

# Positive correlation between wet-day frequency and intensity linked to universal precipitation drivers

Received: 11 May 2022

Accepted: 25 March 2023

Published online: 1 May 2023

 Check for updates

Cameron McErlich<sup>1</sup>✉, Adrian McDonald<sup>1,2</sup>, Alex Schuddeboom<sup>1</sup>, Gokul Vishwanathan<sup>1</sup>, James Renwick<sup>3</sup> & Sapna Rana<sup>4</sup>

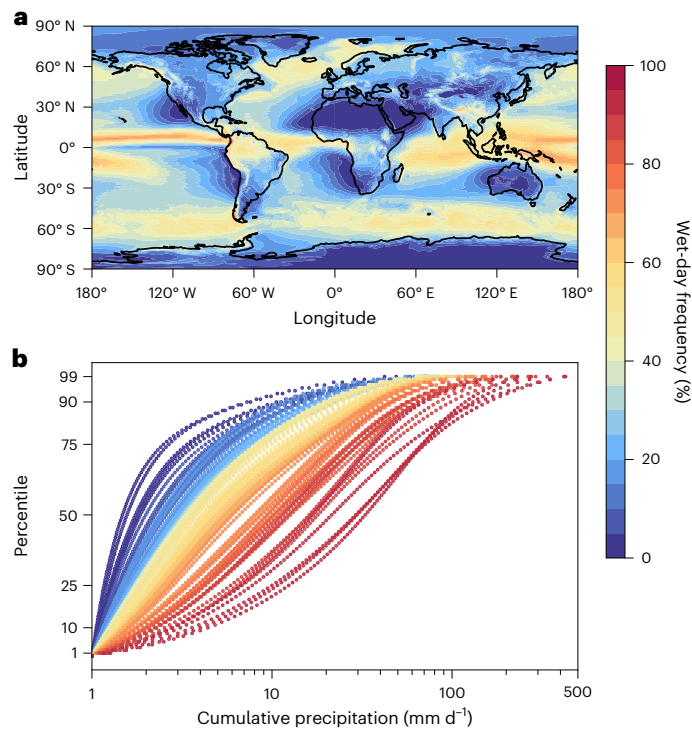
Understanding precipitation is essential for quantifying weather and climate-related risks. Changes in precipitation climatology are typically based on independent analysis of precipitation frequency and intensity. Here we show that where it rains more often, it also rains harder. When grouping global precipitation reanalysis data and observations from the past 40 years into regions of similar wet-day frequency, regardless of geographical separation, there is a strong correlation with wet-day intensity distributions. These wet-day-frequency regions are also more physically coherent than regions based on geographical location. We find the coherent relationship between wet-day frequency and intensity distributions is partially explained by wet-day-frequency regions having similar vertical velocity and convective available potential energy distributions, once polar regions are excluded. These represent dynamic and thermodynamic processes that indicate how conducive wet-day-frequency regions are to large-scale and convective precipitation. This suggests that the main drivers of precipitation are universal. We also show that extreme-precipitation metrics are dependent on wet-day frequency within our framework. Our results imply that wet-day frequency could be used to derive estimates of extreme-precipitation climate indices and corresponding uncertainties, these uncertainties being related to local processes.

The spatio-temporal distribution of precipitation has critical impacts on the availability of water resources<sup>1</sup>, ecosystems<sup>2</sup>, economic growth rates<sup>3</sup> and is also an important contributor to major hazards<sup>4</sup>. Precipitation distributions and how they might change due to climate change are also critical for quantifying societal and economic risks<sup>4</sup> and are of particular importance in extreme events. Wet-day intensity and wet-day-frequency metrics are also commonly used independently to evaluate the simulation of precipitation within numerical weather-prediction models, reanalyses and climate models<sup>5,6</sup>.

The exact distribution of wet-day intensity is important because the impact of extreme-precipitation events escalate drastically with increasing intensity<sup>7</sup>. A small number of heavy-rainfall days contribute disproportionately to total precipitation compared with a large number of days of light precipitation<sup>8</sup>. For example, across Global Historical Climatology Network (GHCN) weather stations, half the annual precipitation falls during the wettest 12 days of the year<sup>8</sup>. The importance of the wet tail of the distribution is also highlighted in work that examined the geographic distribution of rainfall in different intensity classes and the frequency of each class<sup>6,9</sup>.

<sup>1</sup>School of Physical and Chemical Sciences, University of Canterbury, Christchurch, New Zealand. <sup>2</sup>Gateway Antarctica, University of Canterbury, Christchurch, New Zealand. <sup>3</sup>School of Geography Environment and Earth Science, Victoria University of Wellington, Wellington, New Zealand.

<sup>4</sup>MetService, Wellington, New Zealand. ✉e-mail: [cameron.mcerlich@pg.canterbury.ac.nz](mailto:cameron.mcerlich@pg.canterbury.ac.nz)



**Fig. 1 | Relationship between wet-day frequency and intensity.** **a**, The geographic distribution of wet-day frequency above a  $1 \text{ mm d}^{-1}$  threshold. **b**, Distributions of wet-day intensity showing cumulative precipitation values for aggregated regions corresponding to each percentage of the wet-day frequency shown in **a**. Here precipitation is displayed using a logarithmic scale. **a** and **b** are derived using ERA5 total precipitation output between 1980 and 2019.

In this Article, ECMWF Reanalysis v5 (ERA5)<sup>10</sup> output is examined to determine the frequency of precipitation, which is used to group regions by their wet-day frequency. Wet-day intensity distributions are then derived over these regions. This analysis is repeated for satellite, gauge and other reanalysis datasets to determine the robustness of these relationships. The statistical relationship between the wet-day frequency and mean wet-day intensity is shown to have a strong positive correlation in this framework. A range of extreme-precipitation metrics are also shown to be strongly positively correlated to the wet-day frequency. Potential drivers for the observed relationships, such as large-scale and convective-precipitation processes, vertical motion and convective available potential energy are examined to determine the physical basis for these relationships.

## Relationship between wet-day frequency and intensity

To understand the relationship between the frequency and intensity of precipitation, we use a framework for analysing precipitation over regions with a similar wet-day frequency. Figure 1a shows the geographic pattern of the wet-day frequency across 40 years (1980–2019) of ERA5 output. This frequency is determined using a wet-day threshold of  $1 \text{ mm d}^{-1}$ , which is commonly used within the community<sup>11,12</sup>. Previous work<sup>13</sup> has examined precipitation changes under changing climate averaged over wet and dry regions. We take this further by averaging intensity-distribution data from regions with the same wet-day frequency together. This is in contrast to many studies<sup>14–17</sup> that analyse precipitation over geographical regions. Distributions of the wet-day intensity derived over consistent wet-day-frequency regions are shown in Fig. 1b and show the likelihood of particular rainfall intensities occurring over a range of wet-day frequencies.

Figure 1a shows physically interpretable structures, such as regions of high wet-day frequency over the inter-tropical convergence

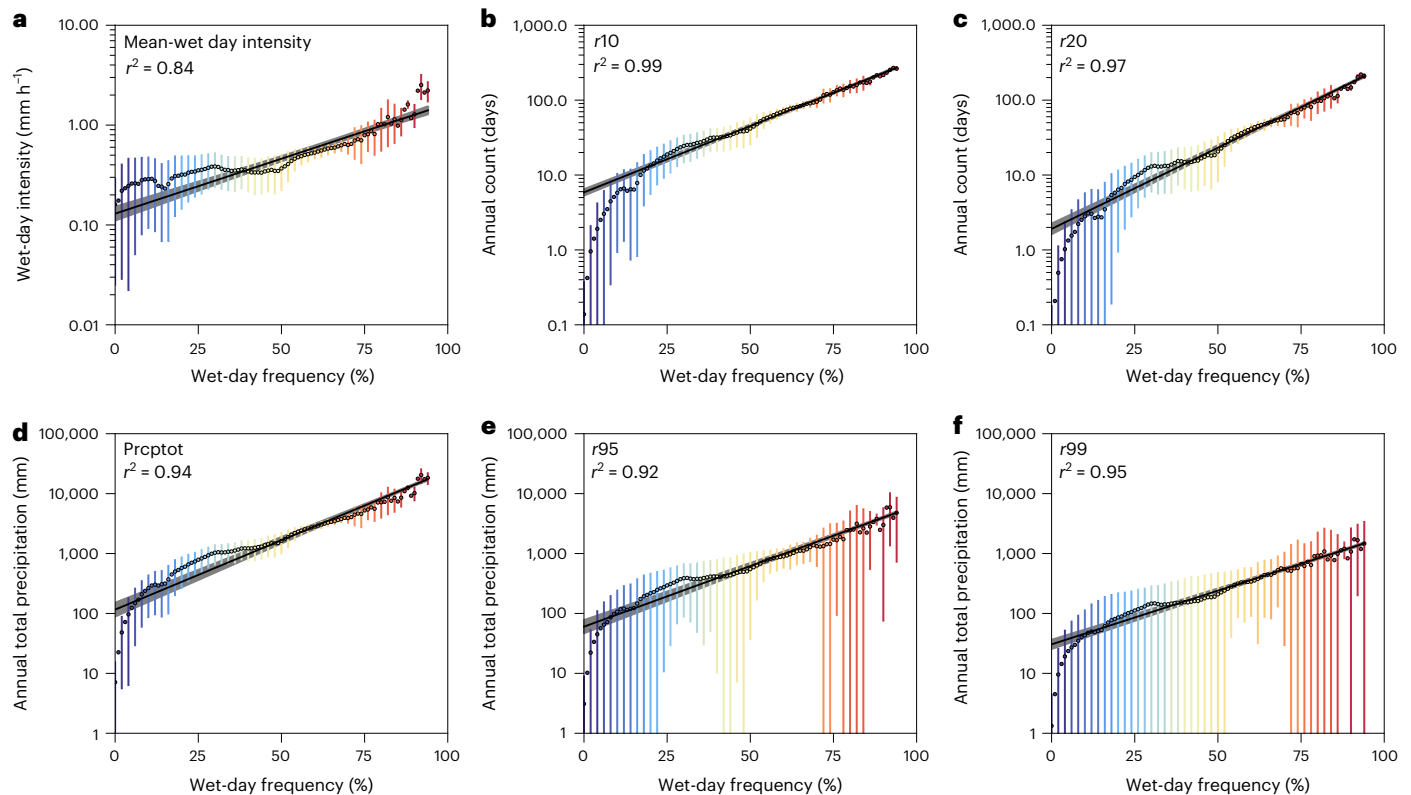
zone and regions of low-precipitation frequency over desert regions. The effects of orography on the wet-day frequency can also be seen over the Andes and Himalayas. As the wet-day frequency increases in Fig. 1b, the distribution of wet-day intensity distributions shift towards higher values. These changes in behaviour demonstrate that where it rains more often, it also rains more intensely. Examination of Fig. 1b also shows that the shape of the wet-day intensity-distribution changes with wet-day frequency. In particular, the form of the curves in Fig. 1b change from sub-exponential at low wet-day frequencies to super-exponential as the wet-day frequency increases. This highlights that regions with low wet-day frequencies tend to experience larger proportions of their precipitation at low-precipitation rates. Conversely, regions with higher wet-day frequency experience larger proportions of their precipitation at greater precipitation rates.

This result is robust to the choice of the wet-day threshold (Extended Data Figs. 1 and 2). Separating precipitation into land and ocean regions also shows our results are relatively invariant to surface type (Extended Data Fig. 3). We also show that aggregating precipitation measurements by wet-day frequency tends to produce similar results (in terms of wet-day percentiles and internal variance) to analysing precipitation across similarly sized geographical regions (Extended Data Fig. 4). This suggests that subsetting precipitation data by the wet-day frequency is as valid as analysis based on geographical region, even though it averages spatially disparate regions together.

It has already been established that there can be a substantial disparity between observational and reanalysis products when it comes to daily precipitation intensity<sup>18–20</sup>. To demonstrate that the relationships displayed in Fig. 1 are not an artefact of the ERA5 reanalysis, they also have been calculated using the Global Historical Climatology Network (GHCN) gauge dataset<sup>21</sup>, Multi-Source Weighted-Ensemble Precipitation (MSWEP) v2.8<sup>22</sup>, CPC MORPHing technique (CMORPH)<sup>23</sup>, Integrated Multi-satellite Retrievals for GPM (IMERG)<sup>24</sup> satellite datasets and Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA2)<sup>25</sup> reanalysis. Once geographic sampling differences between the ERA5 and GHCN datasets are accounted for (as explained in Methods), the two show very similar results (Extended Data Fig. 5). The same aggregation methodology applied to the other datasets (Extended Data Fig. 6) also shows results consistent with the patterns observed in ERA5. Thus, this key relationship between how often it rains and how much it rains is not a product of how precipitation is represented in the ERA5 reanalysis but represents a previously unidentified universal relationship.

## Connecting the wet-day frequency to climate indices

To quantify the connection between wet-day frequency and wet-day intensity values observable in Fig. 1b, mean wet-day precipitation intensities were determined for regions aggregated by wet-day frequency. A weighted linear least-squares regression was then calculated between the natural logarithm of the mean wet-day intensity and the wet-day frequency (Fig. 2a). Data were weighted based on the inverse square of the ratio between the standard deviation and value at each wet-day frequency. Figure 2a shows a strong positive correlation between the wet-day frequency and the logarithm of the mean wet-day intensity across the aggregated wet-day-frequency regions. To test whether this correlation was a result of averaging the precipitation intensity, the correlation between wet-day intensity and wet-day frequency was computed for each precipitation–intensity percentile (assessing the correlation horizontally across Fig. 1b). Extended Data Fig. 7 displays a range of coefficients of determination between  $r^2 = 0.65$  and  $r^2 = 0.92$ , with the weakest correlations found in the lowest-intensity percentiles. This consistent positive correlation is an indicator that the frequency of precipitation is a strong control of precipitation intensity. These relationships are not dependent on an assumption about the form of the wet-day intensity distribution, unlike previous work that



**Fig. 2 | Correlation between the wet-day intensity and the wet-day frequency and a number of climate indices.** **a**, The relation between the wet-day frequency and the mean wet-day intensity. **b–f**, The five  $r_{10}$ ,  $r_{20}$ , prcptot,  $r_{95}$  and  $r_{99}$  climate indices. Data are coloured based on the wet-day frequency as in Fig. 1a.

Error bars are represented for every second point as the mean  $\pm$  1 standard deviation. Black lines indicate the fitted polynomial of linear least-squares regression ( $n = 100$ ) while the shaded regions represent the confidence interval around the regression.  $r^2$  is the coefficient of determination.

needed to assume an exponential distribution<sup>26,27</sup>. These strong relationships provided motivation to complete similar statistical tests on extreme-precipitation indices commonly used in climate studies<sup>28</sup>.

We examine two extreme-precipitation indices based on the number of days per year above a precipitation threshold. Here  $r_{10}$  and  $r_{20}$  are defined for accumulated precipitation greater than or equal to  $10 \text{ mm d}^{-1}$  and  $20 \text{ mm d}^{-1}$ , respectively. Two metrics that quantify the annual total precipitation above the 95th and 99th percentiles,  $r_{95}$  and  $r_{99}$ , are also examined. Finally, the prcptot metric defines the annual total precipitation on wet days (precipitation greater than or equal to  $1 \text{ mm d}^{-1}$ ). Figure 2b–f shows strong positive correlations between the precipitation-extreme climate indices and wet-day frequency in each case. Looking at the standard deviation around the data, Fig. 2a–d displays small variability, except at low wet-day frequencies. This is not the case for the  $r_{95}$  and  $r_{99}$  metrics, which show notably higher variability within each wet-day frequency. This suggests that in our framework, many extreme-precipitation climate indices are heavily dependent on the wet-day frequency, but large variability in  $r_{95}$  and  $r_{99}$  reduces the strength of our result for these metrics.

One might argue that this strong correlation is due to a sampling bias introduced when averaging to 100-wet-day-frequency regions. However, corresponding analysis, which completes spatial averaging over 100 geographically coherent regions (Extended Data Fig. 8), shows a much-poorer correlation than those shown in Fig. 2b–f. There is therefore an advantage to working in the aggregated wet-day-frequency framework to identify these relationships.

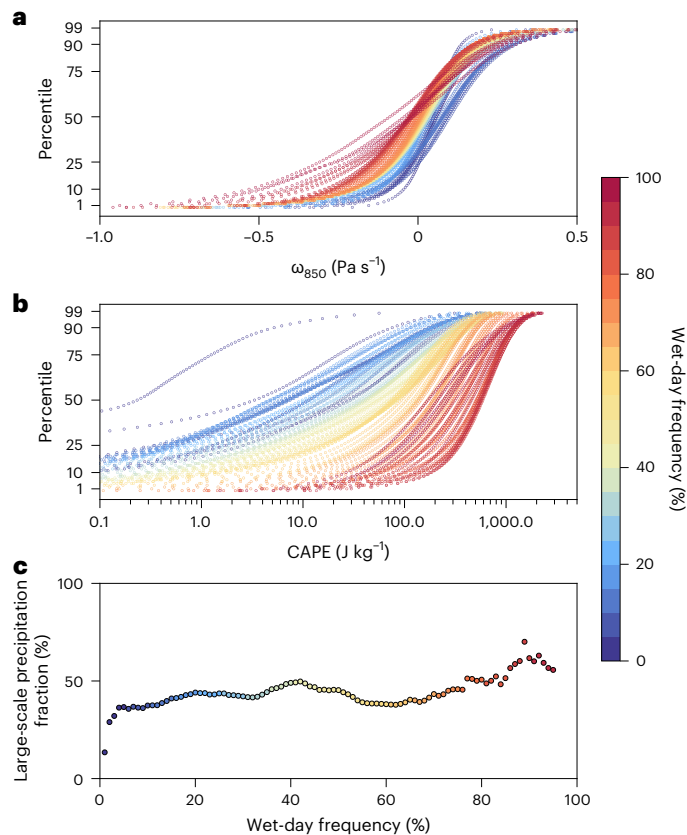
## Identifying physical drivers

Our analysis shows that the strong relationships between wet-day occurrence and other precipitation metrics is only notable when geographically disparate data are clustered based on wet-day frequency.

A myriad of processes impact precipitation including large-scale atmospheric dynamics, meso-scale convective and storm dynamics and local precipitation microphysics<sup>29</sup>. To provide physical justification for why precipitation changes across these regions coherently, the different processes that control precipitation were investigated.

Previous studies<sup>7,30–32</sup> have shown changes in vertical velocity are an important influence on precipitation with larger upward velocities relating to larger precipitation intensities. Given this, we derived distributions of ERA5 vertical velocities (in pressure coordinates, negative values identifying ascent) grouped based on wet-day frequency. The literature also identifies both dynamical and thermodynamical drivers can impact the wet-day intensity<sup>29,31</sup>. Therefore, we also analysed the convective available potential energy (CAPE) metric<sup>33,34</sup> to examine how the vertical thermodynamic structure of the atmosphere might impact the observed relationship. Figure 3a shows distributions of ERA5 vertical velocity at the 850 hPa level, and Fig. 3b shows ERA5 CAPE grouped based on the wet-day-frequency values. To connect back to precipitation processes, the amount of precipitation associated with both large-scale and convective precipitation in the ERA5 dataset was also determined for wet-day frequency. As previous work has identified errors due to detection of snowfall over high latitudes<sup>35–37</sup>, areas polewards of  $60^\circ$  latitude have been removed from this analysis. For latitude-restricted wet-day intensity distributions, Fig. 3c shows the amount of large-scale precipitation as a fraction of the total precipitation.

Vertical velocity distributions in Fig. 3a show that regions with larger wet-day frequencies and therefore higher precipitation–intensity rates relate to regions with higher ascent rates and have distributions skewed towards ascent. A correlation between the vertical velocity and wet-day frequency (Extended Data Fig. 9) displays a coefficient of determination of  $r^2 = 0.27$ , meaning that changes in vertical velocity potentially account for 27% of the variability in the wet-day frequency.



**Fig. 3 | Physical drivers behind the relation between wet-day frequency and intensity.** **a**, Distributions for vertical velocity at 850 hPa for all days grouped based on wet-day frequency. **b**, Distributions for CAPE for all days grouped based on wet-day frequency. **c**, The amount of large-scale precipitation as a fraction of the total precipitation. At each point, the remaining fraction corresponds to convective precipitation.

Figure 3b shows that regions with the highest frequency and therefore intensity also occur in regions that have the greatest CAPE. A similar spatial analysis (Extended Data Fig. 9) displays a coefficient of determination of  $r^2 = 0.23$ . These distributions of higher ascent rates and higher values of CAPE for regions with higher wet-day frequency can both be interpreted as reflecting higher occurrence of environments conducive to precipitation. Furthermore, Fig. 3c shows a shift in the fraction of large-scale precipitation as the wet-day frequency increases. Excluding polar regions, convective precipitation dominates at places with low wet-day frequency while large-scale precipitation dominates elsewhere.

Figure 3 indicates that most extreme-precipitation events are dominated by large-scale processes, such as cyclones, fronts and the slow ascent of air in synoptic systems. This is coherent, because while convective-precipitation events are related to high rainfall rates, they generally have a short duration. Conversely, large-scale precipitation events have lower rainfall rates occurring over a larger portion of the day<sup>38,39</sup>. Thus, when precipitation intensities are averaged over a day, large-scale precipitation will produce higher precipitation intensities. Our result that the distributions of vertical velocity and CAPE change considerably as a function of wet-day frequency therefore identifies that both these metrics are important drivers of large-scale precipitation-generating processes. This provides a physical explanation of the relationship between wet-day frequency and the wet-day intensity distribution observed in Fig. 1b.

## Discussion and outlook

Our work focuses on analysing precipitation over regions of similar wet-day frequency, effectively aggregating wet and dry regions in a

consistent way globally. The central result of this study is that there is a strong relationship between the frequency of wet days and the intensity distribution of precipitation on those days across disparate regions of the globe. We have shown wet-day-frequency regions have similar variability to geographical regions (Extended Data Fig. 4) and correspond to places with distributions of vertical velocity and CAPE that are conducive to precipitation. This suggests that using wet-day frequency to analyse precipitation is more physically coherent than using geographical regions, as our framework displays a connection between precipitation and its dynamic and thermodynamic drivers. This previously unquantified relationship provides a framework for understanding precipitation and how it might change in a warming world, which we believe could have important ramifications for model evaluation and fundamental understanding of the relative importance of drivers of precipitation in different regions. Our results imply that fundamentally, the processes that drive precipitation are universal.

Our framework provides an aggregated overview of the properties of precipitation and important physical drivers. When assessing local distributions of precipitation, other factors and processes would cause uncertainties in any estimates. Extended Data Fig. 9 shows that CAPE and vertical velocity account for approximately 50% of the variance in the wet-day frequency. However, we would expect local factors, such as orographic forcing to dominate over other regions and cause deviations from the distributions of wet-day intensity seen on Fig. 1b. This provides an interesting possibility to look at anomalies from the norm to identify regionalized drivers. We also note that the temporal resolution of the available datasets used will potentially impact the results of aggregation, given that different precipitation types (convective and stratiform) scale differently<sup>38,40</sup>.

While some studies identify the importance of the combination of wet-day frequency and wet-day intensity, they have not explored fully the relationship between these factors. An inspection of previous literature does show that this relationship is previously implied<sup>8,17,19,27,41</sup>. For example, one study<sup>27</sup> used principal component analysis to derive the most important relations between observed precipitation and the two precipitation metrics from thousands of rain gauges. However, the intensity–frequency relationship would have been very difficult to identify given the limited geographic coverage of the gauge network globally. Another study<sup>41</sup> looking at the climatological characteristics of precipitation identifies a link between the geographical pattern of the most common precipitation intensity and the geographical pattern of the precipitation–frequency peak. While those authors show the existence of a frequency–intensity link, they do not explore the link between frequency and the rainfall–intensity distribution detailed in this study or examine the implications of their findings.

One common set of precipitation metrics used to compare across gauge-based, satellite, reanalyses and climate-model outputs are extreme climate indices. Their standardization by the Expert Team on Climate Change Detection and Indices (ETCCDI)<sup>42</sup> allows for comparison across various timescales and resolutions, even for the gauge-based datasets that suffer from sampling biases due to sparse coverage and short observational records. Our results show that when averaging using wet-day frequency, there is a strong positive correlation between the wet-day frequency and many of the precipitation-related extreme climate indices. While variance within these mean statistics exists, especially for the  $r_{95}$  and  $r_{99}$  indices, the appearance of such a strong correlation when framing the analysis in a different way highlights that many of these climate metrics are not independent. However, most studies of precipitation-related climate indices<sup>43–45</sup> assess them as independent metrics and ultimately could be impacted by our result.

Effectively, results in this study suggest that getting the wet-day frequency correct can provide understanding of some of the more complex climate indices. This is useful in the case of extreme-precipitation-related indices (for example,  $r_{95}$ ,  $r_{99}$ ) as long-term datasets are required to get an accurate representation of

the tail of the precipitation distribution. In addition, the relationships between wet-day frequency and intensity could be used to help further identify the relative importance of physical drivers of precipitation and their impact on the precipitation–intensity distribution. Research efforts have created global analyses fitting generalized extreme-value parameters to precipitation–intensity distributions geographically<sup>46</sup> but have not offered opportunities for any physical insight given their statistical nature.

Past research suggests that many precipitation products should be used in studies, as none provide a single best estimate of precipitation<sup>47,48</sup>. While six different datasets were used in this study to verify our findings, a limitation of this study remains a lack of intercomparison. Forthcoming work will focus on evaluating a larger number of commonly used precipitation datasets, reanalyses and Coupled Model Intercomparison Project Phase 6 (CMIP6) climate-model outputs to more completely assess how the precipitation frequency–intensity relation, climate indices and the physical processes that drive them are represented across these products.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41561-023-01177-4>.

## References

- Konapala, G., Mishra, A. K., Wada, Y. & Mann, M. E. Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation. *Nat. Commun.* **11**, 3044 (2020).
- Seddon, A. W. R., Macias-Fauria, M., Long, P. R., Benz, D. & Willis, K. J. Sensitivity of global terrestrial ecosystems to climate variability. *Nature* **531**, 229–232 (2016).
- Kotz, M., Levermann, A. & Wenz, L. The effect of rainfall changes on economic production. *Nature* **601**, 223–227 (2022).
- Morss, R. E., Wilhelmi, O. V., Meehl, G. A. & Dilling, L. Improving societal outcomes of extreme weather in a changing climate: an integrated perspective. *Ann. Rev. Environ. Resour.* **36**, 1–25 (2011).
- Bosilovich, M. G., Chen, J., Robertson, F. R. & Adler, R. F. Evaluation of global precipitation in reanalyses. *J. Appl. Meteorol. Climatol.* **47**, 2279–2299 (2008).
- Lau, W. K.-M., Wu, H.-T. & Kim, K.-M. A canonical response of precipitation characteristics to global warming from CMIP5 models. *Geophys. Res. Lett.* **40**, 3163–3169 (2013).
- Dai, P. & Nie, J. What controls the interannual variability of extreme precipitation? *Geophys. Res. Lett.* **48**, e2021GL095503 (2021).
- Pendergrass, A. G. & Knutti, R. The uneven nature of daily precipitation and its change. *Geophys. Res. Lett.* **45**, 11980–11988 (2018).
- Kay, J. E. et al. Scale-aware and definition-aware evaluation of modeled near-surface precipitation frequency using CloudSat observations. *J. Geophys. Res. D: Atmos.* **123**, 4294–4309 (2018).
- Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
- Polade, S. D., Pierce, D. W., Cayan, D. R., Gershunov, A. & Dettinger, M. D. The key role of dry days in changing regional climate and precipitation regimes. *Sci. Rep.* **4**, 4364 (2014).
- Schar, C. et al. Percentile indices for assessing changes in heavy precipitation events. *Climatic Change* **137**, 201–216 (2016).
- Donat, M. G., Lowry, A. L., Alexander, L. V., O’Gorman, P. A. & Maher, N. More extreme precipitation in the world’s dry and wet regions. *Nat. Clim. Change* **6**, 508–513 (2016).
- Rajah, K. et al. Changes to the temporal distribution of daily precipitation. *Geophys. Res. Lett.* **41**, 8887–8894 (2014).
- Utsumi, N., Kim, H., Kanae, S. & Oki, T. Relative contributions of weather systems to mean and extreme global precipitation. *J. Geophys. Res. D: Atmos.* **122**, 152–167 (2017).
- Papalexiou, S. M. & Montanari, A. Global and regional increase of precipitation extremes under global warming. *Water Resour. Res.* **55**, 4901–4914 (2019).
- Myhre, G. et al. Frequency of extreme precipitation increases extensively with event rareness under global warming. *Sci. Rep.* **9**, 16063 (2019).
- Herold, N., Alexander, L. V., Donat, M. G., Contractor, S. & Becker, A. How much does it rain over land?. *Geophys. Res. Lett.* **43**, 341–348 (2016).
- Sun, Q. et al. A review of global precipitation data sets: data sources, estimation, and intercomparisons. *Rev. Geophys.* **56**, 79–107 (2018).
- Beck, H. E. et al. Daily evaluation of 26 precipitation datasets using stage-IV gauge-radar data for the conus. *Hydrol. Earth Syst. Sci.* **23**, 207–224 (2019).
- Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E. & Houston, T. G. An overview of the global historical climatology network-daily database. *J. Atmos. Oceanic Technol.* **29**, 897–910 (2012).
- Beck, H. E. et al. MSWEP v2 global 3-hourly 0.1° precipitation: methodology and quantitative assessment. *Bull. Am. Meteorol. Soc.* **100**, 473–500 (2019).
- Xie, P. et al. Reprocessed, bias-corrected CMORPH global high-resolution precipitation estimates from 1998. *J. Hydrometeorol.* **18**, 1617–1641 (2017).
- Huffman, G. J., Bolvin, D. T., Nelkin, E. J. & Tan, J. Integrated multi-satellite retrievals for GPM (IMERG) technical documentation (NASA/GSFC, 2019).
- Gelaro, R. et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* **30**, 5419–5454 (2017).
- Benestad, R. E., Nychka, D. & Mearns, L. O. Spatially and temporally consistent prediction of heavy precipitation from mean values. *Nat. Clim. Change* **2**, 544–547 (2012).
- Benestad, R. E., Nychka, D. & Mearns, L. O. Specification of wet-day daily rainfall quantiles from the mean value. *Tellus A: Dyn. Meteorol. Oceanogr.* **64**, 14981 (2012).
- Zhang, X. et al. Indices for monitoring changes in extremes based on daily temperature and precipitation data. *WIREs Clim. Change* **2**, 851–870 (2011).
- Lenderink, G. & van Meijgaard, E. Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nat. Geosci.* **1**, 511–514 (2008).
- Emori, S. & Brown, S. J. Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate. *Geophys. Res. Lett.* **32**, L17706 (2005).
- Pfahl, S., O’Gorman, P. A. & Fischer, E. M. Understanding the regional pattern of projected future changes in extreme precipitation. *Nat. Clim. Change* **7**, 423–427 (2017).
- Chen, G., Ming, Y., Singer, N. D. & Lu, J. Testing the Clausius-Clapeyron constraint on the aerosol-induced changes in mean and extreme precipitation. *Geophys. Res. Lett.* **38**, L04807 (2011).
- Emori, S., Hasegawa, A., Suzuki, T. & Dairaku, K. Validation, parameterization dependence, and future projection of daily precipitation simulated with a high-resolution atmospheric GCM. *Geophys. Res. Lett.* **32**, L06708 (2005).
- Myoung, B. & Nielsen-Gammon, J. W. Sensitivity of monthly convective precipitation to environmental conditions. *J. Clim.* **23**, 166–188 (2010).

35. Kulie, M. S., Bennartz, R., Greenwald, T. J., Chen, Y. & Weng, F. Uncertainties in microwave properties of frozen precipitation: implications for remote sensing and data assimilation. *J. Atmos. Sci.* **67**, 3471–3487 (2010).
36. Levizzani, V., Laviola, S. & Cattani, E. Detection and measurement of snowfall from space. *Remote Sens.* **3**, 145–166 (2011).
37. Beck, H. E. et al. PPDIST, global 0.1° daily and 3-hourly precipitation probability distribution climatologies for 1979–2018. *Sci. Data* **7**, 302 (2020).
38. Berg, P. & Haerter, J. Unexpected increase in precipitation intensity with temperature—a result of mixing of precipitation types. *Atmos. Res.* **119**, 56–61 (2013).
39. Berg, P., Moseley, C. & Haerter, J. O. Strong increase in convective precipitation in response to higher temperatures. *Nat. Geosci.* **6**, 181–185 (2013).
40. Eggert, B., Berg, P., Haerter, J. O., Jacob, D. & Moseley, C. Temporal and spatial scaling impacts on extreme precipitation. *Atmos. Chem. Phys.* **15**, 5957–5971 (2015).
41. Pendergrass, A. G. & Deser, C. Climatological characteristics of typical daily precipitation. *J. Clim.* **30**, 5985–6003 (2017).
42. Peterson, T. & Manton, M. Monitoring changes in climate extremes: a tale of international collaboration. *Bull. Am. Meteorol. Soc.* **89**, 1266–1271 (2008).
43. Brown, P. J., Bradley, R. S. & Keimig, F. T. Changes in extreme climate indices for the northeastern United States, 1870–2005. *J. Clim.* **23**, 6555–6572 (2010).
44. Deniz, A., Toros, H. & Incecik, S. Spatial variations of climate indices in Turkey. *Int. J. Climatol.* **31**, 394–403 (2011).
45. Duan, W. et al. Changes of precipitation amounts and extremes over Japan between 1901 and 2012 and their connection to climate indices. *Clim. Dyn.* **45**, 2273–2292 (2015).
46. Courty, L. G., Wilby, R. L., Hillier, J. K. & Slater, L. J. Intensity-duration-frequency curves at the global scale. *Environ. Res. Lett.* **14**, 084045 (2019).
47. Herold, N., Behrangi, A. & Alexander, L. V. Large uncertainties in observed daily precipitation extremes over land. *J. Geophys. Res. D: Atmos.* **122**, 668–681 (2017).
48. Bador, M., Alexander, L. V., Contractor, S. & Roca, R. Diverse estimates of annual maxima daily precipitation in 22 state-of-the-art quasi-global land observation datasets. *Environ. Res. Lett.* **15**, 035005 (2020).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

## Methods

### ERAS

ECMWF Reanalysis v5 (ERA5) is a climate reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), combining observational measurements with model output to provide a consistent gridded product for recent climate<sup>10</sup>. ERA5 total, convective and large-scale precipitation components, pressure vertical velocity and convective available potential energy output between 1980 and 2019 are used in this study. ERA5 total precipitation consists of both rain and snow at Earth's surface and is the sum of the ERA5 large-scale and convective-precipitation fields, which are also used in this study. This variable does not include precipitation that evaporates in the atmosphere before it lands at Earth's surface or fog and dew<sup>10</sup>. All the ERA5 outputs used in this study are available on a  $0.25^\circ \times 0.25^\circ$  spatial grid and at an hourly temporal resolution, which was sampled at a three hourly interval. In the underlying ERA5 model, large-scale precipitation is produced from the large-scale (stratiform) cloud microphysical processes and denotes both rain and snow, while convective precipitation is produced by the convective parameterization scheme and relates to a single-grid box column when convection is diagnosed. Convective available potential energy is an metric of the stability of the atmosphere and can be used to assess the potential for the development of convection. Using pressure as the vertical coordinate, negative vertical velocities represent regions of ascent and positive vertical velocities represent descending air. Regions of ascent are important in the formation of precipitation<sup>30</sup>.

### GHCN station data

In addition to the ERA5 reanalysis dataset, gauge data included in the Global Historical Climatology Network<sup>21</sup> (GHCN) were also analysed to investigate the robustness of ERA5 results. Gauge data must be quality controlled to ensure data artefacts are removed. In this study, only gauge stations that operated from 1980 to 2019 were considered. We also used the GHCN quality flag system to remove potentially low-quality data and we remove stations above  $60^\circ$  N or below  $60^\circ$  S to avoid regions with a risk of frequent snow in the gauges. These stations were then aggregated into a global grid with the same resolution as the ERA5 data to reduce sampling biases due to the uneven distribution of stations. There will, however, still be strong sampling biases in the gauge data due to the regions with large areas of no station coverage, such as Asia and Africa, and the complete absence of data over the oceans. Because of these biases, when this data are compared to ERA5 (Extended Data Fig. 5), only ERA5 grid cells that have a corresponding GHCN measurement are examined.

### MSWEP

We also used the Multi-Source Weighted-Ensemble Precipitation (MSWEP) v2.8 dataset<sup>22</sup>; this uses a combination of satellite, gauge and reanalysis datasets between 1980 and 2019. The MSWEP v2.8 combines IMERG data<sup>24</sup> with other datasets (including ERA5) to produce a consistent dataset over both the ocean and land at daily resolution on a  $0.1^\circ \times 0.1^\circ$  grid. Given the presence of gauge and satellite data in MSWEP v2.8 and coverage over the entire 40-year period used for ERA5, we believe MSWEP is a valuable product for assessing the robustness of our conclusion between datasets.

### CMORPH

We also used the CPC Morphing technique (CMORPH) high-resolution global satellite precipitation dataset<sup>23</sup>. Satellite data are merged into a precipitation product that is then bias corrected using CPC daily gauge data over land and Global Precipitation Climatology Project (GPCP) gauge data over the ocean. We use data between 2001 and 2019 at daily resolution on a  $0.25^\circ \times 0.25^\circ$  grid. Note that CMORPH data have only spatial coverage between  $60^\circ$  N and  $60^\circ$  S.

### IMERG

We also used the Integrated Multi-satellite Retrievals for GPM (IMERG) V05 precipitation dataset<sup>24</sup>. IMERG combines information from multiple satellites present in the core GPM constellation. The IMERG dataset aims at deriving precipitation by intercalibrating and merging 'all' the available microwave sensors, along with microwave-calibrated infrared satellite estimates and precipitation gauge analyses. V05 of the IMERG product has a  $0.1^\circ \times 0.1^\circ$  spatial resolution, and daily data were used between 2001 and 2019. Due to an observed discontinuity in the wet-day frequency at  $60^\circ$  N and  $60^\circ$  S, polar regions were excluded.

### MERRA2

Finally, we use the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2) dataset<sup>25</sup>. MERRA2 is a reanalysis dataset that combines observations data with climate-model output to provide a consistent gridded product. We use data between 1980 and 2012 at a daily resolution on a  $0.625^\circ \times 0.5^\circ$  spatial grid.

### Wet-day frequency

The geographical distribution of wet-day frequency was derived for each grid cell of the observational products irrespective of their resolution. A period was marked as precipitating if the accumulated precipitation was greater than a 1 mm-per-day threshold and non-precipitating otherwise; this threshold is commonly used within the community<sup>11,12</sup> and is also used in a number of extreme-precipitation indices examined<sup>28</sup>. Products with sub-daily resolution were identified as precipitating if their accumulation over the period since the last measurement corresponded to an equivalent accumulation of 1 mm over a full day.

### Wet-day intensity

Precipitation data were then aggregated into regions where the wet-day frequency was the same. For these regions, precipitation was grouped together to derive cumulative precipitation–intensity distributions. As for the wet-day frequency, a 1 mm d<sup>-1</sup> equivalent threshold was applied such that only wet days were considered. This process averages spatially disparate regions together, but this averaging produces a range of distributions that are similar to those identified when calculating regional or zonal means (Extended Data Fig. 4), which is common in the literature<sup>14–17</sup>. We also demonstrate later that key properties that help define the likelihood and intensity of precipitation, such as vertical velocity<sup>30–32</sup> and convective available potential energy<sup>33,34</sup> are coherent across these regions.

### Climate indices

Commonly used extreme-precipitation indices based on the recommendations of the Expert Team on Climate Change Detection and Indices (ETCCDI) are also examined in this study<sup>28,42</sup>. Five indices (prcptot,  $r10$ ,  $r20$ ,  $r95$  and  $r99$ ) were chosen, with a particular focus on indices that represent the extreme precipitation.  $r10$  and  $r20$  are defined as the annual count of days where precipitation is greater than or equal to 10 mm d<sup>-1</sup> and 20 mm d<sup>-1</sup>, respectively. prcptot is defined as the annual total precipitation on wet days where precipitation is greater than or equal to 1 mm d<sup>-1</sup>.  $r95$  and  $r99$  are defined as the annual total precipitation that occurs above the 95th and 99th percentiles of the precipitation distribution. Climate indices were determined globally across the 1980–2019 period for each ERA5 grid cell and then averaged when the wet-day-frequency clustering was applied.

### Data availability

The ERA5 reanalysis products were obtained from the Copernicus Climate Data Store (<https://cds.climate.copernicus.eu/>). GHCN station data are available from National Oceanic and Atmospheric Administration National Centers for Environmental Information website (<https://www.ncei.noaa.gov/data/global-historical-climatology-network-daily>). MSWEP data are available through Google Drive with the access process

detailed on their website (<http://www.gloh2o.org/mswep/>). CMORPH data are available from NOAA National Centers for Environmental Information website (<https://doi.org/10.25921/w9va-q159>). IMERG data are available from NASA. MERRA2 data are available at the Goddard Earth Sciences (GES) Data and Information Services Center (DISC) (<https://disc.gsfc.nasa.gov/datasets?project=MERRA-2>). Data used to visualize the figures are available at <https://doi.org/10.5281/zenodo.6929755>.

### Code availability

Custom code generated in this study is available at <https://doi.org/10.5281/zenodo.6929755>.

### Acknowledgements

We thank the European Centre for Medium-Range Weather Forecasts for generating the ERA5 dataset, the MSWEP team for the development of the MSWEP dataset and the National Centers for Environmental Information for curating the GHCN dataset. We would like to acknowledge funding from the Whakahura: Extreme Events and the Emergence of Climate Change Ministry of Business, Innovation and Employment endeavour programme (grant number RTVU1906).

### Author contributions

C.M. conceptualized the research, conducted analysis and wrote the original manuscript draft. A.M. supervised the study, helped

conceptualize the work, conducted analysis and contributed to the production and review of the manuscript. A.S. conducted analysis and contributed to the production and review of the manuscript. G.V. conducted analysis and contributed to the review of the manuscript. J.R. and S.R. supervised the study and contributed to the review of the manuscript.

### Competing interests

The authors declare no competing interests

### Additional information

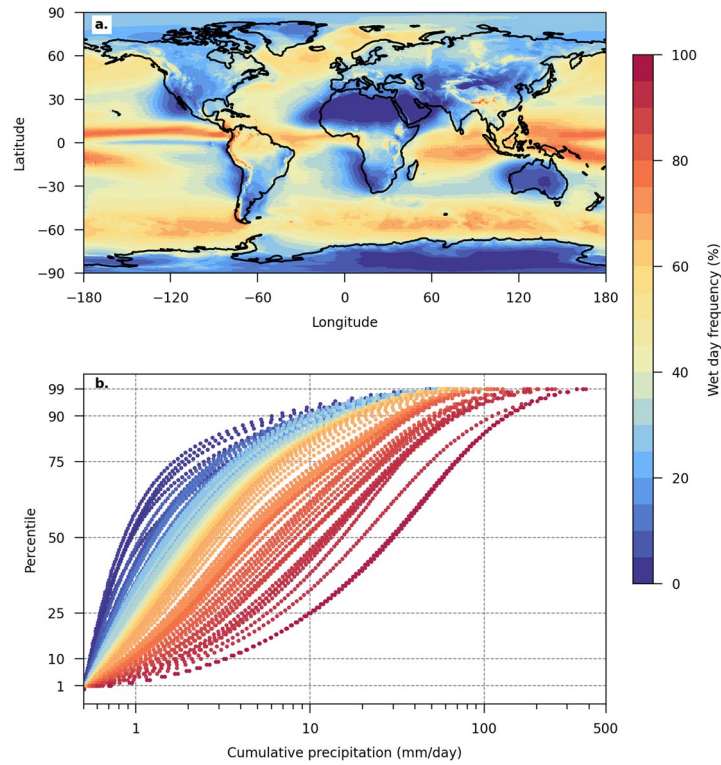
**Extended data** Extended data is available for this paper at <https://doi.org/10.1038/s41561-023-01177-4>.

**Correspondence and requests for materials** should be addressed to Cameron McErlich.

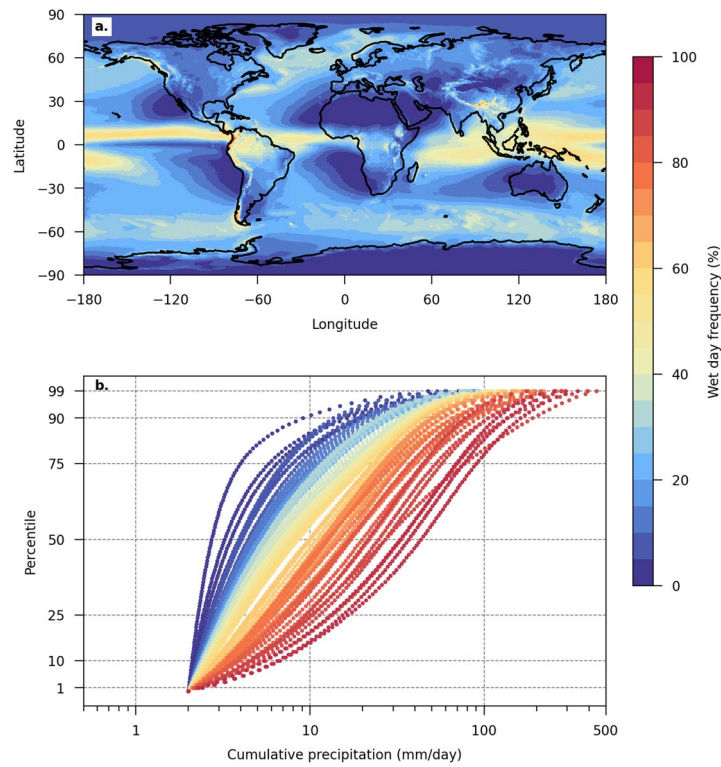
**Peer review information** *Nature Geoscience* thanks Francesco Marra and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary handling editor: Tom Richardson, in collaboration with the *Nature Geoscience* team.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

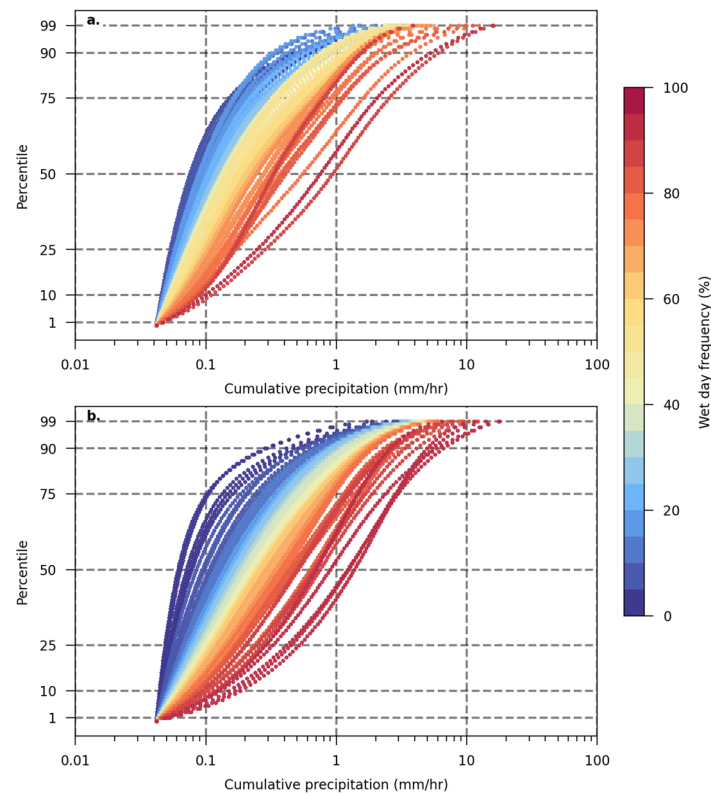




**Extended Data Fig. 1 | Changing the wet-day threshold to 0.5 mm.** In order to test the robustness of the result in Fig. 1, the wet-day threshold was changed from 1 mm/day to 0.5 mm/day a) the geographic distribution of wet-day frequency above a 0.5 mm/day threshold and b) distributions of wet-day intensity for regions corresponding to each percentile of the wet-day frequency.

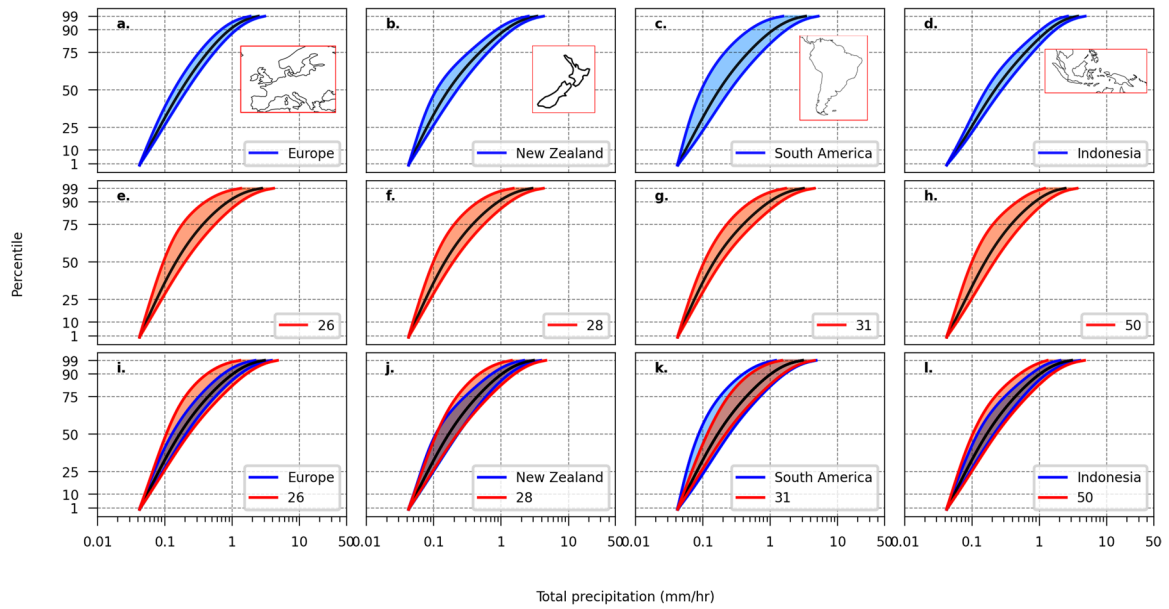


**Extended Data Fig. 2 | Changing the wet-day threshold to 2 mm.** In order to test the robustness of the result in Fig. 1, the wet-day threshold was changed from 1 mm/day to 2 mm/day a) the geographic distribution of wet-day frequency above a 2 mm/day threshold and b) distributions of wet-day intensity for regions corresponding to each percentile of the wet-day frequency.



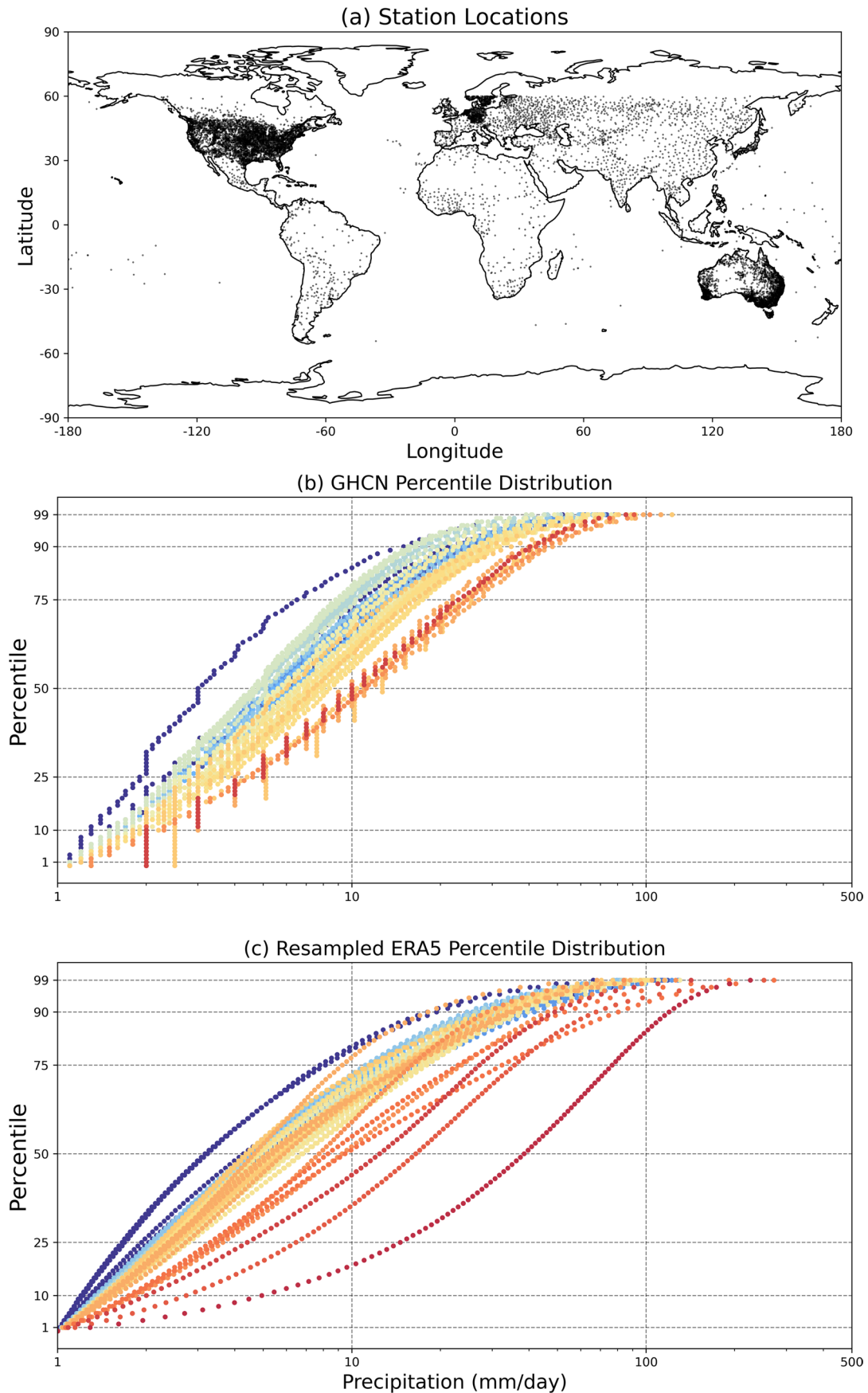
**Extended Data Fig. 3 | Wet-day intensity distributions separated over ocean and land.** Further analysis was done to assess whether aggregating regions of a similar wet-day frequency over only ocean or land would impact the results seen

on Fig. 1 a) distributions of wet-day intensity derived using ERA5 over ocean for regions corresponding to each percentage of the wet-day frequency. b) same as a) but over land.



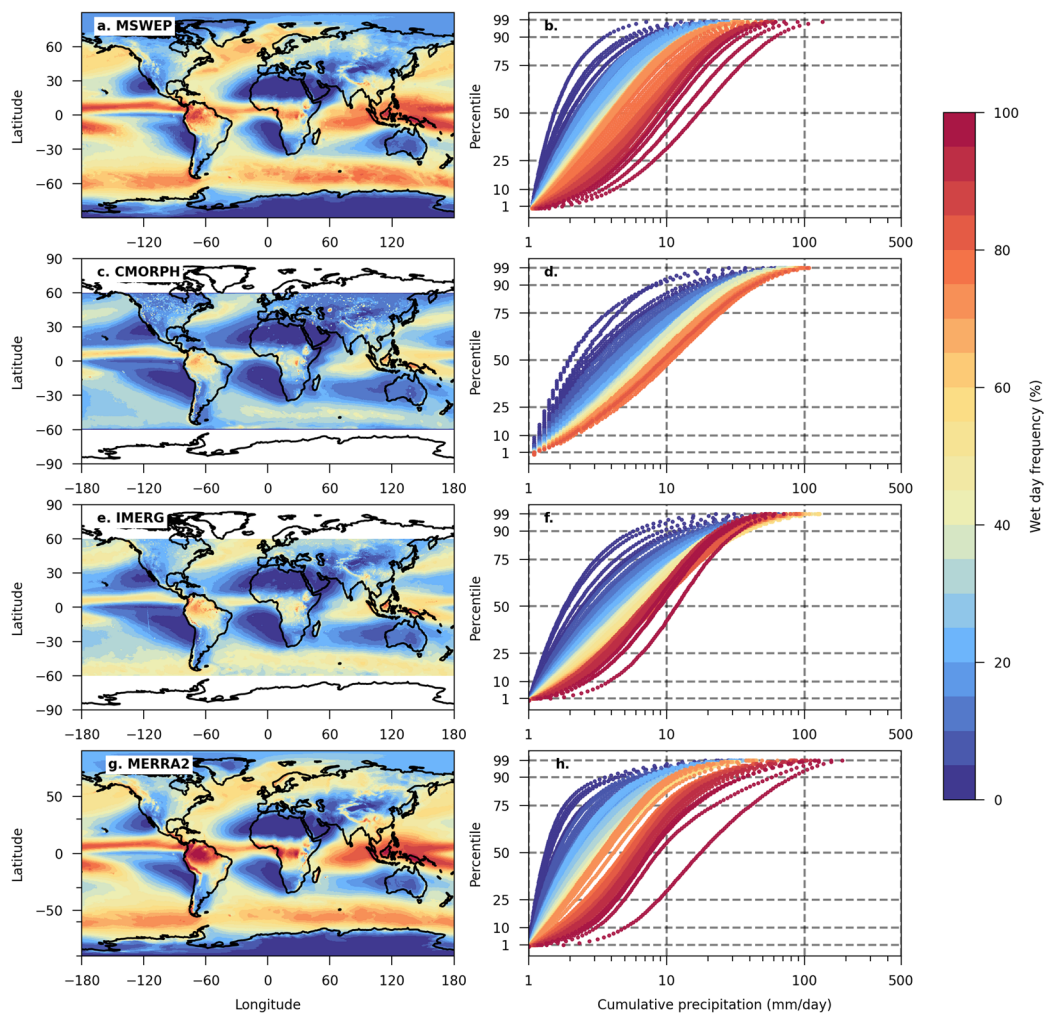
**Extended Data Fig. 4 | Variability of wet-day frequency regions against geographic regions.** Variability of precipitation intensity across (a–d) geographical regions and (e–h) specific wet-day frequencies. (i–l) differences between the two, with means adjusted to match and standard deviation converted based on a percentage of the unadjusted mean. Black lines indicate

the mean while the coloured lines and shaded region indicates the standard deviation. The bounds of chosen geographic regions (inset on a–d) are selected to be comparable to the area covered by a given wet-day frequency, and the wet-day frequencies are chosen as the mean frequency within each geographical region.



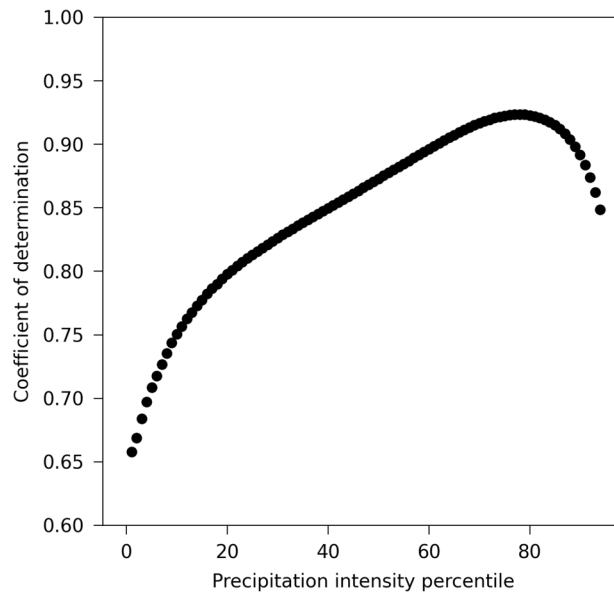
**Extended Data Fig. 5 | Wet-day percentiles of GHCN station data.** To explore the robustness of the precipitation distributions derived from the ERA5 data we also look at the GHCN rain gauge data a) the locations of the grid cells for which the GHCN has valid measurements and as such make up the data used in panels

b) and c) b) distributions of wet-day intensity for each percentage of the wet-day frequency from the GHCN. c) a recreation of b) exclusively using ERA-5 data from the same grid points as the included GHCN analysis.

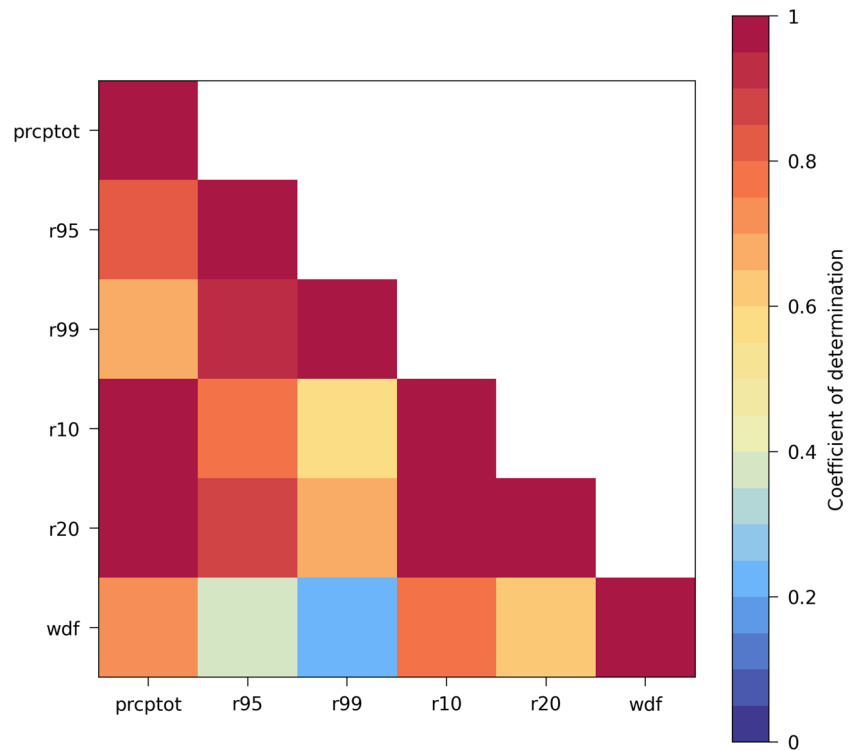


**Extended Data Fig. 6 | Wet-day percentiles of MSWEP, CMORPH, IMERG and MERRA2 data.** a) the geographic distribution of wet-day frequency above a 1 mm/day threshold between 1980–2019 derived using the MSWEP dataset. b) distributions of wet-day intensity for regions corresponding to each percentage of the wet-day frequency derived using the MSWEP dataset. c) the same as a) but

for CMORPH between 2001–2019 and 60N/S. d) same as b) but for CMORPH. e) the same as a) but for IMERG between 2001–2019 and 60N/S. f) same as b) but for IMERG g) same as a) but for MERRA2 between 1980–2012 h) same as b) but for MERRA2.



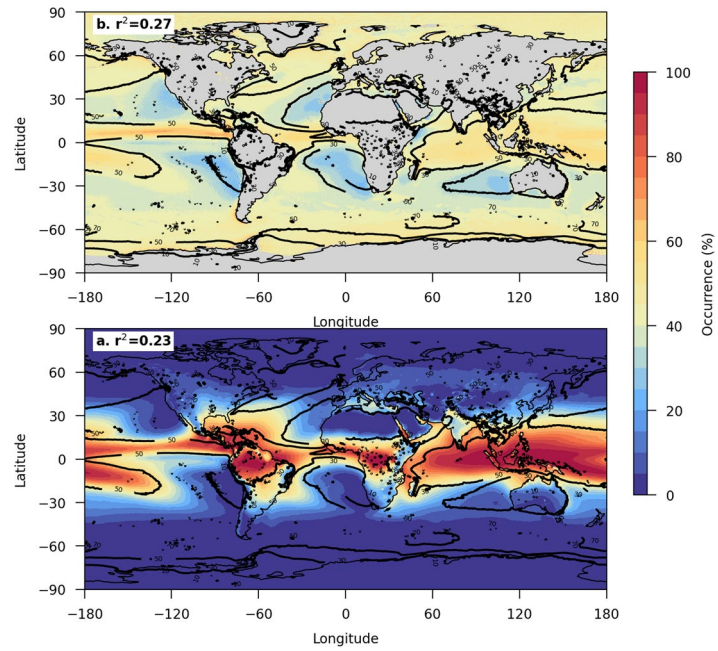
**Extended Data Fig. 7 | Wet-day frequency correlation at a given wet-day intensity percentile.** The coefficient of determination between the wet-day frequency and wet-day intensity values from Fig. 1b at a given intensity percentile. This assesses the correlation horizontally across Fig. 1b.



**Extended Data Fig. 8 | Spatial correlation of precipitation related climate indices.** Matrix comparing the spatial correlation wet-day frequency with a number of precipitation-related climate metrics. Because analysis in Fig. 2 works

out a correlation across 100 regions of wet-day frequency, the ERA-5 data was resampled to a spatial grid with 100 regions (a ten by ten grid across the globe) to explore any potential bias associated with averaging.





**Extended Data Fig. 9 | Spatial correlation of CAPE and vertical velocity.** Frequency of occurrence of a) ERA vertical velocity days where ascent occurs and b) ERA5 CAPE output. Both a) and b) are overlaid with the wet-day frequency from Fig. 1a (black contours).