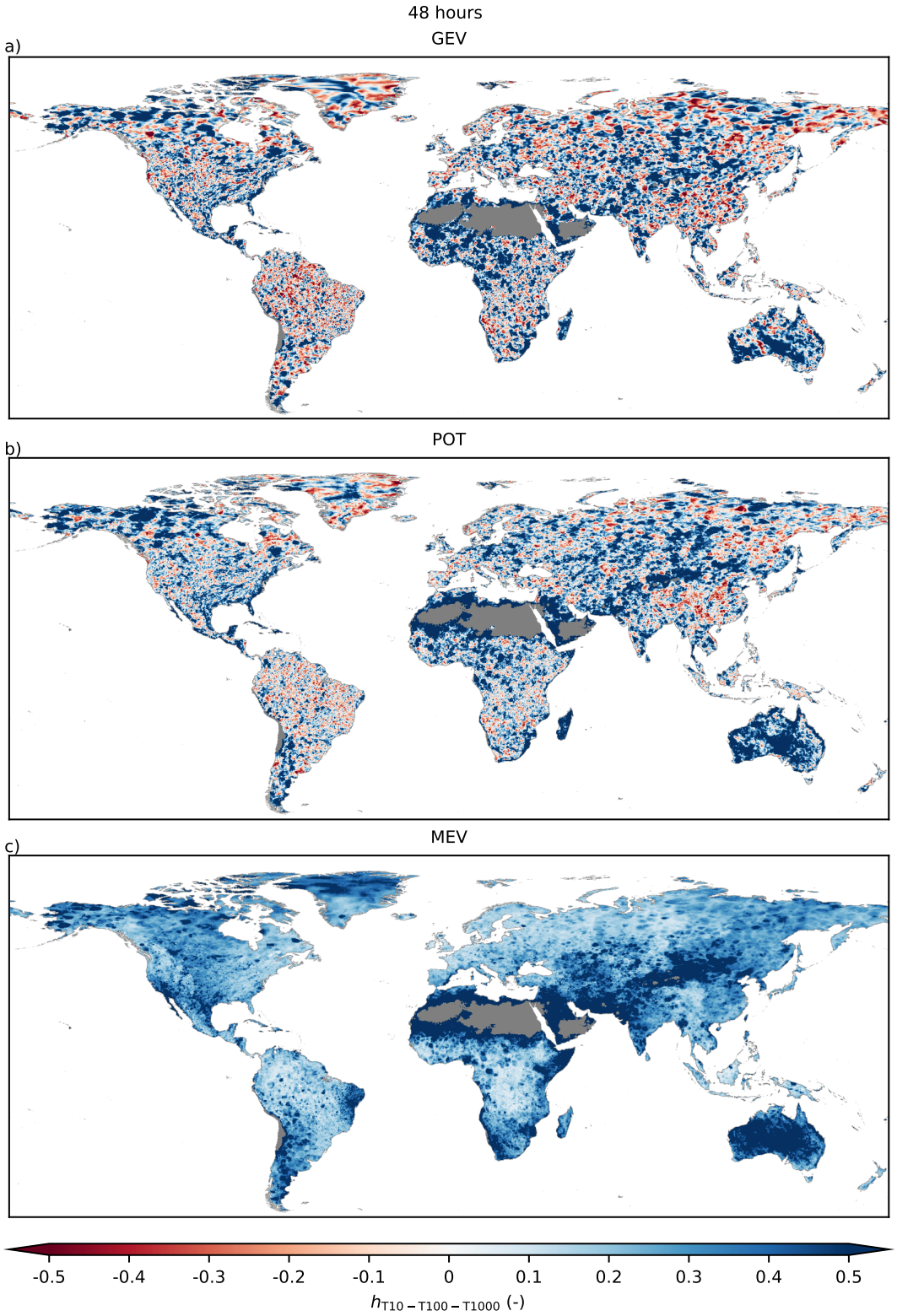

Figure S10: The heaviness amplification factor $h_{\mathrm{T10-T100-T1000}}$ (Eq. 5) for 2-day precipitation calculated for different extreme value methods: (a) GEV, (b) POT, (c) MEV. Red indicates a thin tail, white an exponential tail, and blue a heavy tail. See section 2.2.2 for more information on the heaviness metric.

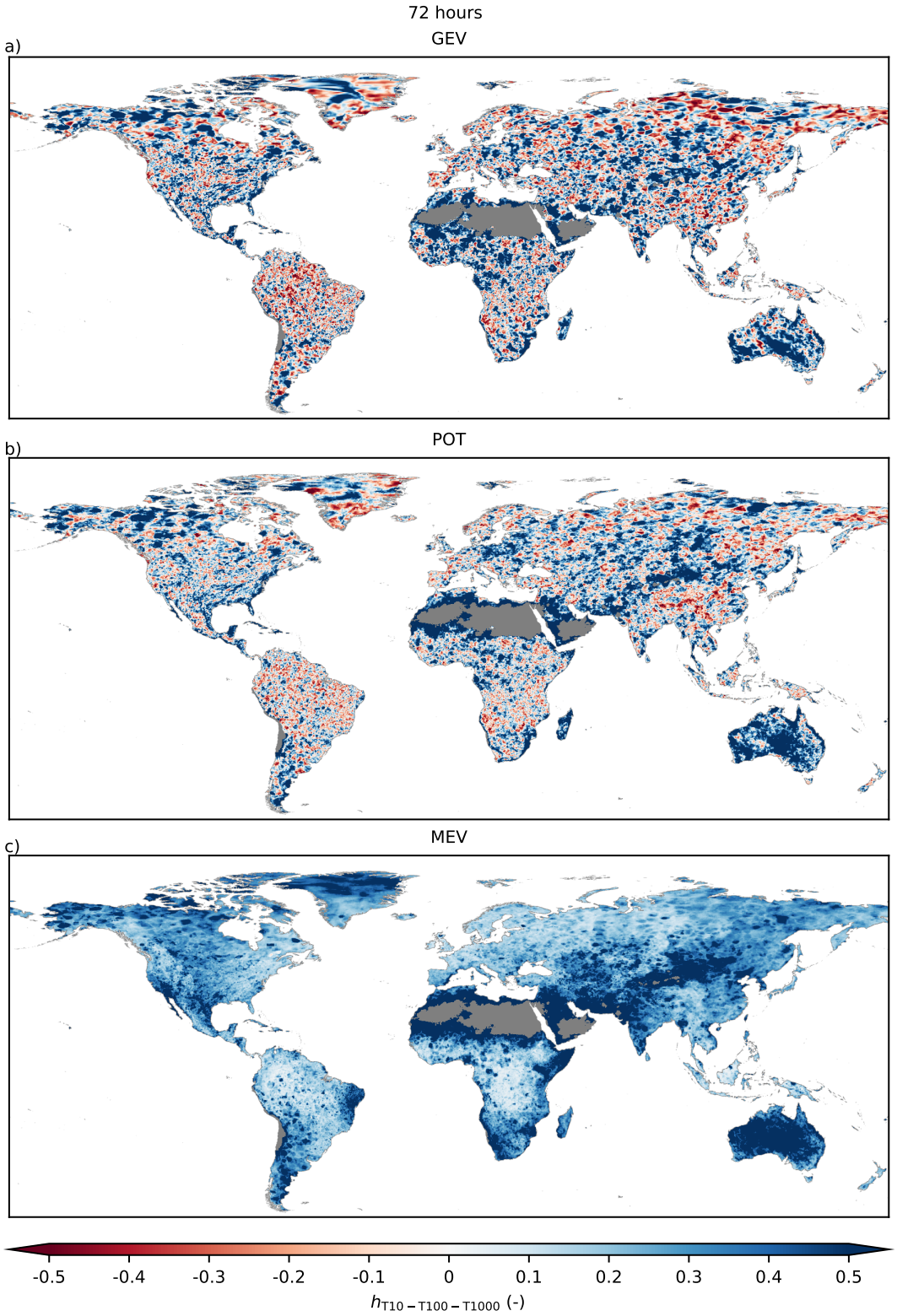

Figure S11: The heaviness amplification factor $h_{\mathrm{T10-T100-T1000}}$ (Eq. 5) for 3-day precipitation calculated for different extreme value methods: (a) GEV, (b) POT, (c) MEV. Red indicates a thin tail, white an exponential tail, and blue a heavy tail. See section 2.2.2 for more information on the heaviness metric.

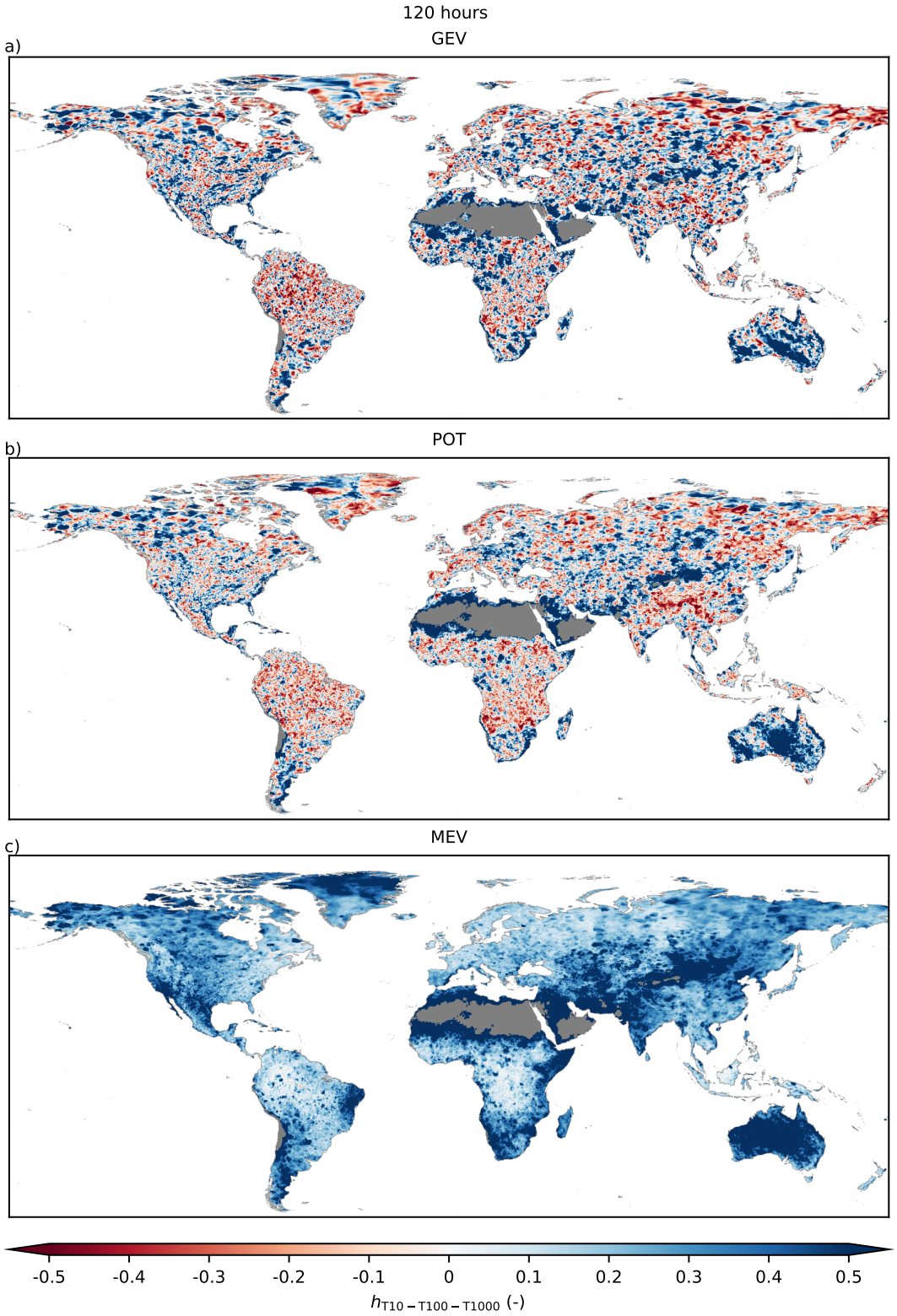

Figure S12: The heaviness amplification factor $h_{\mathrm{T10-T100-T1000}}$ (Eq. 5) for 5-day precipitation calculated for different extreme value methods: (a) GEV, (b) POT, (c) MEV. Red indicates a thin tail, white an exponential tail, and blue a heavy tail. See section 2.2.2 for more information on the heaviness metric.

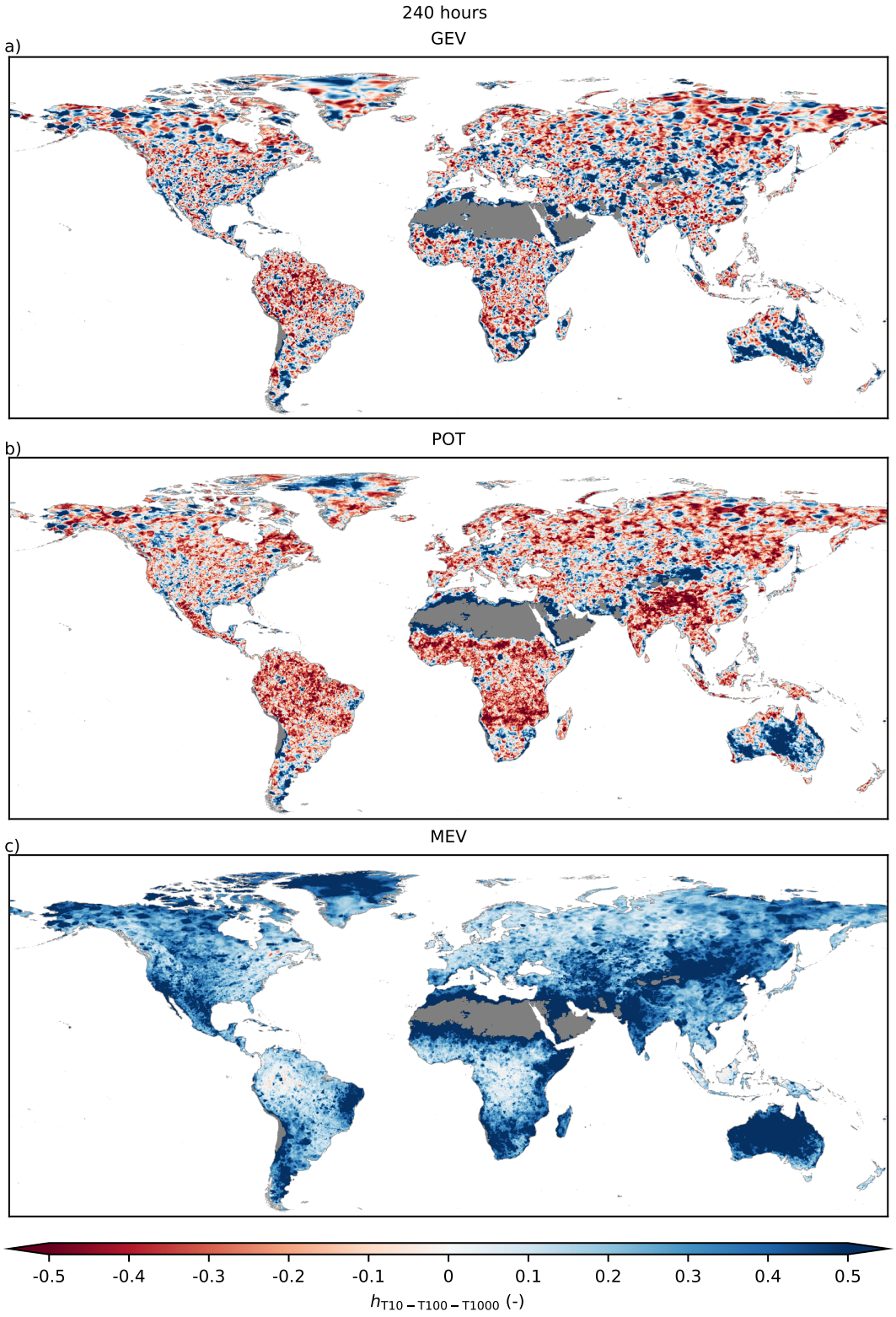

Figure S13: The heaviness amplification factor $h_{\mathrm{T10-T100-T1000}}$ (Eq. 5) for 10-day precipitation calculated for different extreme value methods: (a) GEV, (b) POT, (c) MEV. Red indicates a thin tail, white an exponential tail, and blue a heavy tail. See section 2.2.2 for more information on the heaviness metric.

6 Dataset Usage Notes

The GPEX dataset created in this study is available at the 4TU repository (Gründemann et al., 2021). It provides openly accessible and readily available hydrologically relevant return levels of extreme precipitation estimates worldwide. It contains the precipitation estimates of the three extreme values distributions, the observed estimates, the parameters of the three distributions, as well as a few other variables used in this study (Table S1). Furthermore, the extreme precipitation estimates and parameters of the Gumbel distribution are included in the dataset. In this section we provide some possible uses of the dataset, and instructions and disclaimers for proper use, both for large or regional-scale usage as well as for a single cell or point-scale.

Table S1: Variables included in the GPEX dataset.

Variable	Description
GEV estimate	Extreme precipitation return levels estimated using GEV (mm)
POT estimate	Extreme precipitation return levels estimated using POT (mm)
MEV estimate	Extreme precipitation return levels estimated using MEV (mm)
Gumbel estimate	Extreme precipitation return levels estimated using Gumbel (mm)
Observed estimate	Observed extreme precipitation return levels (mm)
GEV location parameter	Location parameter of the GEV distribution
GEV scale parameter	Scale parameter of the GEV distribution
GEV shape parameter	Shape parameter of the GEV distribution
GEV heaviness	Heaviness amplification factor ($h_{\mathrm{T10-T100-T1000}}$) of the GEV distribution
POT location parameter	Location parameter for a GEV distribution estimated by fitting the GP distribution
POT scale parameter	Scale parameter for a GEV distribution estimated by fitting the GP distribution
POT shape parameter	Shape parameter for a GEV distribution estimated by fitting the GP distribution
POT heaviness	Heaviness amplification factor ($h_{\mathrm{T10-T100-T1000}}$) of the POT distribution
MEV number of events	n parameter of the MEV distribution, number of events per hydrological year
MEV scale parameter	Scale parameter of the MEV distribution
MEV shape parameter	Shape parameter of the MEV distribution
MEV mean number of events	Mean of the n parameter of the MEV distribution, mean number of events per hydrological year
MEV mean scale parameter	Mean of the scale parameter of the MEV distribution
MEV mean shape parameter	Mean of the shape parameter of the MEV distribution
MEV heaviness	Heaviness amplification factor ($h_{\mathrm{T10-T100-T1000}}$) of the MEV distribution
Gumbel scale parameter	Scale parameter of the Gumbel distribution
Gumbel shape parameter	Shape parameter of the Gumbel distribution
Annual maxima	Annual maximum precipitation for each hydrological year (mm)
Start hydrological year	Number indicating the month in which the hydrological year starts
Running parameter	Running parameter (hours) of the declustering method by Marra et al. (2018)
Land mask	Mask used for this study to indicate land cells and ocean cells

6.1 Large-Scale Applications

The GPEX dataset contains global scale extreme precipitation estimates and its parameters at a spatial resolution of 0.1°, covering 3-hourly to 10-day durations. The dataset contains information about precipitation extremes for the entire Earth's land surface except Antarctica. The estimates of three distributions described in the main manuscript, as well as the return levels as observed, and estimated using the Gumbel distribution are included in the dataset. Among the three distributions, the traditional GEV and POT provide comparable large-scale average extremes, although differences can be substantial at smaller scales. When using the dataset at regional scales, we advise taking the average of the precipitation estimates, as neighboring cells could differ. Note that since only 38 years of data were available, the fitted model parameters and associated return values are subject to considerable uncertainty. Furthermore, we acknowledge that the use of just one dataset does not represent the true uncertainty in the generation of the dataset created. We do not think this affects our results for observed global spatial patterns significantly, but in a practical setting we recommend verifying the estimates with local observations if available, and to reproduce the precipitation return level estimates with a full uncertainty range estimation.

The novel MEV distribution provides more spatially coherent patterns of the extremes. Its mean shape parameter for daily events captures the (heavy-)tail behavior, and follows orographic patterns. The extremes estimated by MEV are higher than those estimated by GEV and POT. However, for large return periods and long durations, MEV can overestimate the

extremes, due to the small number of events available for the fitting. We, therefore, recommend analyzing the extremes of all distributions in this dataset to obtain an indication of the uncertainty.

6.2 Small-Scale Applications

The dataset is also suitable for small-scale applications either in comparative studies or for direct use in data sparse regions, but one should be aware of the different statistical characteristics of point-scale and grid-scale. Due to averaging effects in gridded datasets, precipitation extremes of point-scale observations are higher (Cavanaugh & Gershunov, 2015; De Michele, Kottegoda, & Rosso, 2001; Ensor & Robeson, 2008; Hu et al., 2020; Sivapalan & Blöschl, 1998; Zorzetto & Marani, 2019). Illustrative examples of two locations, Vienna and San Francisco, are included in Figure S14. Analysis of the return level plots shows the estimates of the three distributions discussed in the main manuscript, as well as the estimates using the Gumbel distribution compared to the observed ones. We converted the annual maximum precipitation to 'observed' return levels (Figure S14e+j). It should be kept in mind though that these 'observed' return levels are also different from the 'true' return levels. For (sub-)daily durations and low return periods, there is generally a good agreement between the observed return levels and the estimates of the three EVDs. For longer durations and return periods, however, the estimated extremes deviate from the observed extremes. This is seen in San Fransisco (Figure S14f-j) where MEV overestimates and GEV, POT and Gumbel underestimate the extremes.

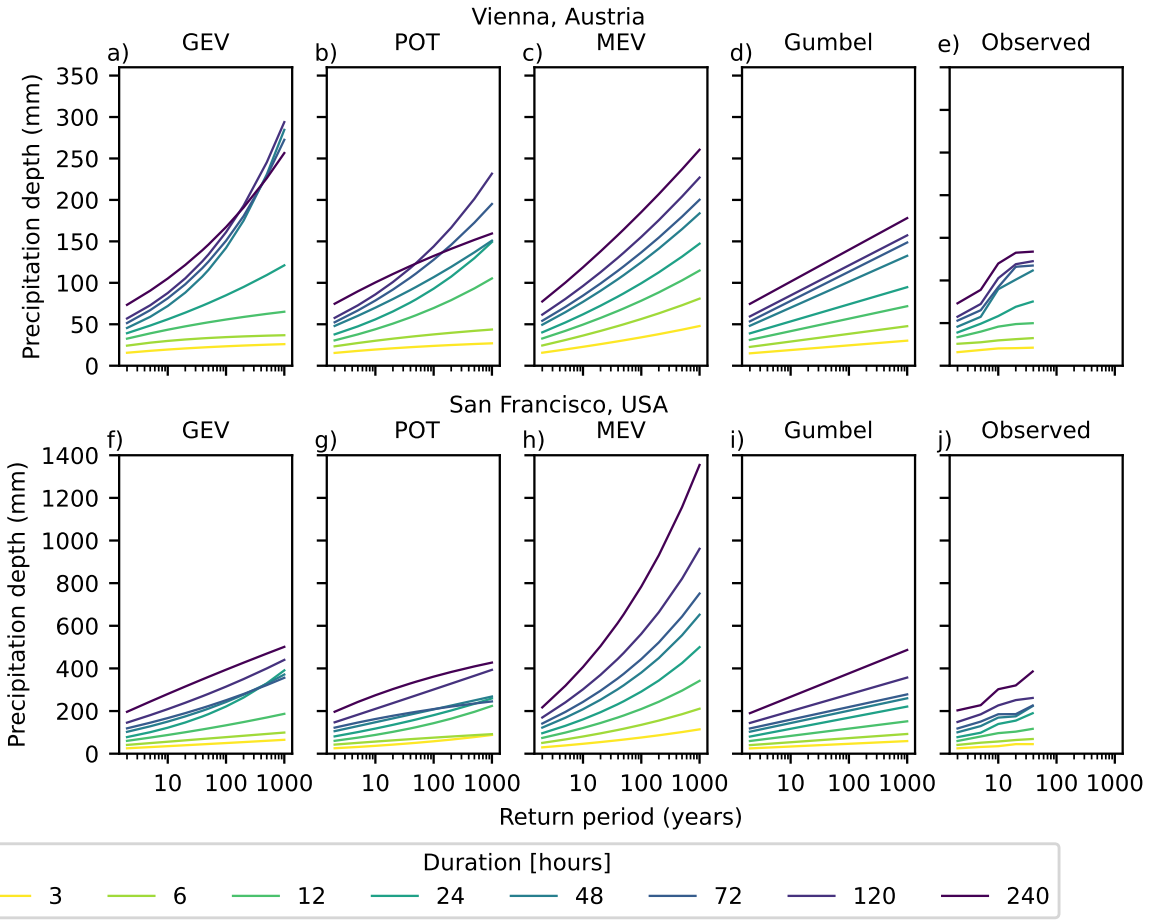

Figure S14: Return level plots for specific locations and different distributions. a–e) Vienna, Austria (48.234°N, 16.415°E) and f–j) San Francisco, California, USA (37.784°N, 122.400°W). Observed (e & j) are the annual maxima converted to return periods.

Furthermore, increasing event durations result in lower shape parameters (less heavy tails), which was seen for all three distributions discussed in the main manuscript. An implication of this is that for long-durations the shape parameter indicates a finite endpoint (GEV and POT), or a very thin tail (MEV), while heavier tails are generally observed for short-durations. When estimating very large return periods (e.g., $T500$), it is therefore possible for shorter duration estimates to be more intense

than the corresponding quantiles computed for longer durations, which is physically impossible (see also Figure S14a,b,f and g). Additionally, Papalexiou and Koutsoyiannis (2013) argued based on long time series that the true population GEV shape parameter of daily precipitation lies between exponential and heavy-tail behavior, and thin-tails lead to an underestimation of the extremes. For applications where underestimation is undesirable, we have included the extreme precipitation estimates using the Gumbel distribution as well. The Gumbel distribution is an exponential distribution, and equal to the GEV distribution where the shape parameter equals zero. Therefore, if the shape parameter of GEV or POT is negative, the Gumbel estimates could be used instead in order to avoid underestimation.

To get a better understanding of the range and uncertainty of a single cell location, we recommend to look at return level plots of the four distributions at the cell of interest in combination with its neighboring cells. This is particularly important for GEV and POT, due to the absence of coherent spatial patterns and the erratic manifestation of the tail behaviors. Previous results (Zorzetto, Botter, & Marani, 2016) show that the benefits of MEV over GEV are greater for large return periods relative to the sample size available for the fit. Hence, for the estimation of large quantiles, MEV may be presumed to be more accurate. Depending on the practical application one could then choose to use the most extreme value, use the MEV value, use the Gumbel value in case of a negative shape for GEV or POT, or use a spatial average of the GEV and POT estimates.

References

Caeiro, F., & Gomes, M. I. (2016). Threshold selection in extreme value analysis. *Extreme value modeling and risk analysis: Methods and applications*, *1*, 69–86.

Cavanaugh, N. R., & Gershunov, A. (2015). Probabilistic tail dependence of intense precipitation on spatiotemporal scale in observations, reanalyses, and GCMs. *Climate Dynamics*, *45*(11-12), 2965–2975. doi: 10.1007/s00382-015-2517-1

De Michele, C., Kottegoda, N. T., & Rosso, R. (2001). The derivation of areal reduction factor of storm rainfall from its scaling properties. *Water Resources Research*, *37*(12), 3247–3252. doi: 10.1029/2001WR000346

Ensor, L. A., & Robeson, S. M. (2008). Statistical characteristics of daily precipitation: comparisons of gridded and point datasets. *Journal of Applied Meteorology and Climatology*, *47*(9), 2468–2476. doi: 10.1175/2008JAMC1757.1

Gründemann, G. J., Zorzetto, E., Beck, H. E., Schleiss, M., van de Giesen, N., Marani, M., & van der Ent, R. J. (2021). *Global Precipitation EXtremes dataset.* doi: 10.4121/uuid:12b5c941-cd54-45db-8d7b-fefaacecaa69

Hu, L., Nikolopoulos, E. I., Marra, F., Morin, E., Marani, M., & Anagnostou, E. N. (2020). Evaluation of MEVD-based precipitation frequency analyses from quasi- global precipitation datasets against dense rain gauge networks. *Journal of Hydrology*, *590*, 125564. doi: 10.1016/j.jhydrol.2020.125564

Langousis, A., Mamalakis, A., Puliga, M., & Deidda, R. (2016). Threshold detection for the generalized pareto distribution: Review of representative methods and application to the noaa ncdc daily rainfall database. *Water Resources Research*, *52*(4), 2659–2681.

Papalexiou, S. M., & Koutsoyiannis, D. (2013). Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, *49*, 187–201. doi: 10.1029/2012WR012557

Serinaldi, F., & Kilsby, C. G. (2014). Rainfall extremes: Toward reconciliation after the battle of distributions. *Water Resources Research*, *50*, 336–352. doi: 10.1002/2013WR014211

Sivapalan, M., & Blöschl, G. (1998). Transformation of point rainfall to areal rainfall: Intensity-duration-frequency curves. *Journal of Hydrology*, *204*(1-4), 150–167.

Zorzetto, E., Botter, G., & Marani, M. (2016). On the emergence of rainfall extremes from ordinary events. *Geophysical Research Letters*, *43*(15), 8076–8082. doi: 10.1002/2016GL069445

Zorzetto, E., & Marani, M. (2019). Downscaling of rainfall extremes from satellite observations. *Water Resources Research*, *55*, 156–174. doi: 10.1029/2018WR022950