ARRCC
Asia Regional Resilience to a Changing Climate

Report

Understanding and quantifying extreme precipitation events in South Asia

Part III – Observational datasets for the assessment of present day monsoon-season rainfall extremes in Nepal

CARISSA Activity 4: Climate services for the water and hydropower sectors in South Asia

March 2022

Delivery Partners:

UKaid from the British people

Met Office

ICIMOD

This study has been produced as part of the UK Aid funded Asia Regional Resilience to a Changing Climate (ARRCC) programme which is being delivered in partnership with the Met Office, the International Centre for Integrated Mountain Development (ICIMOD) and the Nepal Development Research Institute (NDRI). The study was conducted under Work Package 3: Climate Analysis for Risk Information & Services in South Asia (CARISSA) and was produced by Met Office.

Lead Author

Hamish Steptoe, Met Office

Reviewers and Contributors

Rosie Oakes, Met Office

Tim Mitchell, Met Office

Joseph Daron, Met Office

Andrew Cottrell, Met Office

The author also acknowledges valuable discussion and contributions from:

Santosh Nepal, ICIMOD

Mandira Shrestha, ICIMOD

Divas Basnyat, NDRI

Contents

1. Introduction

This report constitutes Part III in a series of reports examining extreme rainfall events in South Asia, that contributes to the CARISSA (Climate Analysis for Risk Information and Services in South Asia) Work Package of the ARRCC (Asia Regional Resilience to a Changing Climate) programme workstream 4, focused on developing climate services for the water and hydropower sectors.

This report focuses on the use of observational datasets for assessing extreme precipitation in Nepal. Further references to precipitation extremes specifically relate to high rainfall accumulations that occur during the Nepal monsoon season, June – September. Within the context of hydropower, damage to infrastructure as a result of precipitation extremes is most often associated with rainfall accumulation occurring in the monsoon season. The remainder of Section 1 summarises the key drivers of precipitation extremes over Nepal, Section 2 reviews a selection of common precipitation datasets and Section 3 compares of subset of 5 datasets from Section 2 with the Met Office seasonal forecast GloSea5. We conclude with some directions for future work in Section 4.

1.1 Drivers of precipitation extremes in Nepal

The South Asian summer monsoon is affected by the El Niño Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD). The El Niño phase of ENSO acts to suppress monsoon rainfall and the La Niña phase acts to enhance rainfall (e.g. Pant & Parthasarathy, 1981; Sikka, 1980). For IOD, in general the positive phase correlates with increased monsoon rainfall and the negative phase with decreased rainfall (e.g. Ashok et al., 2001; Behera et al., 1999). There is also significant intra-seasonal variability in South Asian monsoon rainfall with periods of active and break conditions in the monsoon rains. A more detailed summary of extreme precipitation events in South Asia and their climate drivers can be found in Richardson (2021) and Stacey et al. (2019).

Within Nepal, the synoptic conditions associated with heavy precipitation during the monsoon are low-pressure systems, mid-level troughs, western disturbances and break monsoon conditions (Bohlinger et al., 2017; Nandargi & Dhar, 2011). Heavy rainfall is often associated with northward moving low-pressure systems from the Bay of Bengal and the orographic uplift effect that causes precipitation upon approaching the Himalayan mountain range (Bohlinger et al., 2017). A 'break' in the South Asia summer monsoon shifts the monsoon trough from its normal position over North Indian River Plain, northwards to the foot of the Himalayas (e.g. Ramanadham et al., 1973). The sub-tropical ridge line in the upper troposphere shifts northwards during break and lies approximately above the lower monsoon trough. This seems to provide an effective process of removing ascending air in the lower monsoon trough causing exceptionally heavy rainfall over Assam and along the foot of the Himalayas. Specifically

considering extreme one-day rainfall events, Nandargi & Dhar (2011) suggest that these can occur in both monsoon excess and deficit years, and were most often associated with South Asian summer monsoon break conditions. Karki et al. (2017) show that generally, high-intensity extremes are more intense over the southern lowlands, but during the monsoon there are 3 distinct high-precipitation areas in the regions of Lumle, Gumthang and Num. The complexity in the spatial pattern of extreme rainfall is variable between averaging periods (e.g. annual vs seasonal), creating highly specific regional risks of floods, landslides and droughts, and highlighting the importance of assessing precipitation extremes at high spatial resolution.

2. Data Review

There are many sources of precipitation data over Nepal. Ostensibly similar datasets are known to show marked differences (e.g. Ceglar et al., 2017; Sun et al., 2018), and their reliability is principally defined by the number and spatial coverage of surface stations, satellite algorithms, and the data assimilation models that contribute to their creation. As summarised by Sun at al. (2018), rain gauges provide relatively accurate and trusted measurements of precipitation at single points but are unavailable over many sparsely populated and oceanic areas and can be affected by sampling errors. Satellite observations provide precipitation information with a greater degree of homogeneous spatial coverage but contain non-negligible random errors and biases owing to the indirect nature of the relationship between the observations and precipitation, inadequate temporal sampling given the rate of movement of some extreme rainfall systems, and deficiencies in the data processing algorithms needed to amalgamate their observations. In some cases, these direct observations are incorporated into reanalysis systems that use mathematically defined physical and dynamical processes in order to generate a synthesized estimate of precipitation across a uniform grid, with spatial homogeneity and temporal continuity.

In the context of this report, to make the best estimate of rainfall extremes, a critical review of these datasets is necessary to choose a subset that work best over Nepal. Several authors have undertaken reviews and comparisons of precipitation datasets. The multi-product gridded precipitation comparison of Dahri et al. (2021), evaluated across the western Himalaya and Hindu Kush regions, finds gauge-based and satellite ‘merged’ products performed better in dry regions and during the monsoon season, while reanalysis products provided better estimates in wet areas and during winter months (although this comparison is performed against the authors’ own dataset, described in Dahri et al. (2018), based on gauge data). Chen et al. (2021) suggest that model resolution is important in representing the complex orographic effects on precipitation in mountainous Nepal, and the steepness of slope appears to have a substantial impact on moisture advection/divergence parameterization terms (Zhang & Li, 2016) used in reanalysis datasets. As monsoon water vapor enters Nepal from the east and

moves west, we expect higher-resolution datasets to be a better representation of the true[a] precipitation in eastern Nepal because their model terrain accurately reflects the blocking effects of the high mountains. By comparison, coarse resolution datasets tend to allow more water vapor transport into the western region as they cannot capture terrain blocking effects. Hamm et al. (2020) show that as higher-resolution improves the representation of extreme precipitation, this results in an overall lower spatial mean precipitation but higher extreme events. Although the comparison of observational rain gauge estimates against gridded precipitation estimates is non-trivial, these types of comparisons show that generally gridded products tend to overestimate rainfall amounts in mountainous regions (e.g. Hamm et al., 2020; Zhang & Li, 2016), but higher resolution gridded data tend to compare more favourably. For these reasons, this study favours datasets with high resolution over comparable datasets with a lower resolution. Finally, the comparison by Nguyen et al. (2020) across Monsson Asia shows that there are strong sub-regional differences between datasets, that can be explained by the quantity and quality of rain gauges. Areas with a dense station network are typically found to show high consistency in spatial and temporal patterns, whilst large inter-product spread is found in areas with sparse station density. Furthermore, they show that that satellite products can have the spatial imprints of the underlying in situ data. and cannot necessarily be considered a perfect replacement to the lack of in situ data over data-sparse regions.

Generally, our stakeholder requirements reflect a need to examine precipitation extremes, with an initial focus on maximum 1-day precipitation (RX1day) within the Nepal monsoon season between June and September (JJAS). As these extremes occur on daily time scales, monthly data on its own is unsuitable. It is also important to estimate return periods, as these are widely used within infrastructure sectors, for which robust estimation requires lengthy records. For maximum decision making utility, a further translation of extreme rainfall to river flow will be necessary which we hope to perform in collaboration with partner organisations.

The following sections review literature associated with several common precipitation datasets from in situ, satellite, blended and reanalyses sources.

[a] In so far as we can ever be certainly of what is ‘truth’ given the limitations in observational methods already discussed.

2.1 ERA5 and related reanalysis datasets

Name	Period	Spatial Resolution	Temporal Resolution	Citation
ERA5	1950-present	0.25°[b]	Hourly	Hersbach et al. (2018)
ERA5-Land	1950-present	0.1°[c]	Hourly	Muñoz-Sabater (2021)
AgERA5	1979-present	0.1°	Daily	
WFDE5 **v1.1**	1979-2018	0.5°	Hourly	Lange (2019)

The ERA5 reanalysis combines the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) model with observations to provide a globally complete and consistent dataset. A global comparison against the Global Precipitation Climatology Project (GPCP) dataset (Nogueira, 2020) shows improved representation of convective rainfall, moisture convergence patterns and general reduction in systematic rainfall flux bias over South Asia, with respect to ERA-Interim, but no improvement to precipitation trends (for 1979-2018). When considering extreme rainfall events over India, Bhattacharyya et al. (2022) suggests that ERA5 is the best performing reanalysis product (out of 4[d]) but it still shows a notable dry bias of ~70mm for RX1day extremes against India Meteorological Department (IMD) gauge observations over high-altitude India.

ERA5-Land is a re-run of the land component of ERA5 (after Hersbach et al., 2018) at an enhanced native resolution of 9 km (vs native 31 km in ERA5). It uses ERA5 fields as atmospheric forcing without explicit data assimilation, but with improvements to the water and energy cycles model schemes at surface level (Muñoz-Sabater et al., 2021). The land-surface improvements do not directly benefit precipitation: in this case ERA5-Land total precipitation is based on the linear interpolation of ERA5 total precipitation over a triangular mesh (Muñoz-Sabater et al., 2021) which is unlikely to realistically deal with the directional heterogeneity of Nepal terrain, however, ERA5-Land improves estimates of soil moisture and river discharge estimations. Also note that unlike ERA5, uncertainty estimates are not (yet) available for ERA5-Land.

AgERA5 tailors ERA5 towards users in the agricultural domain, and includes daily aggregates of agronomic relevant variables, tuned to local day definitions and adapted to the finer topography, land use pattern and land-sea delineation of the ECMWF High-Resolution Forecast (HRES) operational model. It is downscaled (interpolated) and variables (except those related to precipitation) are bias corrected against HRES using a multiple linear

[b] Native resolution is 31 km, regridded to 0.25° for download purposes
[c] Native resolution is 9 km, regridded to 0.1° for download purposes
[d] Considering ERA5, MERRA-2 (Gelaro et al., 2017), PGF (Sheffield et al., 2006), and JRA-55 (Kobayashi et al., 2015)

regression approach. Seeing as precipitation is excluded from the bias correction scheme, there is no apparent advantage to using AgERA5 over ERA5-Land, and in both cases the naïve interpolation method used to obtain higher-resolution output is unlikely to benefit precipitation fields in a physically meaningful way.

WFDE5 is a bias-adjusted version of ERA5 aimed for use with impact models that provide a selection of near-surface meteorological variables (including rainfall flux) corrected against Climatic Research Unit (CRU) data and Global Precipitation Climatology Centre (GPCC) data. Although it provides hourly data, it is discounted due to its relatively coarse 0.5° spatial resolution.

2.2 IMDAA (Indian Monsoon Data Assimilation and Analysis) reanalysis dataset

Name	Period	Spatial Resolution	Temporal Resolution	Citation
IMDAA v0.3	1979-2019	0.12°	Hourly	Rani et al. (2021)

The Indian Monsoon Data Assimilation and Analysis (IMDAA) project uses the Met Office Unified Model (UM) with data assimilation to produce a 12-km resolution gridded dataset for the South Asia monsoon region. It combines observations (gauge and satellite) from ECMWF and Met Office archives, with additional IMD observations not otherwise available in other datasets. Evaluation of IMDAA against ERA5 over India (Rani et al., 2021), suggest these additional observations contribute a modest (possibly insignificant) improvement to the monsoon onset date compared to the IMD observational baseline. Comparing daily accumulated rainfall, IMDAA has greater monsoon season accumulation and fewer no-rain or light-rain days than both IMD and ERA5. Rani et al. suggest that this may be due to improved microphysics parametrisation in ERA5, but also note that the increased resolution of IMDAA, compared to ERA5, could also explain the rainfall intensity differences. However, it is not clear if or how Rani et al. account for these resolution differences in their analysis. Examining the monsoon circulation, a key driver of monsoon intensity, Rani et al. suggest that there are negligible differences in the strength of the low and upper level flow between ERA5 and IMDAA, suggesting that the large-scale drivers are similar. On this basis, differences in in precipitation accumulation could be primarily due to resolution differences and the associated microphysics that can be resolved.

The hilly regions of northern India may be a reasonably proxy for Nepal in the absence of any Nepal specific analysis. Singh et al. (2021) note an IMDAA wet bias over the foothills of the Himalayas compared to ERA5 and the IMD's lower resolution global reanalysis dataset

(NGFS, ~25 km resolution). Comparing the highest rainfall intensities (≥ 200 mm/day), IMDAA compares more favourably to IMD observations and NGFS global-reanalysis than ERA5, which appears to underestimate the number of very intense rainfall days > 550 mm/day (see T. Singh et al., 2021 Fig. 4). Further analysis specific to Nepal, and in relation to extreme rainfall events, is currently lacking for this dataset.

2.3 APHRODITE (Asian Precipitation - Highly-Resolved Observational Data Integration Towards Evaluation)

Name	Period	Spatial Resolution	Temporal Resolution	Citation
APHRODITE-2 v1901	1998-2015[e]	0.25°	Daily	Yatagai et al. (2012)

The APHRODITE-2 dataset (Yatagai et al., 2012) v1901 provides daily gridded precipitation at 0.25° over the monsoon Asia region. The dataset collates, quality controls and interpolates in-situ rain gauge data, including data from Nepal's Department of Hydrology and Meteorology (DHM). Version 1901 provides data for the period 1998-2015. Earlier versions of the data extend back to 1951, but recent versions improve the unification of the daily accumulation values, described in Yatagai et al. (2020), and are not comparable.

APHRODITE-2 includes ancillary information regarding the number of valid 0.05° cells (out of a possible maximum of 25) with at least one rain gauge that contributes to the value for each 0.25° grid cell in the final interpolated data set. As APHRODITE-2 is solely based on rain gauge data, it is important to consider the availability of gauges, especially in areas of complex terrain and low gauge density, such as Nepal. **Figure 1** shows the spatial and temporal variability in gauge data across Nepal: (a) shows that 61% of possible APHRODITE-2 grid cells over Nepal, do not have any gauge data within their boundary that contributes to their value. In these cases, their value in the final dataset is interpolated from the next nearest grid cells (as detailed in Yatagai et al., 2012).

APHRODITE-2 is compared against a number of alternative datasets: earlier versions of APHRODITE (V1003R1) correlate highly with GPCC and GPCP rainfall (for both summer and winter) but with an overall dry bias (Palazzi et al., 2013). Evaluation over Central Asia (not including Nepal) by Lai et al. (2020) suggests that APHRODITE (V1101) generally underestimates extreme precipitation threshold values versus gauge data, while

[e] Earlier versions of APHRODITE extend back to 1951, but it is not contiguous with this version due to differences in interpolation procedures

overestimating the total numbers of extreme precipitation events, particularly over the mountainous areas, and that these biases are more evident between May and September (MJJAS). Bai et al. (2018) show that over mountainous regions, gridded datasets based on traditional mathematical methods (e.g. Inverse Distance Weighted interpolation or Kriging) usually have smaller area-averaged precipitation than gridded datasets that account for windward, leeward and vertical (topographic) distribution of precipitation. Anecdotally, Dr Divas Basnyat from the Nepal Development Research Institute (NDRI) also suggests that APHRODITE underestimates rainfall over Nepal compared with observed (gauged) seasonal and annual average precipitation (not extreme rainfall) and runoff in some selected catchments in Nepal (personal communication, 2020).

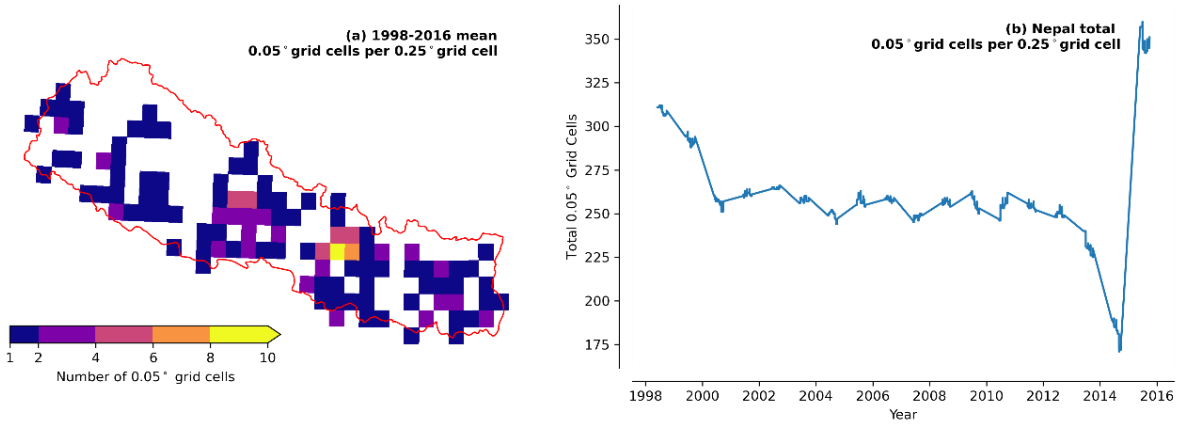

Figure 1 Diagnostic plots of data quantity in the APHRODITE-2 dataset over Nepal. (a) shows the 1998-2016 mean number of 0.05° grid cells, containing at least one rain gauge, that contribute to each 0.25° grid cell. For each 0.25° grid cell there are a total of 25 possible 0.05° grid cells that could contribute to its value. (b) shows how the total number of 0.05° grid cells with at least one rain gauge over Nepal varies between 1998 – 2016.

2.4 CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data)

Name	Period	Spatial Resolution	Temporal Resolution	Citation
CHIRPS v2	1981-present	0.25° & 0.05°	Daily, Pentad	Funk et al. (2014, 2015)

CHIRPS v2 is a quasi-global (50°S - 50°N) gridded dataset derived from: (i) the monthly precipitation climatology product CHPClim (based on satellite and gauge data); (ii) quasi-global geostationary thermal infrared satellite observations from the Climate Prediction Center (CPC, Janowiak et al., 2001) and the National Climatic Data Center (NCDC) GridSat-B1 (Knapp et al., 2011); (iii) the Tropical Rainfall Measuring Mission (TRMM, Huffman et al.,

2007); (iv) atmospheric model rainfall fields from the NOAA Climate Forecast System, v2 (CFSv2, Saha et al., 2010) and (v) in-situ precipitation observations. CHIRPS primarily focuses on 5-day rainfall accumulation (pentads) for the purposes of modelling agricultural drought. Shrestha et al. (2017) note that there is a ~90% reduction in the number of rain gauge stations used in the CHIRPS v2 blending process after 1992, for unknown reasons.

Ceglar et al. (2017) find CHIRPS v2 (0.25° version) shows greater seasonal variation and greater JJA mean precipitation (up to 420 mm / JJA wetter in some regions) than APHRODITE over Nepal. Over mountainous SW China, Bai et al. (2018) show CHIRPS to underestimate precipitation due to TRMM overestimating cloud top temperatures, and suggest that the blending techniques of alternative products such as CMOPRH (see Section 2.5) and MSWEP (see Section 2.8) should be preferred. Other studies also note that CHIRPS can struggle to differentiate rain vs no-rain conditions compared to APHRODITE (e.g. Tan et al., 2020).

2.5 CMORPH (CPC MORPHing technique)

Name	Period	Spatial Resolution	Temporal Resolution	Citation
CMORPH v1	1998-present	8 km (0.072°)	30 min[f]	Xie et al. (2019)

The CMORPH technique (Joyce et al., 2004), another quasi-global (60°S - 60°N) dataset, uses precipitation estimates derived from low orbit satellite passive microwave observations. Low orbit satellites have a relatively narrow band of view, referred to as the cross-track swath width, which in this case is ~2200 km. To propagate rainfall features between successive orbit periods, CMORPH uses feature information from geostationary satellite infrared (IR) data to interpolate the microwave-derived precipitation estimates in space and time.

Direct comparison against gridded Nepal DHM gauge data (Krakauer et al., 2013) suggests CMORPH is generally drier (especially at elevations > 1000m) but with higher interannual variability. Seasonally, CMORPH is notably drier (on the order of 100 – 150 mm/month) during JJAS compared to DHM stations (but also APHRODITE and TRMM). A more limited comparison in the Khumb region of northeast Nepal (Yamamoto et al., 2011) agrees that CMORPH underestimates JJAS precipitation, and also has a tendency to overestimate precipitation in the pre- (April-May) and post- (October) monsoon seasons.

[f] Daily version available at 0.25° resolution

2.6 IMERG (Integrated Multi-satellitE Retrievals for GPM) & related datasets

Name	Period	Spatial Resolution	Temporal Resolution	Citation
IMERG v06	1998-present	0.1°	30 min	Huffman et al. (2020)
AIMERG[g]	2000-2015	0.1°	30 min	Ma et al. (2020)

The Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (known as IMERG) v06 algorithm combines information from the Global Precipitation Measurement (GPM) satellite constellation (2014 – present) with earlier Tropical Rainfall Measuring Mission (TRMM) satellite measurements (2000 – 2015) to estimate precipitation over the majority of the Earth's surface. Both GPM (and formerly TRMM) satellite constellations host several types of passive microwave radiometers, but the IMERG products also combine input from geosynchronous IR and precipitation gauge analyses.

Although IMERG captures the seasonal precipitation cycle associated with the monsoon over Nepal (e.g. Sharma et al., 2020), Nepal et al. (2021) show that IMERG v06 has a dry bias (~ −2.5 mm/day) compared with DHM station data (most notable during the monsoon season) and overestimates the number of consecutive wet days. Assessing R1X and R5X indices, they calculate correlation coefficients of 0.37 and 0.53 respectively between DHM stations and IMERG. Over the Tibetan Plateau, Xu et al. (2017) suggest that IMERG overestimates the frequency of intense rainfall events (>10 mm/day), but underestimates the amount of light rain events (0 –1 mm/day) compared to gauge data from the China Meteorological Administration.

Recently, an Asia specific bias-corrected version of IMERG, AIMERG, has been released by Ma et al. (2020), combing IMERG with APHRODITE and the China Merged Precipitation Analysis (CMPA). This revised version appears to improve the capture of heavy rainfall events in terms of both the systematic bias and random error, although this assessment is based on evaluating rainfall associated with a typhoon over eastern China. Ma et al. show that the calibration (bias-correction) of AIMERG generally results in it being drier than IMERG over most of its domain.

[g] Asia only

2.7 HAR (High Asia Refined analysis)

Name	Period	Spatial Resolution	Temporal Resolution	Citation
HAR v2	2000-2020	10 km (0.088°)	Hourly	Wang et al. (2021)

HAR (Wang et al., 2021) is a regional atmospheric data set, focusing on high mountain Asia, generated by dynamically downscaling ERA5 using the Weather Research and Forecasting (WRF) model version 4.0.3 (Skamarock et al., 2019). It uses a daily re-initialization strategy, with a 12 hour spin-up period, and includes a snow depth correction to account for an ERA5 cold bias in this region.

Comparisons between water balance-derived precipitation and corresponding satellite data by Li et al. (2020) suggests that HAR (v1) data in the high mountain areas perform better than the precipitation data sets from APHRODITE, satellite remote sensing (including TRMM, GPM and GPCC) and coarse-resolution atmospheric reanalysis (including ERA5). Maussion et al. (2014) suggest that qualitatively HAR (v1) realistically reproduces both the relationship between the standard deviation and the mean of precipitation, and the characteristics of precipitation events on the Tibetan Plateau. Further analysis by Pritchard et al. (2019) shows HAR v1 (at 10 km) to have reasonable correspondence with most in situ measurements, general consistency with observed runoff, and to be able to reproduce the strong vertical precipitation gradients of the upper Indus basin.

HAR v2 extends the spatial domain, has longer temporal coverage and uses an updated version of the WRF model. Looking explicitly at indices of precipitation, Hamm et al. (2020) find that higher resolution datasets show the overall highest values for very wet days (> 10 or 20 mm precipitation accumulation), but have much lower mean values and lower maximum values for the count of wet-days. This implies that individual grid cells in higher resolution products (e.g., HAR v2) can experience more extreme precipitation events in multiple grid cells than coarser products. They also note that HAR v2 uses a cumulus parameterization scheme, which can, in rare occasions, lead to extremely high values (e.g. 500 mm in one day), which are not found in any of the other products. This feature isn't observed in a test 2 km version of HAR, in which convective systems are explicitly resolved. Comparing the HAR v2 with ERA5-Land (both at a similar spatial resolution) the difference between the linear interpolation of ERA5-Land and WRF-downscaling become clear. For example, while the grid-cell with the maximum total precipitation (in the 5-month period May-Sep) in ERA5-Land has 2400 mm, the maximum in HAR v2 totals 3 times as much (7865 mm). At the time of writing there are no explicit comparisons of HAR v2 data over Nepal.

2.8 MSWEP (Multi-Source Weighted-Ensemble Precipitation)

Name	Period	Spatial Resolution	Temporal Resolution	Citation
MSWEP v2.8[h]	1979-present	0.1°	3-hourly	Beck et al. (2019)

MSWEP (Beck et al., 2019) is a global combined product, merging rain gauge estimates with ERA5 and IMERG (as of v2.8). The authors also implement distributional bias corrections, mainly to improve the precipitation frequency, a correction of systematic terrestrial precipitation biases using river discharge data, and apply monthly climatological corrections using Climatologies at high resolution for the earth's land surface areas (CHELSA) dataset (of Karger et al., 2017).

Performance of MSWEP (v2.1, using ERA-Interim rather than ERA5) against India Meteorological Department rain gauge data (Mondal et al., 2018) compares well in the Indian Himalayas (annual averages) and appears to outperform CMORPH and two other satellite based products. A number of other studies compare MSWEP with several other common precipitation products (e.g. Derin et al., 2019; Ullah et al., 2019) and typically find MSWEP overestimates the precipitation at both low and high altitudes relative to gauge observations, but compares well when used to drive hydrological models, comparing hydrological output e.g. Chen et al. 2020). However, it is hard to judge the performance of v2.8 based on studies of v2.2 or earlier, due to the recent release of v2.8 which newly incorporates ERA5, and removes ERA-Interim.

[h] Released February 4, 2021 – see Technical Documentation

2.9 Seasonal Prediction Systems

Name	Period	Ensemble Members	Spatial Resolution	Temporal Resolution	Citation
GloSea 5 (GC2)	1993-2016	24 24	0.83° x 0.56° [i] 0.35° x 0.23° [j]	Daily	MacLachlan et al. (2015) Scaife et al. (2019)
GloSea 6 (GC3)	1993-present	28	0.83° x 0.56°	Daily	Davis et al. (2020)
DePreSys 3 (GC2)	1960-2018	80[k]	0.83° x 0.56°	Monthly	Dunstone et al. (2016)
SEAS5	1981-present	25[l]	0.83° x 0.56° [m]	Daily	Johnson et al. (2019)

There are several seasonal prediction systems, providing ensemble output for both forecasts and hindcast (re-forecast) periods of at least 24 years. The large number of ensemble members means these systems can be used to explore unobserved extreme events, a methodology commonly referred to as the UNprecedented Simulated Extreme ENsemble (UNSEEN) approach, by pooling ensemble members to give a larger-than-observed sample of plausible years of data (e.g. Jain et al., 2020; Kelder et al., 2020; Kent et al., 2022; Thompson et al., 2017).

A selection of seasonal prediction systems include:

- **GloSea5** (MacLachlan et al., 2015) is built around the Met Office HadGEM3 atmosphere-land-ocean-sea-ice coupled climate model and incorporates the Met Office operational 4D-Var data assimilation scheme into the model initial atmospheric conditions.
- **GloSea6** (Davis et al., 2020) currently has broadly the same model configuration as GloSea5 but with updated model physics (using atmosphere configuration GC3), a more realistic treatment of land-surface initialization and will eventually have more ensemble members.
- **DePreSys3** operating in both 'decadal' and 'interannual' modes, is also based on Met Office HadGEM3-GC2. It includes a full-field data assimilation scheme that nudges the model towards the observed analyses in the atmosphere, ocean and sea-ice (Dunstone et al., 2018; Dunstone, Smith, Scaife, et al., 2016).

[i] Also referred to as N216. Operational version, currently using GC3 configuration.
[j] Also referred to a N512. Research version, not operationally available, used GC2. Only 1993-2015.
[k] 40 ensemble members, initialised twice per year in May and November.
[l] The hindcasts have 25 ensemble members, but the forecasts have 51 members.
[m] Also referred to as spectral truncation T319

- **SEAS5** the ECMWF seasonal forecast, provides 25 ensemble members at a similar resolution to operational GloSea5 (at N216 resolution) for a 36 year re-forecast period. The method for generating ensemble members in SEAS5 is different to Met Office models (see Johnson et al., 2019 for details) and therefore it may provide useful additional atmospheric sampling for future refinements of our analysis.

A more comprehensive comparison of seasonal prediction systems is available in Stacey et al. (2021).

Each of the seasonal forecast models favour a time-lagged ensemble approach (e.g. Dalcher et al., 1988; Hoffman & Kalnay, 1983), whereby at any given point in time, each ensemble member has been initialised from a spread of date-times equal to or preceding the given point in time. GloSea5 and SEAS5 hindcast members run for c. 7 months and DePreSys3 interannual (decadal) hindcasts run for 16 months (~5.5 years).

The key differences between GloSea5 and DePreSys3 lies in their initialisation strategies and external forcing datasets. DePreSys3 uses a weakly coupled data initialisation strategy, where the atmosphere, ocean and sea-ice are nudged towards observed analyses in a single assimilation run of the model. In contrast, GloSea5 combines atmosphere, ocean and sea-ice initial conditions at the point of forecast initialisation. From an atmosphere perspective, the initialisation data is essentially the same for the hindcasts of the two systems, as both use ERA-Interim reanalysis. However, DePreSys3 nudges towards the ERA fields, whilst GloSea5 takes the instantaneous field at the hindcast initialisation time. From an ocean perspective, DePreSys3 and GloSea5 have very different assimilation approaches: GloSea5 covers a relatively data-rich period for sub-surface ocean coverage, whereas DePreSys3 runs hindcasts back to 1960, where sub-surface ocean data was extremely sparse (see Dunstone, Smith, & Hermanson, 2016 for further details).

The strengths of seasonal forecasts can be assessed based on the predictive capability of the models. In this context, GloSea5 has a good ability to predict early/normal/late south Asia monsoon onset (e.g. Chevuturi et al., 2019) and to represent key ENSO variability, but has erroneous sea-surface temperature anomalies related to the IOD, both of which are key drivers of Indian rainfall (e.g. Johnson et al., 2017). GloSea5 also poorly represents changes to the Siberian high (e.g. Lim et al., 2018; Lu et al., 2017) which has been shown to be an important component of South Asian summer monsoon rainfall seasonal forecasting (R. Singh et al., 2021), and does not properly represent the shifting of the Somali jet axis towards the equator during break phases of the monsoon, an important feature of the monsoon intraseasonal oscillation (Jayakumar et al., 2017). Stacey et al. (2021) show that GloSea5 and SEAS5 have a significant and positive correlation with CHIRPS, for mean-daily JJAS precipitation over western Nepal, and receiver operating characteristic (ROC) scores > 0.6 (i.e. the false positive rate versus the true positive rate of the model as compared to CHIRPS observations) for JJAS. From the point of view of providing a realistic representation of current

precipitation climatology, analysis by Jain et al. (2020) shows GloSea5 model fidelity to be sufficiently similar to Indian precipitation, being within ±2σ of the IMD JJA-mean monsoon rainfall observations, albeit for an India area average.

On decadal time scales, DePreSys3 is also shown to have modest but significant interannual predictive capability in the South Asia monson region (Dunstone et al., 2020), but in this case we prefer the higher spatial and temporal resolution of GloSea5 N512, which has over 5.5x the number of grid cells for a given area compared to DePreSys.

Practically, GloSea5 N512 (0.35° x 0.23°) provides 24 years x 122 JJAS days x 24 members (3 initialisation dates per year x 8 realisations at each initialisation date) = 70,272 ensemble days, versus DePreSys3 which provides 59 years x 4 months x 40 members x 2 initialisations = 18,880 ensemble months. Based on the different initialisation strategies and hindcast periods, DePreSys3 data will probably sample a greater range of plausible rainfall situations, but in the context of this work, the higher resolution of GloSea5 (N512 vs N216) and daily (vs monthly) output is preferred over DePreSys3. For this work, we prioritise the ease of access to our in-house seasonal forecast products as it appears necessary to retrieve SEAS5 data via MARS (rather than the C3S Climate Data Store) to obtain full resolution data for the full hindcast period.

2.10 Data Selection

Based on this literature review, we suggest basing our evaluation of current precipitation extremes on:

(i) MSWEP v2.8, HAR v2 and IMDAA represent a range of methods used to define the best estimates of observed precipitation intensity, and

(ii) GloSea5-N512 for the best estimate of unobserved precipitation variability.

This selection is somewhat arbitrary, but aims to balance data sources and methodological approaches to constructing gridded data, whilst maintaining high grid resolution and trying to avoid multiple datasets that share common source data which would unduly bias later data processing efforts. Future updates could consider incorporating IMERGv06 data, and additional estimates of precipitation variability from either DePreSys3 or SEAS5.

Datasets that were also considered for discussion in this Section but were dismissed include: GPCC-FD (Schneider et al., 2011, 2017) and GPCP v2.3 (Adler et al., 2003, 2018) due to their relatively coarse resolutions (≥ 0.5°); individual satellite products, such as TRMM and GPM products, are dismissed in favour of the combined satellite products, such as IMERG and CMOPRH; and the PERSIANN CCS-CDR dataset (Ashouri et al., 2015) looked promising

(and should be considered for future updates) but is only available as flat binary files, which we failed to get working with common data conversion software. We also regret that we were unable to obtain access to the full Nepal DHM in-situ gauge dataset. Every effort should be made to incorporate this into future analysis.

Figure 2 and Figure 3 present an initial visualisation of these selected datasets. For the common period 2000 – 2015 they show mean JJAS precipitation accumulation (mm) and the maximum 1-day (RX1day) JJAS accumulation rainfall respectively, against common baseline datasets APHRODITE and ERA5. For illustrative purposes, each dataset is plotted at their native resolution, as is immediately available to users, although we note that for a formal comparison it is important to re-grid data to a common resolution in order to compare equivalent grid-box averages. Nonetheless, even amongst these well respected datasets, the degree of variability is striking. APHRODITE appears drier than both ERA5 and GloSea5 (both at similar resolutions), and the patterns of rainfall intensity amongst the higher resolution models is notably different for both rainfall climatology and extremes. The following section explores these differences further.

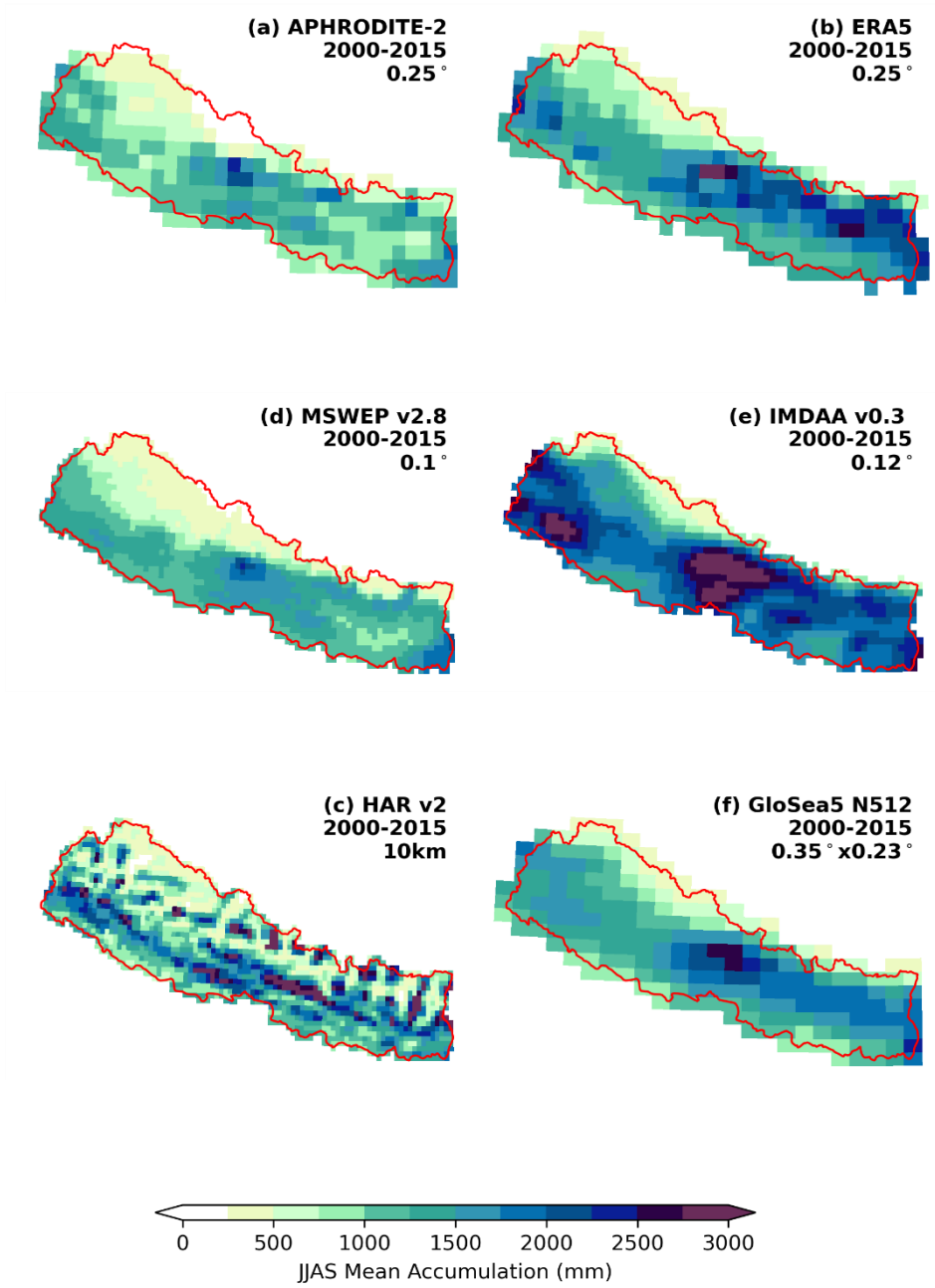

Figure 2 Comparison of mean JJAS precipitation accumulation (mm) for the common period 2000-2015 for 6 datasets. Note the different resolutions of each dataset implies that JJAS mean accumulation values are based on different spatial averages.

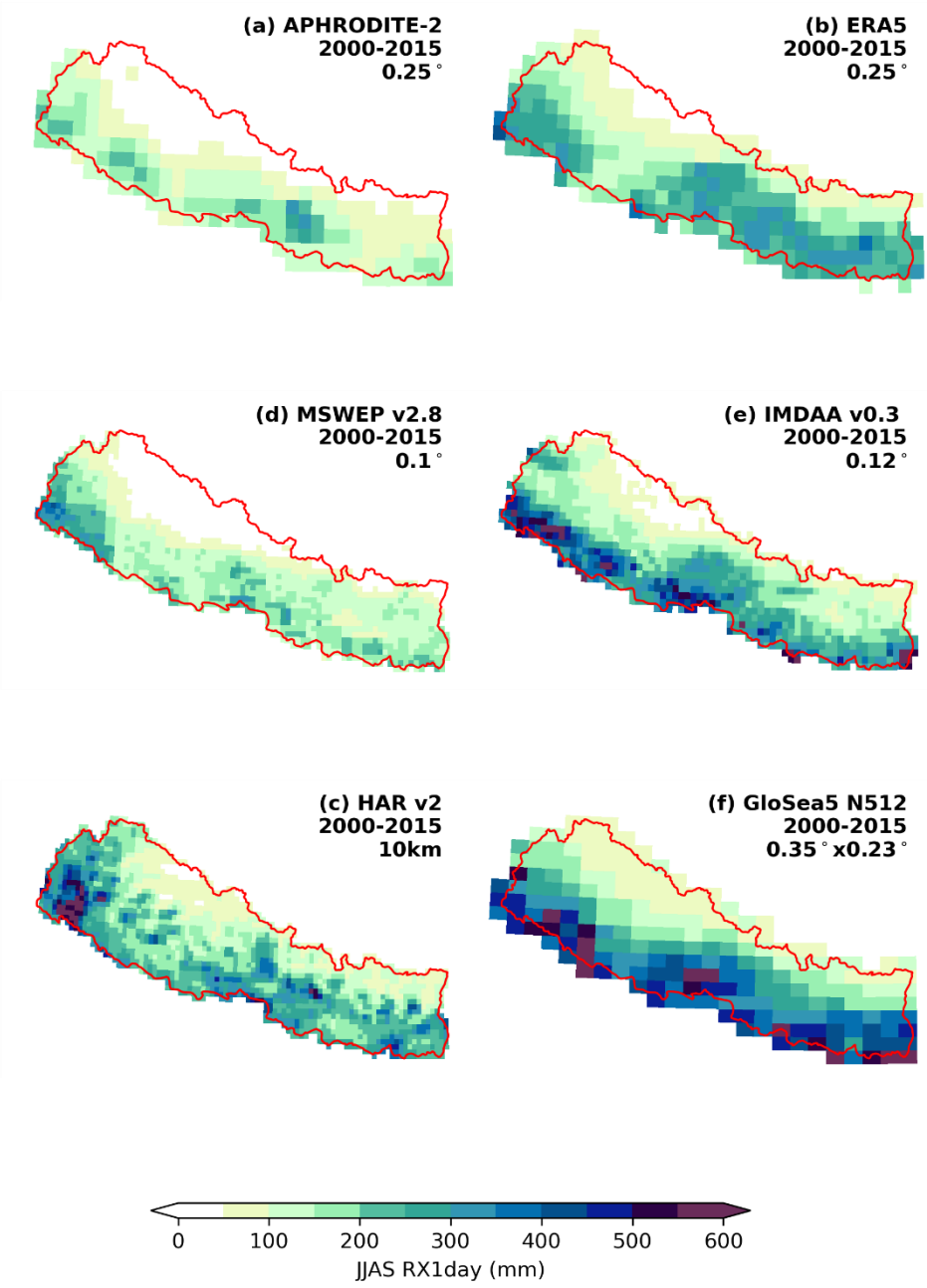

Figure 3 Comparison of maximum RX1day JJAS precipitation (mm) for the common period 2000-2015 for 6 datasets. As for Figure 2, the different resolutions of each dataset implies that RX1day values are based on different spatial averages.

3. Comparison of Observations and GloSea5

To formally compare the differences between precipitation estimates, each data set is first regridded to the coarsest common resolution, in this case GloSea5 N512 (0.35° x 0.23°). We use the 2nd order conservative remapping of Jones (1999) via the Climate Data Operators (CDO) toolset of Schulzweida (2020) to facilitate a like-for-like comparison of grid cell averages.

Our primary interest is then to assess the realism, or fidelity, of the GloSea5 simulations, so as to judge if GloSea5 can provide a supplementary source of unobserved precipitation data, in addition to the limited observational record. This comparison takes place from two different points of reference, across a common period: (i) comparing dataset climatology, (ii) comparing extremes. In principle, to have confidence in the use of GloSea5 data over Nepal, the comparisons of GloSea5 climatology should be similar to other datasets. By design, we expect the comparison of GloSea5 extremes to be different to observational datasets, but undertake a formal comparison so as to understand the magnitude and direction of the differences.

Our testing focus is the climatological representation of wet days (where daily accumulation ≥ 1 mm) as this is the primary interest of this ARRCC workstream. By extension, we do not explicitly test the fidelity of dry data (where daily accumulation < 1 mm), and we do not necessarily expect the fidelity conclusions derived for wet data to apply to dry data.

Where possible, the fidelity comparison is performed across the full 24 year GloSea5 N512 period (1993-2016), but the APHRODITE and HAR datasets impose additional temporal limits. The comparisons for these datasets are made for the period 1998-2015 (18 years) and 2000-2016 (17 years) respectively. This is half of the typical 30 year period used for assessing climatology, but we are restricted to a single common period to facilitate a fair comparison.

We use Generalised Additive Models (GAMs, Hastie & Tibshirani, 2017; Wood, 2017) to facilitate data pooling between grid cells, to enhance the signal and reduce grid cell noise in the spatial domain. We make two general assumptions with respect to both the precipitation climatology and precipitation extremes: (1) that any long-term climate change signal is negligible during these periods, and (2) there is no impact of multi-decadal climate variability. Consequently, the GAMs only have longitude and latitude covariates.

3.1 Comparison of Precipitation Climatology

For precipitation climatology, we examine the spatial variability of seasonal (JJAS) accumulation. We fit a Tweedie distributions to each dataset using GAMs, and compare the Tweedie mean parameter (Figure 4) and the effect size (Figure 5). In both cases, we

primarily compared GloSea5 against each of the other datasets. **Figure 6** shows an example of the underlying data distributions at one grid cell, centred on Kathmandu.

The effect size specifically compares the difference in means (μ), scaled by the control data standard deviation (σ). This helps to contextualise the difference between the two data distributions in a more practical and holistic manner, in a way that the mean alone does not.

In this case, we define the absolute effect size[n] as follows:

$$\text{Effect Size} = \left| \frac{\mu_{GloSea5} - \mu_{Control}}{\sigma_{Control}} \right| \quad (1)$$

where the control data is the dataset against which GloSea5 is being compared.

Means that are robustly different but which have comparatively large variances (and large overlap between their respective distributions) will intuitively appear more similar than a lone assessment of their means would suggest. A large effect size implies there is a larger percentage of non-overlap between the two distributions, and can be interpreted in terms of the Common Language Effect Size (CLES) metric of McGraw & Wong (1992). The CLES describes the probability that a JJAS accumulation year from GloSea5 will be higher than from the control dataset if both are chosen at random. Figure 5 shows the link between effect size and CLES.

GloSea5 is almost uniformly wetter (Δμ ≥ 100%) than either APHRODITE or MSWEP (Figure 4), but the effect size is only notable (≥ 0.9, CLES=74%) over the Himalayas (Figure 5), predominantly in western regions. For the IMDAA dataset, although there is a strong east-west gradient difference, often ≥ 100%, with GloSea5 being wetter in the west (possibly due to resolution-dependent moisture advection mechanism described by Chen et al., 2021) then effect size is mostly ≤ 0.6 (CLES=66%). Differences between GloSea5 and EAR5 or HAR are much lower, also with negligible effect sizes.

We summarise that GloSea5 JJAS accumulation is not robustly different to ERA5 or HAR. For IMDAA, MSWEP and APHRODITE, some GloSea5 annual JJAS accumulation is robustly different in specific locations. In particular, GloSea5 is expected to be wetter in ≥ 66 out of 100 years over most of Nepal compared to the IMDAA datasets, ≥ 74 out of 100 years over the Himalayas compared to MSWEP, and ≥ 90 out of 100 years for APHRODITE. For all other

[n] In principle, the effect size can be positive or negative which indicates the direction of difference. In this case, this direction of difference is already reflected by examining the different in means (Figure 4) and looking at the absolute value simplifies the visualisation in Figure 5.

regions (Tarai and non-Himalyan areas) the difference in means is large, but the effect size is smaller (≤ 0.6) for MSWWP and APHRODITE.

Visualisation of the data distribution for a grid cell containing Kathmandu shown in Figure 6 helps rationalise the interplay between the difference in means and effect size. For this location, although the HAR mean JJAS accumulation is significantly larger than GloSea5 mean JJAS accumulation (Figure 4f), the spread in the GloSea5 estimate is large enough to fully encompass the range of HAR estimates (Figure 6, right panel), such that the effect size (Figure 5d) is small. Although a fidelity test based on differences in means might fail in this case, incorporating the data variance via the effect size shows that this difference is not significant.

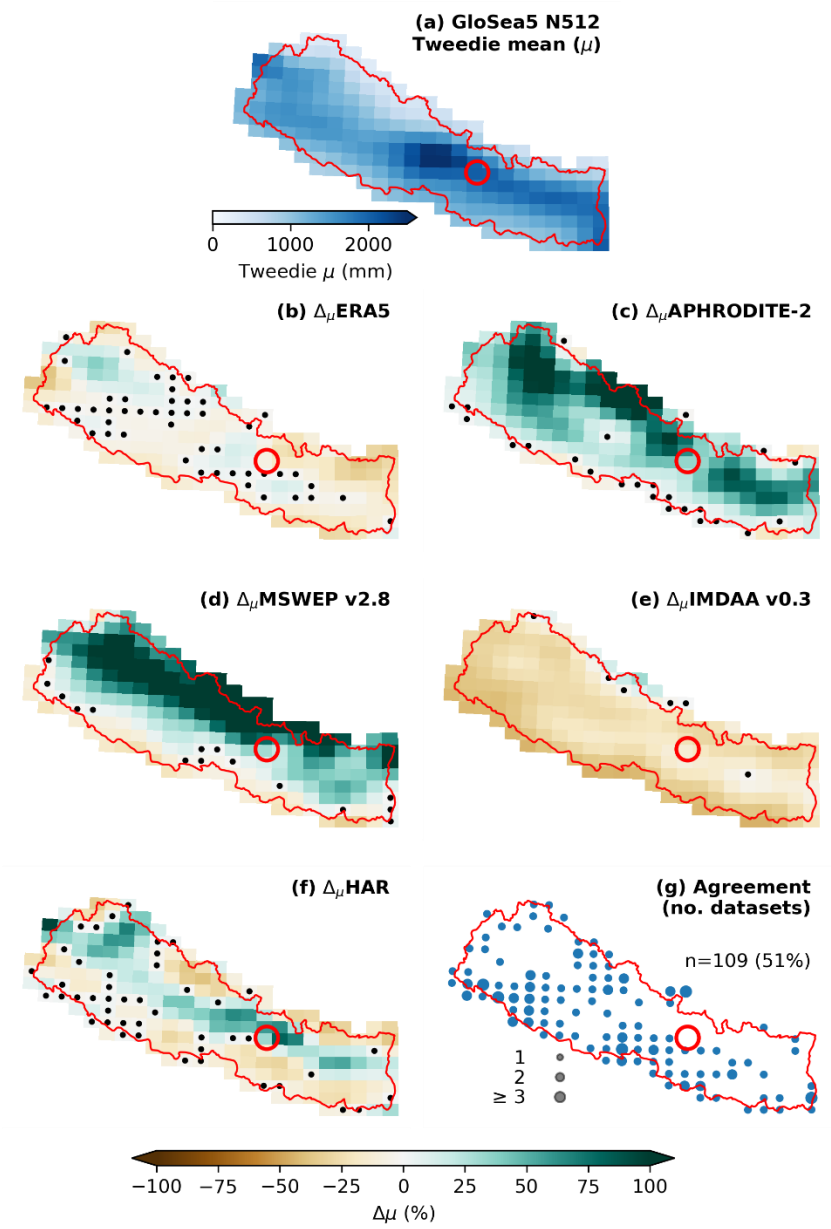

Figure 4 GloSea5 climatological mean (μ) based on a fitted Tweedie distribution (a) and the percentage difference between fitted mean parameters for the 5 observational datasets against GloSea5, Δμ (b – f). Black dots show locations where the difference between the two location parameters is not significantly different to zero at the 5% level, based on 10,000 simulated mean parameter estimates. GloSea5 mean is larger (lower) than observations for positive (negative) Δμ. A summary of grid cells with insignificant differences between GloSea5 and observations (b – f) is shown in (g). Larger circles show greater intermodal agreement, where the differences are insignificant in more of the comparisons. *n* is the total number of insignificant grid cells. All comparisons are done at GloSea5 resolution (0.35° x 0.23°) for common time periods. The red circle denotes the grid cell containing Kathmandu.

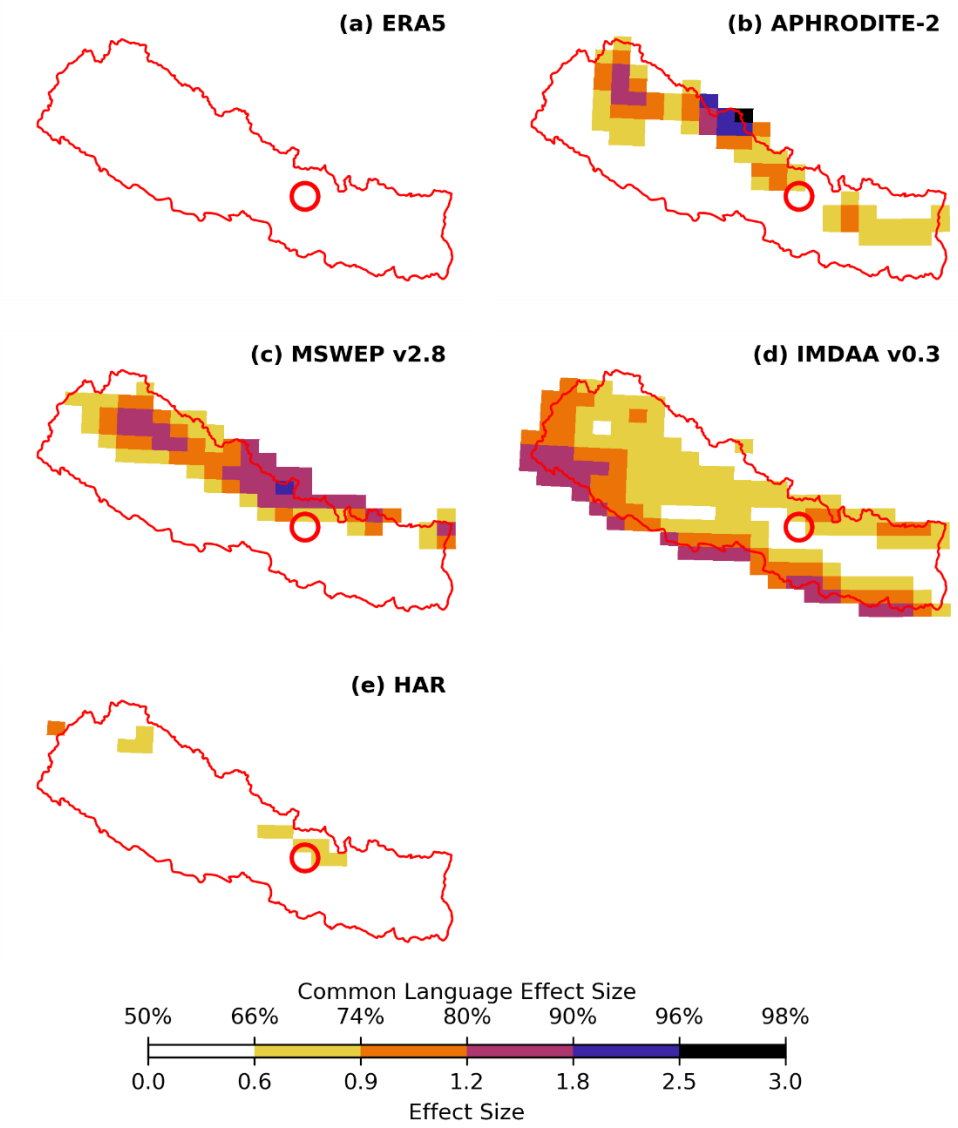

Figure 5 Effect size as per Equation (1) comparing GloSea5 against datasets (b – f) in Figure 4. The colour scale shows the calculated effect size and the corresponding Common Language Effect Size (CLES) as defined by McGraw & Wong (1992). All comparisons are done at GloSea5 resolution (0.35° x 0.23°) for common time periods.

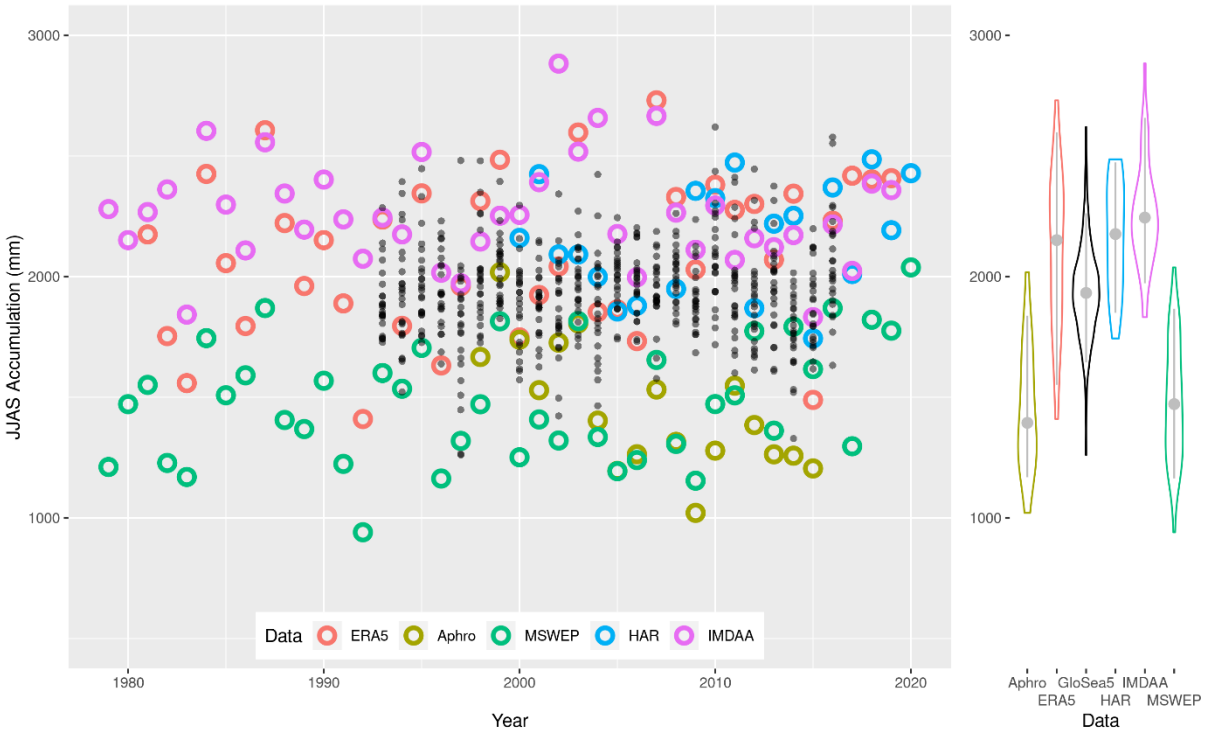

Figure 6 JJAS annual accumulation at one grid cell centred on Kathmandu (centred on 27.8N 85.3E) for each model in Figure 4. The time series of JJAS accumulation (left) shows GloSea5 ensemble members in black, with other datasets as coloured rings. These are aggregated (right) into violin plots of the overall distribution, with grey dots marking the 50th percentile, and grey lines showing the 5th and 95th percentile extent.

3.2 Comparison of Precipitation Extremes

For comparing extremes, we use an extreme value analysis (EVA) approach (e.g. Kelder et al., 2020; Kent et al., 2022) based on the comparison of seasonal block-maxima. This fidelity comparison of extremes, makes use of the *evgam* package (Youngman, 2020) to fit GAMs as for Section 3.1. As the focus is on the monsoon season (June – September, JJAS) we examine the JJAS block maxima of 1-day maximum rainfall (RX1day) for the common range of years available from each dataset, fitting a Generalise Extreme Value (GEV) distribution. Figure 7 is equivalent to Figure 6, showing an example of the temporal variability of JJAS block-maxima from each dataset at the grid cell containing Kathmandu.

Figure 7 Temporal variability in JJAS maximum one-day accumulation (RX1day). As for Figure 6, this plots datasets for the grid cell containing Kathmandu (centred on 27.8N 85.3E).

As before, the GAM only has longitude and latitude covariates, in this case for both the location (μ, the centre of the GEV distribution) and scale (σ, the spread of the GEV distribution) parameters. Although there appears to be some debate on the variability of the shape parameter (ξ) in the literature (e.g. Ragulina & Reitan, 2017), due to the relatively small spatial extent of the Nepal domain, we keep ξ constant, so it does not vary in space. Typically we would expect $0 < \xi < 0.1$ for precipitation extremes (e.g. Papalexiou & Koutsoyiannis, 2013). Brown (2018) notes that ξ is often poorly constrained, even with large datasets, and it is generally difficult to discern robust spatial variability. Finally, we note that although topographic elevation is expected to have explanatory power in terms of rainfall extremes, it has a strong directional bias (elevation change across Nepal is aligned NE-SW) such that we expect strong concurvity (i.e. the non-parametric analogue of multicollinearity) with longitude and latitude that could lead to an underestimation of σ. Hence, for this analysis we exclude topography from the GAM.

In terms of GloSea5, ensemble members are derived from 8 ensembles started from 3 different initialisation times. In practise we assume each ensemble member is independent of the others, although further detailed analysis is required to check this.

Differences are assessed in terms of μ, σ and the 20-year JJAS-max RX1day return period. The significance of differences is based on the difference of 10,000 simulated model parameters (μ and σ) from the GEV fits for each dataset, and 20 year return period values estimated from the GEV sampling distribution (after Youngman, 2020), between GloSea5 and each other model. From the 10,000 values, the 5th to 95th percentile range is calculated. Variability in simulated estimates reflects GEV fit uncertainty. Where the range in the difference between simulated parameters from the GloSea5 GEV model and the other datasets includes zero, we conclude that there is no significant difference at the 5% level. Although the shape parameter is not explicitly compared (it varies between models, but is constant in space), it is used when calculating the 20-year return level.

3.2.1 GEV Location

The fitted GEV location (μ) from GloSea5 shows opposing differences compared to the two baseline observational datasets APHRODITE-2 and ERA5 (Figure 8). With significant differences in almost all grid cells, GloSea5 (median μ = 59 mm, range [20, 121] mm across Nepal) typically has lower extreme RX1day rainfall[o] than ERA5 (Nepal median Δ_{μ}ERA5 = -14 mm, [-60, 19] mm) but larger RX1day rainfall than APHRODITE-2 (Δ_{μ}APHRODITE-2 = 16 mm, [-10, 43] mm). The only spatial consensus between these two datasets, is that the annual-maxima RX1day precipitation in the NW border region is slightly higher in GloSea5, than either ERA5 or APHRODITE-2.

Comparison against other observational-based datasets is also varied. Nepal summaries show that GloSea5 RX1day is typically higher than MSWEP (8 mm, [-31, 38] mm) and lower than IMDAA (-3 mm, [-13, 39] mm) and HAR (-12 mm, [-53, 23] mm). Across all five datasets, consensus differences suggest GloSea5 RX1day to be higher than observations in the High Mountain areas of central provinces Gandaki and Bagmati (in 4 out of 5 datasets), lower in the lowlands of Sudurpashchim province (western Nepal, 3 of 5 datasets) and lower in the eastern provinces (Province 1 and 2, 3 out of 5 datasets). Out of 213 possible grid cells, 102 cells (48%) have at least one model with no significant difference in μ compared to GloSea5.

3.2.2 GEV Scale

Comparison of GloSea5 GEV scale (σ) parameters shows similar consistency in inter-model differences, but a larger number of grid cells at which the differences are not judged to be significantly different to zero. Comparing ERA5 and APHRODITE-2, GloSea5 σ parameter (20 mm [6, 50] mm) is larger in almost all grid cells than APHRODITE-2 (7 mm [-9, 28] mm) and typically smaller than ERA5 (-4 mm [-33, 15] mm), but note that there are a greater number of insignificant differences with ERA5. There is a large degree of consensus between the areas that have greater σ in GloSea5 than ERA5 or APHRODITE, including most of the Middle

[o] GloSea5 location parameter is larger (lower) than observations for positive (negative) Δμ

Hills and Himalayas. Both MSWEP, HAR and IMDAA reflect similar differences to those seen for the location parameter (Section 3.2.1), with Δ_{μ}MSWEP = 8 mm, [- 31, 38] mm, Δ_{μ}HAR = -6 mm [-41, 15] mm and Δ_{μ}IMDAA = -3 mm [-14, 9] mm.

Across all datasets, there is a consensus that GloSea5 is under-spread in the far west (in 3 of 5 models), with a smaller over-spread area to the west of Kathmandu (3 of 5). However, it's notable that there are a much greater total number of insignificant grid cell differences (i.e. the spread in GloSea5 is not significantly different) across datasets (in 167 cells, 78%) compared to μ.

3.2.3 20-year JJAS RX1day

Differences between return estimates reflect the difference in the combined effect of μ, σ and ξ parameters. In this case, we focus on a 20-year return period, such that where GloSea5 is wetter (drier), its 20-year RX1day return level is larger (smaller). The median Nepal 20-year return level for GloSea5 = 138 mm [44, 310] mm, with differences between APHRODITE-2 (55 mm, [-25, 146] mm) and MSWEP (23 mm, [-67, 130] mm) being generally drier, and ERA5 (-19 mm, [-131, 72] mm), IMDAA (-9 mm, [-50, 67] mm) and HAR (-22 mm, [-171, 61] mm) being generally wetter. Comparing the spectrum of models, the GloSea5 20-year estimate is not significantly different to at least one dataset in 164 out of 213 possible grid cells (77%).

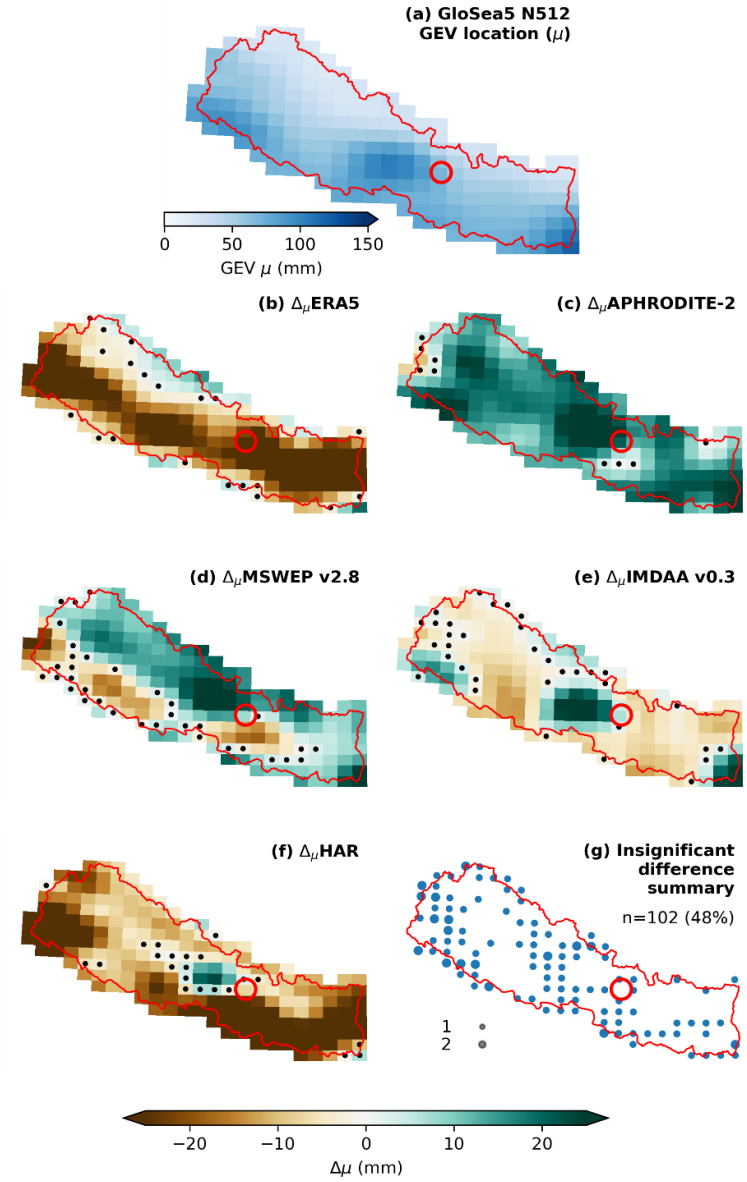

Figure 8 GloSea5 GEV location parameter (50% percentile, a) and the difference between GEV location parameters for the 5 observational datasets against GloSea5, Δμ (b – f). Black dots show locations where the difference between the two location parameters is not significantly different to zero at the 5% level, based on 10,000 location parameter estimates. GloSea5 location parameter is larger (lower) than observations for positive (negative) Δμ. A summary of grid cells with insignificant differences between GloSea5 and observations (b – f) is shown in (g) by coloured circles. Larger circles show greater intermodal agreement, where the differences are insignificant in more of the comparisons. *n* is the total number of insignificant grid cells. All comparisons are done at GloSea5 resolution (0.35° x 0.23°) for common time periods. The red circle denotes the grid cell containing Kathmandu, the location for Figure 7.

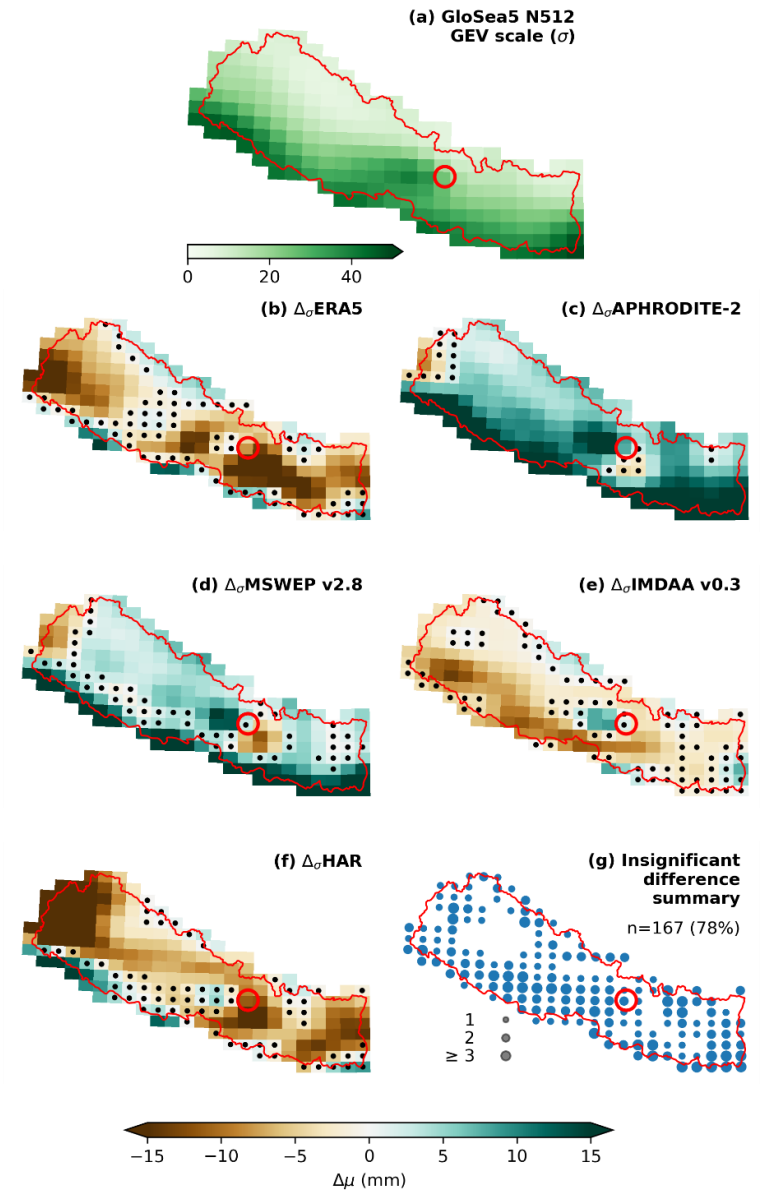

Figure 9 *As for Figure 8*, but GloSea5 GEV scale parameter (a) and the difference between GEV scales parameters for the 5 observational datasets against GloSea5, Δσ (b – f). Note the change in colour bar scales compared to *Figure 8*. A summary of grid cells with an insignificant difference between GloSea5 and observations (b – f) is shown in (g) by coloured circles. Larger circles show greater intermodal agreement, where the differences are insignificant in more of the comparisons.

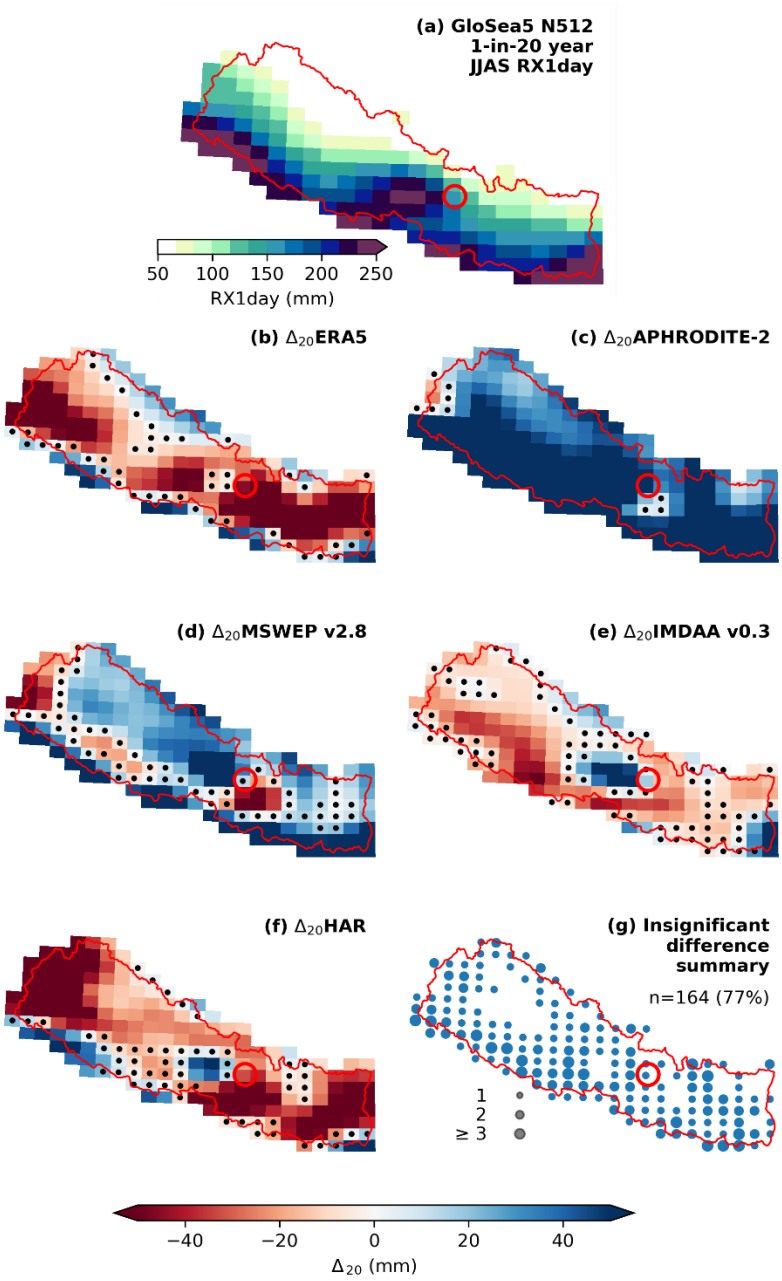

Figure 10 As for Figure 8, but GloSea5 1-in-20 year JJAS RX1day return level (a) and the difference between 1-in-20 year JJAS RX1day for the 5 observational datasets against GloSea5, Δ_{20} (b – f). Note the change in colour bar scales compared to Figure 8. A summary of grid cells with an insignificant difference between GloSea5 and observations (b – f) is shown in (g) by coloured circles. Larger circles show greater intermodal agreement, where the differences are insignificant in more of the comparisons.

4. Summary

Observational-based datasets estimating precipitation events over Nepal show a considerable spread. Following a review of 16 well-known datasets, we suggest that MSWEP v2.8, HAR v2 and IMDAA represent a balance of resolving capability and methodological spread that should reasonably sample the true precipitation variability. For the purposes of comparison, to this set of 3 datasets we add ERA5 and APHRODITE-2 as well-known baselines (albeit noting the relative scarcity of gauges contributing to the APHRODITE-2 estimates over Nepal, as shown in Figure 1). Although these five datasets are deliberately chosen to represent different methodologies and sources of raw data, they are each versions of the same reality. For a data user wishing to estimate metrics relating to extreme precipitation events, this diversity in estimates is typically unhelpful. Unlike ensemble weather and climate data, it is substantially harder to judge the relative merits and/or independence of each dataset in such a way as to be able to draw statistically valid summary statistics or uncertainty estimates.

In the context of comparing the performance of the Met Office seasonal forecast model GloSea5, against these five datasets, we find a similar level of disagreement. The process of 'fidelity testing' typically compares summary statistics of the GloSea5 ensemble against an observational baseline. In this case, there is no obvious baseline dataset to use. In general, comparisons of climatological differences (JJAS total rainfall accumulation) show that GloSea5 is not significantly different to ERA5 or HAR (effect size ≤ 0.6). For the other datasets, GloSea5 JJAS accumulation is notably larger than APHRODITE-2 and MSWEP over the Himalayas (effect size ≥ 0.6), and notably smaller than IMDAA over most of Nepal. Comparing extremes, we note that these differences in climatologies do not hold for GloSea5 RX1day annual-maxima. For 1-in-20 year return periods, GloSea5 generally predicts a wetter 1-in-20 year event than APHRODITE-2 and MSWEP, but is generally drier than ERA5, HAR and IMDAA. Without preferring one comparison over the others, we find the GloSea5 1-in-20 year return estimates (at each grid cell) to be insignificantly different to the observed 1-in-20-year return estimates for 78% of Nepal.

Lacking a single robust data set for estimating precipitation extremes, future work within ARRCC will consider a Bayesian methodology for data melding. The aim will be to develop a framework for assessing the credibility of precipitation extremes across Nepal, accounting for the variation in precipitation data as observed in this report. If possible, a fully Bayesian approach should facilitate the incorporation of prior uncertainty for each dataset, and account for the natural hierarchy of data when incorporating ensemble data from GloSea5 with deterministic estimates from MSWEP, IMDAA and HAR data sources.

References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., et al. (2003). The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). *Journal of Hydrometeorology*, *4*(6), 1147–1167. https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2

Adler, R. F., Sapiano, M., Huffman, G., Wang, J.-J., Gu, G., Bolvin, D., et al. (2018). The Global Precipitation Climatology Project (GPCP) Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation. *Atmosphere*, *9*(4), 138. https://doi.org/10.3390/atmos9040138

Ashok, K., Guan, Z., & Yamagata, T. (2001). Impact of the Indian Ocean dipole on the relationship between the Indian monsoon rainfall and ENSO. *Geophysical Research Letters*, *28*(23), 4499–4502. https://doi.org/10.1029/2001GL013294

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., et al. (2015). PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. *Bulletin of the American Meteorological Society*, *96*(1), 69–83. https://doi.org/10.1175/BAMS-D-13-00068.1

Bai, L., Shi, C., Li, L., Yang, Y., & Wu, J. (2018). Accuracy of CHIRPS Satellite-Rainfall Products over Mainland China. *Remote Sensing*, *10*(3), 362. https://doi.org/10.3390/rs10030362

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., et al. (2019). MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society*, *100*(3), 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1

Behera, S. K., Krishnan, R., & Yamagata, T. (1999). Unusual ocean-atmosphere conditions in the tropical Indian Ocean during 1994. *Geophysical Research Letters*, *26*(19), 3001–3004. https://doi.org/10.1029/1999GL010434

Bhattacharyya, S., Sreekesh, S., & King, A. (2022). Characteristics of extreme rainfall in different gridded datasets over India during 1983–2015. *Atmospheric Research*, *267*, 105930. https://doi.org/10.1016/j.atmosres.2021.105930

Bohlinger, P., Sorteberg, A., & Sodemann, H. (2017). Synoptic conditions and moisture sources actuating extreme precipitation in Nepal. *Journal of Geophysical Research: Atmospheres*, *122*(23), 12,612-653,671. https://doi.org/10.1002/2017JD027543

Brown, S. J. (2018). The drivers of variability in UK extreme rainfall. *International Journal of Climatology*, *38*, e119–e130. https://doi.org/10.1002/joc.5356

Ceglar, A., Toreti, A., Balsamo, G., & Kobayashi, S. (2017). Precipitation over Monsoon Asia: A Comparison of Reanalyses and Observations. *Journal of Climate*, *30*(2), 465–476. https://doi.org/10.1175/JCLI-D-16-0227.1

Chen, J., Li, Z., Li, L., Wang, J., Qi, W., Xu, C.-Y., & Kim, J.-S. (2020). Evaluation of Multi-Satellite Precipitation Datasets and Their Error Propagation in Hydrological Modeling in a Monsoon-Prone Region. *Remote Sensing*, *12*(21), 3550. https://doi.org/10.3390/rs12213550

Chen, Y., Sharma, S., Zhou, X., Yang, K., Li, X., Niu, X., et al. (2021). Spatial performance of multiple reanalysis precipitation datasets on the southern slope of central Himalaya. *Atmospheric Research*, *250*(November 2020), 105365. https://doi.org/10.1016/j.atmosres.2020.105365

Chevuturi, A., Turner, A. G., Woolnough, S. J., Martin, G. M., & MacLachlan, C. (2019). Indian summer monsoon onset forecast skill in the UK Met Office initialized coupled seasonal forecasting system (GloSea5-GC2). *Climate Dynamics*, *52*(11), 6599–6617. https://doi.org/10.1007/s00382-018-4536-1

Dahri, Z. H., Moors, E., Ludwig, F., Ahmad, S., Khan, A., Ali, I., & Kabat, P. (2018). Adjustment of measurement errors to reconcile precipitation distribution in the high-altitude Indus basin. *International Journal of Climatology*, *38*(10), 3842–3860. https://doi.org/10.1002/joc.5539

Dahri, Z. H., Ludwig, F., Moors, E., Ahmad, S., Ahmad, B., Shoaib, M., et al. (2021). Spatio-temporal evaluation of gridded precipitation products for the high-altitude Indus basin. *International Journal of Climatology*, (February), joc.7073. https://doi.org/10.1002/joc.7073

Dalcher, A., Kalnay, E., & Hoffman, R. N. (1988). Medium Range Lagged Average Forecasts. *Monthly Weather Review*, *116*(2), 402–416. https://doi.org/10.1175/1520-0493(1988)116<0402:MRLAF>2.0.CO;2

Davis, P., Ruth, C., Scaife, A. A., & Kettleborough, J. (2020). A Large Ensemble Seasonal Forecasting System: GloSea6. In *AGU Fall Meeting Abstracts* (Vol. 2020, pp. A192-05). AA(Met Office Hadley Centre, Exeter, United Kingdom), AB(Met Office Hadley Centre, Exeter, United Kingdom), AC(Met Office Hadley Centre, Exeter, United Kingdom), AD(Met Office Hadley Centre, Exeter, United Kingdom).

Derin, Y., Anagnostou, E., Berne, A., Borga, M., Boudevillain, B., Buytaert, W., et al. (2019). Evaluation of GPM-era Global Satellite Precipitation Products over Multiple Complex Terrain Regions. *Remote Sensing*, *11*(24), 2936. https://doi.org/10.3390/rs11242936

Dunstone, N., Smith, D., & Hermanson, L. (2016). DePreSys3 – collaboration guide. Exeter: Met Office.

Dunstone, N., Smith, D., Scaife, A., Hermanson, L., Eade, R., Robinson, N., et al. (2016). Skilful predictions of the winter North Atlantic Oscillation one year ahead. *Nature Geoscience*, *9*(11), 809–814. https://doi.org/10.1038/ngeo2824

Dunstone, N., Smith, D., Scaife, A., Hermanson, L., Fereday, D., O'Reilly, C., et al. (2018). Skilful Seasonal Predictions of Summer European Rainfall. *Geophysical Research Letters*, *45*(7), 3246–3254. https://doi.org/10.1002/2017GL076337

Dunstone, N., Smith, D., Yeager, S., Danabasoglu, G., Monerie, P.-A., Hermanson, L., et al. (2020). Skilful interannual climate prediction from two large initialised model ensembles. *Environmental Research Letters*, *15*(9), 94083. https://doi.org/10.1088/1748-9326/ab9f7d

Funk, C., Peterson, P. J., Landsfeld, M. F., Pedreros, D. H., Verdin, J. P., Rowland, J. D., et al. (2014). A Quasi-Global Precipitation Time Series for Drought Monitoring. *U.S. Geological Survey Data Series*, *832*, 4. https://doi.org/110.3133/ds832

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., et al. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, *2*, 150066. https://doi.org/10.1038/sdata.2015.66

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Hamm, A., Arndt, A., Kolbe, C., Wang, X., Thies, B., Boyko, O., et al. (2020). Intercomparison of

Gridded Precipitation Datasets over a Sub-Region of the Central Himalaya and the Southwestern Tibetan Plateau. *Water*, *12*(11), 3271. https://doi.org/10.3390/w12113271

Hastie, T. J., & Tibshirani, R. J. (2017). *Generalized additive models*. Routledge.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., et al. (2018). ERA5 hourly data on pressure levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.bd0915c6

Hoffman, R. N., & Kalnay, E. (1983). Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus A*, *35A*(2), 100–118. https://doi.org/10.1111/j.1600-0870.1983.tb00189.x

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., et al. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology*, *8*(1), 38–55. https://doi.org/10.1175/JHM560.1

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., et al. (2020). NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). *Algorithm Theoretical Basis Document (ATBD) Version 06*, 35.

Jain, S., Scaife, A. A., Dunstone, N., Smith, D., & Mishra, S. K. (2020). Current chance of unprecedented monsoon rainfall over India using dynamical ensemble simulations. *Environmental Research Letters*, *15*(9), 094095. https://doi.org/10.1088/1748-9326/ab7b98

Janowiak, J. E., Joyce, R. J., & Yarosh, Y. (2001). A Real–Time Global Half–Hourly Pixel–Resolution Infrared Dataset and Its Applications. *Bulletin of the American Meteorological Society*, *82*(2), 205–217. https://doi.org/10.1175/1520-0477(2001)082<0205:ARTGHH>2.3.CO;2

Jayakumar, A., Turner, A. G., Johnson, S. J., Rajagopal, E. N., Mohandas, S., & Mitra, A. K. (2017). Boreal summer sub-seasonal variability of the South Asian monsoon in the Met Office GloSea5 initialized coupled model. *Climate Dynamics*, *49*(5–6), 2035–2059. https://doi.org/10.1007/s00382-016-3423-x

Johnson, S. J., Turner, A., Woolnough, S., Martin, G., & MacLachlan, C. (2017). An assessment of Indian monsoon seasonal forecasts and mechanisms underlying monsoon interannual variability in the Met Office GloSea5-GC2 system. *Climate Dynamics*, *48*(5–6), 1447–1465. https://doi.org/10.1007/s00382-016-3151-2

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., et al. (2019). SEAS5: the new ECMWF seasonal forecast system. *Geoscientific Model Development*, *12*(3), 1087–1117. https://doi.org/10.5194/gmd-12-1087-2019

Jones, P. W. (1999). First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates. *Monthly Weather Review*, *127*(9), 2204–2210. https://doi.org/10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2

Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *Journal of Hydrometeorology*, *5*(3), 487–503. https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., et al. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, *4*(1), 170122. https://doi.org/10.1038/sdata.2017.122

Karki, R., Hasson, S. ul, Schickhoff, U., Scholten, T., & Böhner, J. (2017). Rising Precipitation Extremes across Nepal. *Climate*, *5*(1), 4. https://doi.org/10.3390/cli5010004

Kelder, T., Müller, M., Slater, L. J., Marjoribanks, T. I., Wilby, R. L., Prudhomme, C., et al. (2020). Using UNSEEN trends to detect decadal changes in 100-year precipitation extremes. *Npj Climate and Atmospheric Science*, *3*(1), 47. https://doi.org/10.1038/s41612-020-00149-4

Kent, C., Dunstone, N., Tucker, S., Scaife, A. A., Brown, S., Kendon, E. J., et al. (2022). Estimating unprecedented extremes in UK summer daily rainfall. *Environmental Research Letters*, *17*(1), 014041. https://doi.org/10.1088/1748-9326/ac42fb

Knapp, K. R., Ansari, S., Bain, C. L., Bourassa, M. A., Dickinson, M. J., Funk, C., et al. (2011). Globally Gridded Satellite Observations for Climate Studies. *Bulletin of the American Meteorological Society*, *92*(7), 893–907. https://doi.org/10.1175/2011BAMS3039.1

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, *93*(1), 5–48. https://doi.org/10.2151/jmsj.2015-001

Krakauer, N., Pradhanang, S., Lakhankar, T., & Jha, A. (2013). Evaluating Satellite Products for Precipitation Estimation in Mountain Regions: A Case Study for Nepal. *Remote Sensing*, *5*(8), 4107–4123. https://doi.org/10.3390/rs5084107

Lai, S., Xie, Z., Bueh, C., & Gong, Y. (2020). Fidelity of the APHRODITE Dataset in Representing Extreme Precipitation over Central Asia. *Advances in Atmospheric Sciences*, *37*(12), 1405–1416. https://doi.org/10.1007/s00376-020-0098-3

Lange, S. (2019). Near surface meteorological variables from 1979 to 2018 derived from bias-corrected reanalysis (WFDE5) v1.1. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.20d54e34

Li, D., Yang, K., Tang, W., Li, X., Zhou, X., & Guo, D. (2020). Characterizing precipitation in high altitudes of the western Tibetan plateau with a focus on major glacier areas. *International Journal of Climatology*, *40*(12), 5114–5127. https://doi.org/10.1002/joc.6509

Lim, S.-M., Hyun, Y.-K., Kang, H.-S., & Yeh, S.-W. (2018). Prediction Skill of East Asian Precipitation and Temperature Associated with El Niño in GloSea5 Hindcast Data. *Atmosphere. Korean Meteorological Society*, *28*(1), 37–51. https://doi.org/10.14191/Atmos.2018.28.1.037

Lu, B., Scaife, A. A., Dunstone, N., Smith, D., Ren, H.-L., Liu, Y., & Eade, R. (2017). Skillful seasonal predictions of winter precipitation over southern China. *Environmental Research Letters*, *12*(7), 074021. https://doi.org/10.1088/1748-9326/aa739a

Ma, Z., Xu, J., Zhu, S., Yang, J., Tang, G., Yang, Y., et al. (2020). AIMERG: a new Asian precipitation dataset (0.1°/half-hourly, 2000–2015) by calibrating the GPM-era IMERG at a daily scale using APHRODITE. *Earth System Science Data*, *12*(3), 1525–1544. https://doi.org/10.5194/essd-12-1525-2020

MacLachlan, C., Arribas, A., Peterson, K. A., Maidens, A., Fereday, D., Scaife, A. A., et al. (2015). Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system. *Quarterly Journal of the Royal Meteorological Society*, *141*(689), 1072–1084. https://doi.org/10.1002/qj.2396

Maussion, F., Scherer, D., Mölg, T., Collier, E., Curio, J., & Finkelnburg, R. (2014). Precipitation Seasonality and Variability over the Tibetan Plateau as Resolved by the High Asia Reanalysis*. *Journal of Climate*, *27*(5), 1910–1927. https://doi.org/10.1175/JCLI-D-13-00282.1

McGraw, K. O., & Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin*, *111*(2), 361–365. https://doi.org/10.1037/0033-2909.111.2.361

Mondal, A., Lakshmi, V., & Hashemi, H. (2018). Intercomparison of trend analysis of Multisatellite Monthly Precipitation Products and Gauge Measurements for River Basins of India. *Journal of Hydrology*, *565*, 779–790. https://doi.org/10.1016/j.jhydrol.2018.08.083

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., et al. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data Discussions*, *13*, 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Nandargi, S., & Dhar, O. N. (2011). Extreme rainfall events over the Himalayas between 1871 and 2007. *Hydrological Sciences Journal*, *56*(6), 930–945. https://doi.org/10.1080/02626667.2011.595373

Nepal, B., Shrestha, D., Sharma, S., Shrestha, M. S., Aryal, D., & Shrestha, N. (2021). Assessment of GPM-Era Satellite Products' (IMERG and GSMaP) Ability to Detect Precipitation Extremes over Mountainous Country Nepal. *Atmosphere*, *12*(2), 254. https://doi.org/10.3390/atmos12020254

Nguyen, P.-L., Bador, M., Alexander, L. V., Lane, T. P., & Funk, C. C. (2020). On the Robustness of Annual Daily Precipitation Maxima Estimates Over Monsoon Asia. *Frontiers in Climate*, *2*. https://doi.org/10.3389/fclim.2020.578785

Nogueira, M. (2020). Inter-comparison of ERA-5, ERA-interim and GPCP rainfall over the last 40 years: Process-based analysis of systematic and random differences. *Journal of Hydrology*, *583*(January), 124632. https://doi.org/10.1016/j.jhydrol.2020.124632

Palazzi, E., von Hardenberg, J., & Provenzale, A. (2013). Precipitation in the Hindu-Kush Karakoram Himalaya: Observations and future scenarios. *Journal of Geophysical Research: Atmospheres*, *118*(1), 85–100. https://doi.org/10.1029/2012JD018697

Pant, G. B., & Parthasarathy, S. B. (1981). Some aspects of an association between the southern oscillation and indian summer monsoon. *Archives for Meteorology, Geophysics, and Bioclimatology Series B*, *29*(3), 245–252. https://doi.org/10.1007/BF02263246

Papalexiou, S. M., & Koutsoyiannis, D. (2013). Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, *49*(1), 187–201. https://doi.org/10.1029/2012WR012557

Pritchard, D. M. W., Forsythe, N., Fowler, H. J., O'Donnell, G. M., & Li, X.-F. (2019). Evaluation of Upper Indus Near-Surface Climate Representation by WRF in the High Asia Refined Analysis. *Journal of Hydrometeorology*, *20*(3), 467–487. https://doi.org/10.1175/JHM-D-18-0030.1

Ragulina, G., & Reitan, T. (2017). Generalized extreme value shape parameter and its nature for extreme precipitation using long time series and the Bayesian approach. *Hydrological Sciences Journal*, *62*(6), 863–879. https://doi.org/10.1080/02626667.2016.1260134

Ramanadham, R., Visweswara Rao, P., & Patnaik, J. K. (1973). Break in the Indian summer monsoon. *Pure and Applied Geophysics PAGEOPH*, *104*(1), 635–647. https://doi.org/10.1007/BF00875908

Rani, S. I., T, A., George, J. P., Rajagopal, E. N., Renshaw, R., Maycock, A., et al. (2021). IMDAA: High Resolution Satellite-era Reanalysis for the Indian Monsoon Region. *Journal of Climate*, 1–78. https://doi.org/10.1175/JCLI-D-20-0412.1

Richardson, K. J. (2021). Extreme precipitation events in South Asia: understanding climate drivers

through case studies. Exeter, UK: Met Office.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., et al. (2010). The NCEP Climate Forecast System Reanalysis. *Bulletin of the American Meteorological Society*, *91*(8), 1015–1058. https://doi.org/10.1175/2010BAMS3001.1

Scaife, A. A., Camp, J., Comer, R., Davis, P., Dunstone, N., Gordon, M., et al. (2019). Does increased atmospheric resolution improve seasonal climate predictions? *Atmospheric Science Letters*, *20*(8), e922. https://doi.org/10.1002/asl.922

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., & Ziese, M. (2011). GPCC Full Data Reanalysis Version 6.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. https://doi.org/10.5676/DWD_GPCC/FD_M_V7_050

Schneider, U., Finger, P., Meyer-Christoffer, A., Rustemeier, E., Ziese, M., & Becker, A. (2017). Evaluating the Hydrological Cycle over Land Using the Newly-Corrected Precipitation Climatology from the Global Precipitation Climatology Centre (GPCC). *Atmosphere*, *8*(12), 52. https://doi.org/10.3390/atmos8030052

Schulzweida, U. (2020, October). CDO User Guide. https://doi.org/10.5281/zenodo.4246983

Sharma, S., Chen, Y., Zhou, X., Yang, K., Li, X., Niu, X., et al. (2020). Evaluation of GPM-Era Satellite Precipitation Products on the Southern Slopes of the Central Himalayas Against Rain Gauge Data. *Remote Sensing*, *12*(11), 1836. https://doi.org/10.3390/rs12111836

Sheffield, J., Goteti, G., & Wood, E. F. (2006). Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling. *Journal of Climate*, *19*(13), 3088–3111. https://doi.org/10.1175/JCLI3790.1

Shrestha, N. K., Qamer, F. M., Pedreros, D., Murthy, M. S. R., Wahid, S. M., & Shrestha, M. (2017). Evaluating the accuracy of Climate Hazard Group (CHG) satellite rainfall estimates for precipitation based drought monitoring in Koshi basin, Nepal. *Journal of Hydrology: Regional Studies*, *13*, 138–151. https://doi.org/10.1016/j.ejrh.2017.08.004

Sikka, D. R. (1980). Some aspects of the large scale fluctuations of summer monsoon rainfall over India in relation to fluctuations in the planetary and regional scale circulation parameters. *Journal of Earth System Science*, *89*(2), 179–195. https://doi.org/10.1007/BF02913749

Singh, R., Kishtawal, C. M., & Singh, C. (2021). The Strengthening Association Between Siberian Snow and Indian Summer Monsoon Rainfall. *Journal of Geophysical Research: Atmospheres*, *126*(9), e2020JD033779. https://doi.org/10.1029/2020JD033779

Singh, T., Saha, U., Prasad, V. S., & Gupta, M. Das. (2021). Assessment of newly-developed high resolution reanalyses (IMDAA, NGFS and ERA5) against rainfall observations for Indian region. *Atmospheric Research*, *259*, 105679. https://doi.org/10.1016/j.atmosres.2021.105679

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Zhiquan, L., Berner, J., et al. (2019). A Description of the Advanced Research WRF Model Version 4. *NCAR Technical Note NCAR/TN-475+STR*, 145.

Stacey, J., Richardson, K., Krjinen, J., & Janes, T. (2019). Seasonal Forecasting in South Asia: A Review of the Current Status. *Asia Regional Resilience to a Changing Climate (ARRCC) Work Package 2: Strengthening Climate Information Partnerships – South Asia (SCIPSA)*. Exeter, UK: Met Office.

Stacey, J., Bett, P., Colledge, F., Colman, A., Daron, J., Graham, R., et al. (2021). Skill of South Asian

Precipitation Forecasts in Multiple Seasonal Prediction Systems. *Asia Regional Resilience to a Changing Climate (ARRCC) Work Package 2: Strengthening Climate Information Partnerships – South Asia (SCIPSA)*. Exeter, UK: Met Office.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K. (2018). A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. *Reviews of Geophysics*, *56*(1), 79–107. https://doi.org/10.1002/2017RG000574

Tan, X., Ma, Z., He, K., Han, X., Ji, Q., & He, Y. (2020). Evaluations on gridded precipitation products spanning more than half a century over the Tibetan Plateau and its surroundings. *Journal of Hydrology*, *582*, 124455. https://doi.org/10.1016/j.jhydrol.2019.124455

Thompson, V., Dunstone, N. J., Scaife, A. A., Smith, D. M., Slingo, J. M., Brown, S., & Belcher, S. E. (2017). High risk of unprecedented UK rainfall in the current climate. *Nature Communications*, *8*(1), 1–6. https://doi.org/10.1038/s41467-017-00275-3

Ullah, W., Wang, G., Ali, G., Tawia Hagan, D., Bhatti, A., & Lou, D. (2019). Comparing Multiple Precipitation Products against In-Situ Observations over Different Climate Regions of Pakistan. *Remote Sensing*, *11*(6), 628. https://doi.org/10.3390/rs11060628

Wang, X., Tolksdorf, V., Otto, M., & Scherer, D. (2021). WRF-based dynamical downscaling of ERA5 reanalysis data for High Mountain Asia: Towards a new version of the High Asia Refined analysis. *International Journal of Climatology*, *41*(1), 743–762. https://doi.org/10.1002/joc.6686

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R* (2nd ed.). Chapman and Hall/CRC.

Xie, P., Joyce, R., Wu, S., Yoo, S.-H., Yarosh, Y., Sun, F., & Lin, R. (2019). NOAA Climate Data Record (CDR) of CPC Morphing Technique (CMORPH) High Resolution Global Precipitation Estimates, Version 1. NOAA National Centers for Environmental Information. https://doi.org/10.25921/w9va-q159

Xu, R., Tian, F., Yang, L., Hu, H., Lu, H., & Hou, A. (2017). Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network. *Journal of Geophysical Research: Atmospheres*, *122*(2), 910–924. https://doi.org/10.1002/2016JD025418

Yamamoto, M. K., Ueno, K., & Nakamura, K. (2011). Comparison of Satellite Precipitation Products with Rain Gauge Data for the Khumb Region, Nepal Himalayas. *Journal of the Meteorological Society of Japan. Ser. II*, *89*(6), 597–610. https://doi.org/10.2151/jmsj.2011-601

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., & Kitoh, A. (2012). APHRODITE: Constructing a Long-Term Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges. *Bulletin of the American Meteorological Society*, *93*(9), 1401–1415. https://doi.org/10.1175/BAMS-D-11-00122.1

Yatagai, A., Maeda, M., Khadgarai, S., Masuda, M., & Xie, P. (2020). End of the Day (EOD) Judgment for Daily Rain-Gauge Data. *Atmosphere* . https://doi.org/10.3390/atmos11080772

Youngman, B. D. (2020). evgam: Generalised Additive Extreme Value Models.

Zhang, Y., & Li, J. (2016). Impact of moisture divergence on systematic errors in precipitation around the Tibetan Plateau in a general circulation model. *Climate Dynamics*, *47*(9–10), 2923–2934. https://doi.org/10.1007/s00382-016-3005-y

Met Office
FitzRoy Road
Exeter
Devon
EX1 3PB
United Kingdom