"The climatic water balance captures evolving water resources pressures on the margins of the Himalaya"

Authors: Forsythe, Nathan D[1] [ORCID: 0000-0002-4593-8233]; Tiwari, Prakash Chandra[2]; Pritchard, David M W[1] [ORCID: 0000-0002-9215-7210]; Walker, David W.[1,3] [ORCID: 0000-0002-2486-4677], Joshi, Bhagwati[2]; Fowler, Hayley J[1] [ORCID: 0000-0001-8848-3606];

Affiliations:

[1] Water resources research group, School of Engineering, Newcastle University, United Kingdom

[2] Department of Geography, Kumaun University, India

[3] Wageningen University & Research, the Netherlands

<abstract>

Evaluation of the climatic water balance (CWB) – i.e. precipitation minus potential evapotranspiration – has strong potential as a tool for investigating patterns of variability and change in the water cycle since it estimates the (im)balance of atmospheric moisture near the land surface. Using observations from a middle-Himalaya weather station at Mukteshwar (29.474°N, 79.646°E, Uttarakhand state) in India, we demonstrate a CWB-based set of analytical procedures can robustly characterise local climate variability. Use of the CWB circumvents uncertainties in the soil water balance stemming from limited data on subsurface properties. We also focus on three key input variables used to calculate the CWB: precipitation, mean temperature and diurnal temperature range. We use local observations to evaluate the skill of gridded datasets –specifically meteorological reanalyses – in representing local conditions. Reanalysis estimates of Mukteshwar climate showed large absolute biases but accurately captured the timing and relative amplitude of the annual cycle of these three variables and the CWB. This suggests that the reanalyses can provide insight regarding climate processes in data-sparse regions, but caution is necessary if extracting absolute values. While the local observations at Mukteshwar show clear annual cycles and substantial interannual variability, results from investigation of their time-dependency were quite mixed. Pragmatically this implies that while "change is coming, variability is now." If communities can adapt to the observed historical hydroclimate variability they will have built meaningful adaptive capacity to cope with on-going environmental change. This follows a 'low regret' approach advocated in the face of a substantially uncertain future.

MAIN TEXT

[1] Introduction

[1.1] A conceptual framework for understanding the changing water cycle

When addressing the question of how the water cycle, in a specific location or region, has changed in recent decades, and how it may change in the future, the conceptual framing of the question will guide the response (Milly et al., 2005; Huntington, 2006; Oki and Kanae, 2006, Sheffield and Wood, 2008; Trenberth et al., 2014). For human activities and terrestrial ecology, the partitioning of precipitation between infiltration and runoff is of preponderant importance, because the path water takes to return either to the atmosphere, via evapotranspiration, or to the sea, via stream networks, has great influence on crop production, natural vegetation cover, water supply and freshwater ecosystems. While the key determinant of partitioning is precipitation intensity (rainfall rate), this is modulated by surface characteristics including slope, land cover (permeability) and underlying soil properties (porosity, hydraulic conductivity). These characteristics can vary greatly over short distances, and many catchments, including the focus catchment, and particularly those with substantial human activities, exhibit high degrees of heterogeneity. Where available, detailed spatially comprehensive information on catchment surface characteristics enables the use of precipitation and evapotranspiration data to calculate the soil moisture balance. This is needed to estimate moisture available to meet water requirements of crops and natural vegetation as well as quantifying contributions to groundwater recharge and stream baseflow.

Unfortunately, information on surface characteristics, especially soil properties, is rarely available with sufficient spatial granularity to enable skilled calculation of the soil moisture balance over substantial areas (Grunwald, 2009), unless available river discharge measurements and/or groundwater level observations enable back-calculation of spatially aggregated runoff-infiltration partitioning. Alternatively, the climatic water balance (CWB), i.e. the net quantity of precipitation minus potential (or reference) evapotranspiration, can be evaluated almost everywhere and with relative confidence, particularly if drawing upon gridded datasets such as global meteorological reanalyses. At monthly and longer timescales, the CWB provides a strong indicator of relative moisture abundance or shortfall and is useful for evaluating stresses on, and the potential of forestry and rainfed agriculture for, specific crops and regions (Sharma *et al.*, 2010; Crimmins *et al.*, 2011; Churchill *et al.*, 2013). These stresses are of preponderant concern because, with the exception of high-latitude and high-elevation contexts, moisture rather than energy will be the limiting constraint on plant development through transpiration (Jung et al, 2010) and hence ecosystem benefits and food production.

Furthermore, potential evapotranspiration (PET: Thornthwaite, 1948; Hargreaves, 1994) can be parameterised with reasonable skill from simply daily mean temperature (T_{avg}) and diurnal temperature range (DTR) (Droogers and Allen, 2002; Hargreaves and Allen, 2003). Thus, together with precipitation, the CWB can be determined from three readily observed climate variables. From a purely meteorological standpoint, these three variables together succinctly summarise prevailing weather conditions: dry versus wet, warm versus cold, and clear (high DTR) versus overcast (low DTR) skies. This is reflected in tools such as the RainSim-CRU Weather Generator (Burton *et al.*, 2009; Kilsby *et al.*, 2007) for synthetic time-series generation and stochastic downscaling of climate projections. However, PET can be better estimated by more complex formulae derived from physical principles, e.g. the Penman-Monteith equation (Monteith, 1965) requires net radiation, humidity and windspeed data in addition to temperature, along with parameterisations representing aerodynamic and surface resistances to fluxes. Unfortunately, in many areas where assessment of water availability is required, formal meteorological observations are lacking due to limited density of national monitoring networks. Formal measurements of humidity – as dewpoint temperature, relative humidity or vapour pressure – and windspeed are not as widely available as temperature and precipitation observations. Observations of radiation components (shortwave, longwave) are even more rare. In these cases global meteorological reanalyses provide a promising data source as they assimilate not only available regional surface observations but also a portfolio of other inputs including radiosonde measurements and satellite imagery. Numerical tools and forecasting models then synthesise spatially continuous, physically consistent estimates of climate variables both at the surface and upward through the atmosphere, but these are biased in absolute values compared to observations, particularly in regions of high topographic variability, where elevation biases also play a role.

Changes in the CWB, itself a metric of moisture surplus or deficit, provide a first order indication of whether moisture is tending to become more abundant (CWB increase) or scarce (CWB decrease). These changes – be they increasing surpluses, aggravated deficits or a tendency toward equilibrium – result from increases or decreases in atmospheric supply (precipitation) and demand (potential/reference evapotranspiration) of moisture at the land surface. Thus the individual causal mechanisms of changes in precipitation and (surface) energy – indexed by T_{avg} and DTR – are of great interest. Furthermore, understanding the role of distinct climate processes – such as surface energy balance modulation by cloud radiative effects – as causes of these changes can provide qualitative context to better anticipate likely future CWB evolution and to objectively evaluate available climate model outputs which provide quantitative projections of this evolution. Using Mukteshwar as a case study, the present work advances a framework analytical methodology for addressing these issues at the 'point' (single-site) scale at which a great many scientists and technical professionals will be working to understand the evolution of the hydrological cycle and its implication for interdependent human and natural systems.

[1.2] Case study context

Situated in the 'middle upper reaches' of the Ganges basin, the small headwater sub-catchments of the Kosi river rising from the Gaula and Almora ranges of the Kumaun Lesser Himalaya (KLH) are critical water resources units at both micro and macro scales. These sub-catchments provide valuable insights regarding potential pathways for sustainable resilience to hydroclimate variability. With complex agro-forestry land cover patterns and surface elevations ranging from ~1000m to ~2300m above sea level (asl), these catchments experience a (primarily) subtropical/monsoonal precipitation regime and support multiple crop growing seasons each year. While annual rainfall is sufficient for substantial agricultural production, these catchments also generate important surface runoff (and baseflow) for downstream segments of the middle and lower Ganges basin. This latter area along with the Punjab (in both India and Pakistan) serves as the 'bread basket' of South Asia, encompassing the majority of the region's irrigated farmland and underpinning its food security (Rahaman, 2009). This paper explores potential pressures on local water resources and food security in the KLH due to evolution of the local water cycle through CWB-focused analysis of historical observations from the Mukteshwar meteorological station in Uttarakhand state, India (Figure 1). This station is located on a ridgeline overlooking two headwaters catchments – Ramgad and Dhokane – of the Kosi river tributary to the Ganges.

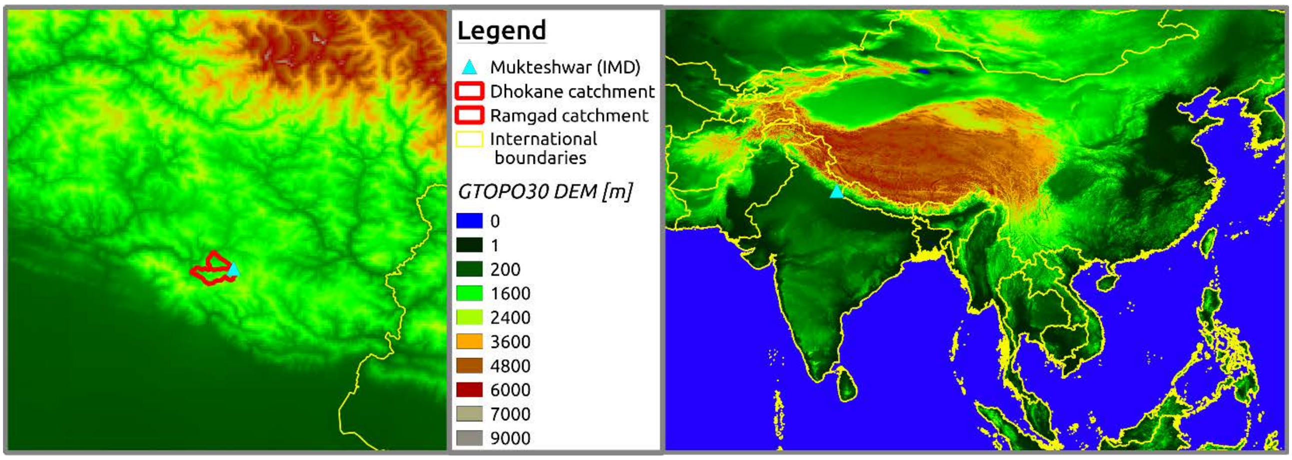

Figure 1: Study area geographical context showing location of Mukteshwar meteorological station (29.474°N, 79.646°E, Uttarakhand state) operated by India Meteorological Department (IMD) in relation to surface elevation and international boundaries in Asia. The left panel shows detail of the Kumaon division of Uttarakhand state while the right panel shows the broader Asian continental context.

[2] <u>Data and Methods</u>

[2.1] Data

[2.1.1] *Local climate observations: IMD Mukteshwar*

The weather observation station at Mukteshwar (29.474°N, 79.646°E) – currently operated by the India Meteorological Department (IMD) – was established in 1897. Along with precipitation, daily maximum and minimum temperature observations (beginning in 1969) were made available by IMD personnel for use in this study. In the absence of sub-daily observations, daily mean temperature was approximated as the mean of recorded daily maximum and minimum. There was an interruption in temperature data recording from September 1993 through August 1997. This study also lacks access to observations of all variables during 2015, with the exception of December of that year. A double mass check with temperature data from New Delhi, accessed via the Global Historical Climate Network dataset (Lawrimore *et al.*, 2011), however, reveals no slope 'break points'. This result mitigates concerns regarding step changes or inhomogeneity in temperature measurements and lends confidence to the results presented in this paper. Precipitation records at Mukteshwar are far more complete with a mean total fraction of missing observations of 4.3% as compared to 14.5% for temperature (Table 1). This study focuses on a common analytic time period of 1980 through 2018 (as complete calendar years).

Table 1 Missing** daily observations from Mukteshwar IMD station, by fraction of record for individual months, 1980 to 2018

Variable	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Mean
Precipitation	0.044	0.049	0.046	0.051	0.057	0.041	0.039	0.044	0.040	0.041	0.045	0.021	0.043
Temperature	0.147	0.155	0.144	0.151	0.154	0.138	0.140	0.140	0.154	0.142	0.149	0.128	0.145

** observations available to this study. There is a period of 11 months in 2015 from January through November where observations were made as demonstrated by inclusion in GHCN-Monthly (v2 for precipitation, v3 for temperature).

[2.1.2] *Global meteorological reanalyses*

Global meteorological reanalyses ingest vast quantities of climate observations ranging from ocean buoys through ground-based measurements, to atmospheric soundings and satellite imagery. They are produced by leading weather/climate forecasting institutes and serve a range of purposes (Bosilovich et al., 2008, Lorentz and Kunstmann, 2012; Vose et al., 2012). For their producers reanalyses projects offer an opportunity to test updates to their data assimilation and weather forecasting systems. For the broader scientific community, reanalyses offer ‘gap free’, i.e. spatiotemporally continuous, estimates of a broad range of climate variables at levels ranging from the ground (or sea) surface to the upper (‘top of’) atmosphere.

Variable estimates from reanalyses are generally grouped in two broad categories: i) analytical outputs which include ‘state’ variables (temperature, humidity, wind speed, etc.) estimated using the data assimilation schemes/components of forecasting systems; b) forecast outputs which include fluxes (precipitation, radiation, etc) estimated using the forecast models themselves. The analytical methods utilised in reanalysis projects are guided by physical processes/relationships. Therefore, their outputs can avoid the potentially spurious results found in ‘observational’ gridded datasets which attempt to fill voids over sparsely observed regions through purely geostatistical techniques. This study utilised data from four independent reanalyses: a) ERA-Interim (Dee et al., 2011) produced by the European Centre for Medium Range Weather Forecasting (ECMWF), b) JRA-55 (Ebita et al., 2011) produced by the Japan Meteorological Agency (JMA), c) MERRA2 (Rienecker et al., 2011) produced by NASA and d) ERA5 (Hersbach et al, 2020) also produced by ECMWF. Key differentiating characteristics of each of the reanalyses are presented in Table 2.

Table 2 Global meteorological reanalyses

Reanalysis	Producer	Start date	Latitude resolution	Longitude resolution	Analytical/synoptic time-step
ERA-Interim	ECMWF	01/01/1979	0.75°	0.75°	6 hours
JRA-55	JMA	01/01/1958	1.25°	1.25°	3 hours
MERRA2	NASA	01/01/1980	0.50°	0.625°	Hourly
ERA5	ECMWF	01/01/1979	0.25°	0.25°	Hourly

[2.2] Methods

[2.2.1] *Calculation of CWB from supply and demand components=*

In the absence of multi-decadal local hydrological observational records – and the detailed local soil characteristic descriptions needed to calculate the soil moisture balance – we focused on the climatic water balance (CWB) as the core indicator of water availability in the KLH in the vicinity of Mukteshwar. In the CWB the atmospheric moisture demand component is represented by potential, or reference, evaporation. For a given set of weather conditions PET quantifies the amount of moisture which, if available, would be transferred to the atmosphere from the land surface, including vegetation (Thornthwaite, 1948). A wide range of equations exist for calculating PET. Here we adopted the United Nations Food and Agriculture Organisation (FAO) Penman Monteith method for calculating reference evapotranspiration (ET_0) (see Allen *et al.*, 1998) as it is a well-established approach with relatively flexible input data requirements: net radiation, humidity and windspeed data in addition to temperature. The equation also uses parameterisations representing aerodynamic and surface resistances to fluxes which vary based on a range of factors including vegetation height. This is based on resistance associated with a 'reference crop', specifically a "well-watered grass 12cm tall" to facilitate both spatiotemporal comparisons and extrapolations to various important crops (through use of coefficients). The approach of calculating a reference from which the potential water requirements of specific crops can be quickly estimated is particularly useful in farming systems such as those used by smallholders in the geographic focus of the study, i.e. the Kumaun Himalaya around Mukteshwar, where a wide range of vegetables and legumes are cultivated.

To calculate the reference evapotranspiration (ET_0) local observations of daily rainfall, minimum and maximum temperature were paired with ensemble mean estimates for the overlying grid cell from the four reanalyses – ERA-Interim, JRA-55, NASA MERRA2 and ERA5 – for radiation, wind speed and relative humidity. These ensemble estimates were made by extracting daily (mean) time-series from the relevant grid cell of each individual reanalysis. Without ground-based data to validate or characterise bias in reanalysis data, a simple ensemble averaging approach was adopted to obtain (reasonable) central estimates.

We also calculated daily estimates for ET_0 directly for each reanalysis ensemble member using its own values for input variables in the grid cell overlying Mukteshwar. This allows us to compare CWB results using the maximum available local observations to estimates purely derived from global gridded datasets.

[2.2.2] *Climatological characterisations and time-series analyses*

Climatological characterisation was approached as statistical (mean, quantiles) description of the annual cycle at a monthly time-step. The use of local observations and global meteorological reanalyses at very different spatial scales requires comparison not only of absolute values but also in relative terms as the large-scale reanalyses are unlikely to provide absolute value matches to local observations in regions of high topographic variability such as Uttarakhand/the Kumaun Himalaya where there is a steep transition from plains to high mountains. We therefore applied simple normalisations to both the gauge and reanalysis data: a) for zero-bounded 'accumulating' variables (precipitation, reference evapotranspiration, net radiation) we normalised the monthly mean and quantiles of individual data sources by dividing absolute values by the period annual mean; b) for 'state' variables (temperature, humidity, wind speed, CWB) we normalised the monthly mean and quantiles of individual data sources by subtracting the annual period mean from absolute values then dividing the result by the amplitude, i.e. maximum period monthly mean minus minimum period monthly mean. This specific normalisation method – as a opposed to a the more widely used standardisation method of subtracting the (period monthly) mean and dividing by the standard deviation – was used to preserve the form (shape) of the annual cycle in order to assess if gridded datasets with strong absolute biases might still provide some useful information content by accurately capturing the interplay of dominant climatic processes and forcings throughout the year.

Time-series analyses were performed to examine changes in CWB and its drivers over the record period. For time-series analyses: a) monthly means/totals were calculated if a minimum of 24 days (~80%) were available; b) annual aggregates of seasonal values were calculated only if all months concerned had met the aggregation criteria for calculation of valid mean/total values, i.e. sufficient daily observations. We used an alternate approach to the standard "p-value" for quantifying the probability of random occurrence of values of specific correlation or trend metrics. This deviation from standard procedure was inspired by recent thinking of Serinaldi *et al.* (2018) that challenges the validity of null hypothesis significance tests (NHSTs) for assessment of long-term patterns in hydro-climatological time series. Serinaldi asserts specifically that "*NHSTs have a logically flawed rationale coming from ill-posed and theoretically unfounded hybridization of Fisher significance tests and Neyman-Pearson hypothesis tests; they do not provide the in-formation that scientists need (i.e., the* *likelihood of H_0 given the data and/or physical significance), do not allow conclusions about the truth of falsehood of any hypothesis, and do not apply to exploratory non-randomized studies…* " The alternate method -- which still utilises the correlation assessment component of the null hypothesis approach -- conserves the observed values for a given variable but randomises ('shuffling') their sequence a large number of times ($n=1x10^6$) to provide a large sample of chaotic/quasi-natural variability. This method is similar to that utilised by Guerreiro et al (2018) to assess whether observed changes in sub-daily precipitation intensity exceed those which might occur through random/natural variability. Each synthetic sample member was tested against the potential causal factor – e.g. time, cloud cover – and the statistical distribution of resultant correlation strengths/trend rates were sampled to identify values corresponding to chances of 'random' (chaotic) occurrence. This method assumes that the observed series of values of a given variable represent a sample of physical plausible "real" values, but their specific sequencing could be the result of natural variability or driven by some strong causal factor. We use this approach to robustly evaluate the likelihood of the correlation (trend rate), indicated by an observed sequence, occurring through natural variability.

[3] <u>Results</u>

We now present the results of characterising the CWB – the climatology of its constituent elements, their temporal variability and evaluation of the potential drivers of this variability – using gauge observations from the Mukteshwar IMD station and equivalent reanalyses estimates. Because the CWB quantifies near-surface atmospheric moisture surplus/deficit status it helps us to understand the water cycle at Mukteshwar. This includes water cycle changes in recent decades, along with their potential causes. This work also demonstrates the utility of the single-site (point-based) CWB approach for characterising climate drivers of water resources in focused geographic areas. The inter-comparison of local observations to meteorological reanalyses further provides insight on the potential to extract useful CWB characterisations in data-sparse regions.

[3.1] Climatologies of individual variables

[3.1.1] *Climatologies of primary (precipitation) and secondary (temperature) variables*

The gauge observations in Figure 2 indicate that Mukteshwar has a strongly monsoonal precipitation regime: roughly 70% of annual precipitation falls in June through September. Due to its high surface elevation at ~2200m asl, the annual cycle/range of (daily) mean near surface air temperature (T_{avg}) exhibits a large amplitude more typical of temperate latitude zones, with the hottest month more than 10°C warmer than the coldest month. The annual cycle of diurnal temperature range (DTR) shows influence of both incoming (top of atmosphere) solar radiation and seasonal cloud cover with relative DTR maxima in the pre- and post-monsoon seasons and annual minimum during the monsoon. In addition to period mean conditions, Figure 2 also shows interannual variability quantified as the 10th and 90^{th} percentiles of the period distribution, i.e. values for a given calendar month from 1980 to 2018. Precipitation logically shows larger absolute variability, expressed as this 10^{th} to 90^{th} percentile range, during the monsoon than in the drier seasons. Year to year variability of monthly mean (daily) temperature is greater in winter and the pre-monsoon (Dec to June), with 10^{th} to 90^{th} percentile ranges of roughly 5°C, than during the monsoon and autumn (July to Nov), with ranges of roughly 2°C. Interannual variability of DTR is greatest in the early monsoon (June/July) and least in the late autumn (Nov/Dec).

The normalised climatogies of these three variables reveal that the reanalyses have strong skill in the (monthly) timing and amplitude of annual cycles (Figure 2, bottom row). For T_{avg} in particular, the contrast of the absence of relative bias with the very large absolute bias can be explained in part by the study area location on the fringe of the Himalaya and by the coarse spatial resolution of the reanalyses. Depending upon the precise position of grid cell boundaries in the individual reanalyses, the grid cell overlying Mukteshwar is likely to be estimated to have a surface elevation either much higher (colder) or lower (warmer) than at the specific (point) location. These differences come from both latitudinal position and simulated elevation of the source grid cells in each of the reanalyses. By taking into account the likely role of elevation differences between the actual Mukteshwar IMD station (2218m asl) and the invariant orography values from each of the reanalyses we can infer the component 'residual' bias. This bias could be due to oversimplification of spatial temperature gradients through coarse spatial resolution and hence oversimplification of land surface cover and its modulation of surface energy balance influences on near surface air temperature. Alternately the biases of individual reanalyses' representation of near surface temperature could be due to errors in surface energy balance or cloud radiative effects. In the case of the Mukteshwar IMD station, all simulated elevations are lower than the 'real world' and differences range from less than 100m lower in JRA55 to nearly 1500m lower in ERA-Interim. The cold bias (Figure 2, Table 3) in JRA55 mean temperature thus cannot be attributed solely to elevation. For the remaining reanalyses, assuming a temperature lapse rate 0.7°C per 100m vertical difference, their respective differences between real and simulated elevation could account for the following amounts of their warm biases: a) ERA-Interim = ~10.5°C; b) NASA MERRA2 = ~8°C; and c) ERA5 = ~5°C. Subtracting these estimates from the calculated mean temperature biases in Table 3 implies that 'elevation corrected' (cold) biases would be roughly 4°C in both ERA-Interim and JRA55 and perhaps less than 2°C in both NASA MERRA2 and ERA5.

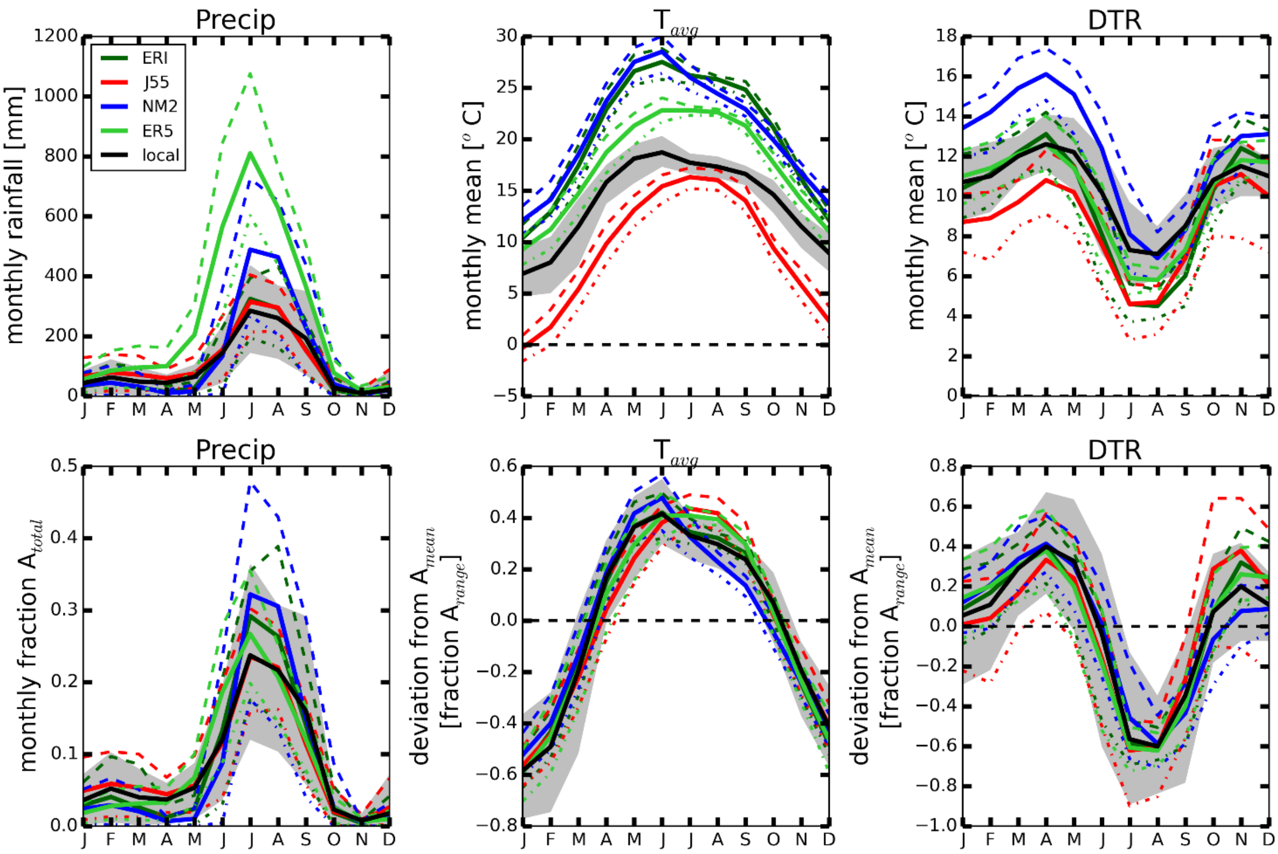

Figure 2: Climatologies of primary (precipitation) and secondary (temperature: Tavg, DTR) variables for the Mukteshwar site from local observations and global meteorological reanalyses. Solid lines indicate period mean values. Areas bounded by grey shading and dashed lines denote ranges of 10^{th} to 90^{th} percentiles respectively for local observations and reanalyses. Top row shows

absolute values. Bottom row shows normalised values (calculated as described in Section 2.2.2) thus comparing de-biased skill at representation of annual cycle timing and amplitude. notes: ERI=ERA-Interim, J55=JRA-55, NM2=NASA MERRA2, ER5=ERA5, local = local observations at Mukteshwar IMD.

The differences between individual reanalysis performance in absolute and normalised terms can be considered in detail by calculating error metrics – the mean bias/error for absolute values and the root mean square deviation (RMSD) for normalised values – of the annual cycle monthly period statistics with the local observations as the reference or 'ground truth' (Table 3). This is not limited to the period mean but can also address interannual variability through quantiles of the distribution. Table 3 shows this for the 10th and 90th percentiles of the distributions of individual calendar months for the 39-year record period. This indicates that the smallest bias for different statistics of a given variable may be from different reanalyses. Furthermore, the smallest biases in absolute terms may differ from those in normalised terms. Despite this, errors in the mean are for the most part smaller than errors in the 'tails' of the distribution, particularly in normalised terms. This is an indication of how gridded datasets struggle to accurately represent interannual variability at the point scale.

Table 3: Identified biases – as mean bias (error) for absolute values and root mean square deviation (RMSD) for normalised values – of annual cycle of monthly period statistics, for individual reanalyses' grid cells overlying Mukteshwar IMD station, 1980 to 2018.

Identified biases		Precipitation [absolute units: mm]				Mean temperature [absolute units: °C]				Diurnal temperature range [absolute units: °C]			
type	statistic	ERI	J55	NM2	ER5	ERI	J55	NM2	ER5	ERI	J55	NM2	ER5
Absolute (mean bias)	10%	5.7	21.1	11.8	110.4	7.3	-4.2	7.5	4.2	-0.6	-2.1	2.0	0.1
	Mean	-7.4	11.5	26.5	152.7	6.5	-4.7	6.9	3.5	-0.8	-1.8	1.9	-0.4
	90%	-22.8	2.4	52.6	186.9	7.1	-5.0	6.5	3.0	-1.0	-1.9b	1.8	-0.8
Normalised (RMSD)	10%	0.021	0.022	0.022	0.031	0.116	0.113	0.128	0.100	0.151	0.099	0.182	0.162
	Mean	0.025	0.015	0.041	0.026	0.033	0.075	0.067	0.054	0.079	0.111	0.078	0.082
	90%	0.036	0.041	0.058	0.053	0.086	0.108	0.086	0.081	0.189	0.205	0.150	0.180

Key to reanalyses labels: ERI = ERA-Interim; J55 = JRA-55; NM2 = NASA MERRA2, ER5 = ERA5.

[3.1.2] *Climatologies of tertiary variables (radiation, humidity and wind speed) from meteorological reanalyses*

Despite the potential shortcomings in the available data and the lack of in-situ observations to provide a 'ground-truthing' benchmark, it is nevertheless interesting to compare the climatologies of net surface radiation (R_{sfcnet}), relative humidity (RH) and windspeed at 10m height (10mWind) from the four global reanalyses, ERA-Interim, JRA-55, NASA MERRA2 and ERA5 (Figure 3). For R_{sfcnet} there is general agreement between the reanalyses, particularly after normalisation: a strong annual cycle in net radiation driven by seasonal variation in incoming shortwave (solar) energy. For RH there is a similar level of agreement, after normalisation, with a pronounced annual minima in the pre-monsoon months (April, May) and a strong maximum during the monsoon (July to Sept). Interestingly, although absolute value estimates differ by a factor of 2, there is also post-normalisation agreement on the shape of the annual cycle in 10mWind.

In the absence of local observations to evaluate biases in the reanalyses' estimates of these variables, the implications for reference evapotranspiration of the mean states of these three variables bears elaboration. R_{sfcnet} contribution to driving evapotranspiration will be greatest prior to the monsoon but only marginally reduced during the rainy season. The evapotranspiration-enhancing vapour pressure deficit (increasing as RH decreases), however, will be substantially greater in the pre-monsoon than during the rains. 10mWind will act in concert with RH as higher windspeeds during the pre-monsoon will further enhance energy and moisture transfer from the surface toward the atmosphere. Lighter winds during the monsoon will further limit what would otherwise, due to strong radiative input, be elevated evapotranspiration rates. Again, given the absence of direct "ground-truthing" observations for the tertiary variables it is worthwhile to point out the strong (logical) similarities – comparing Figures 2 and 3 – in the shapes of the annual cycles of R_{sfcnet} and T_{avg}. Similarly, the shapes of the normalised annual cycles of 10mWind and DTR have much in common. The normalised annual cycle of RH, if inverted, also resembles this latter pattern. These similarities clearly point to the logical use of directly observed 'secondary variables' (T_{avg}, DTR) as potential proxies for the estimates of tertiary variables (R_{sfcnet}, RH, 10mWind) provided by the large-scale reanalyses.

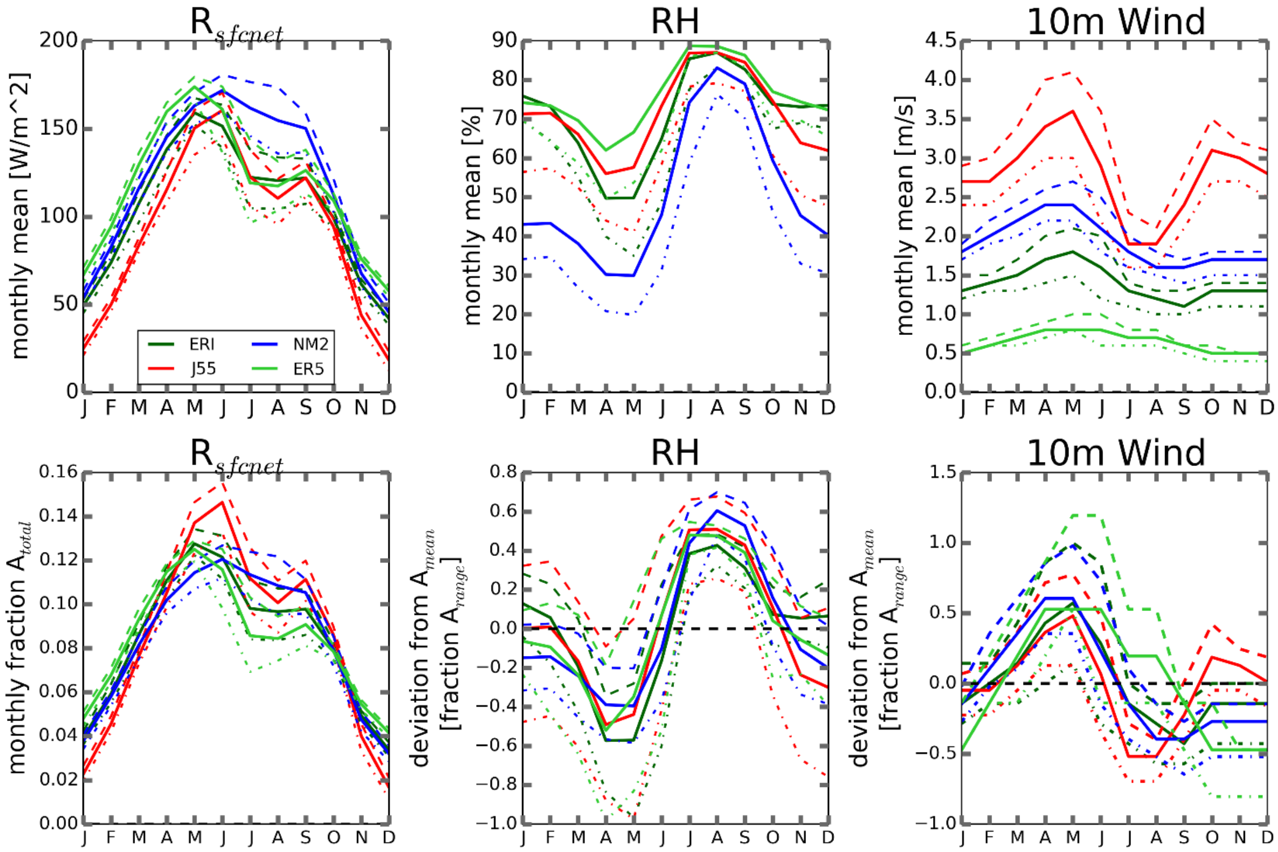

Figure 3: Climatologies of tertiary variables – radiation as Rsfcnet, humidity as RH, wind as 10m windspeed – for the Mukteshwar site from global meteorological reanalyses. Solid lines indicate period mean values. Areas bounded by dashed lines denote ranges of 10th to 90th percentiles. notes: ERI = ERA-Interim, NM2=NASA MERRA2, J55=JRA-55, ER5 = ERA5;

[3.2] CWB climatology

[3.2.1] *CWB estimates derived from local observations*

The annual cycles of precipitation, ET_0 and CWB at Mukteshwar are shown in Figure 4. The shape/form of the reference evapotranspiration (ET_0) annual cycle strongly resembles that of T_{avg} and

R_{sfcnet}, as all three are predominantly influenced by the seasonal variations of incoming solar radiation. The annual cycle of the CWB is (logically) dominated by the moisture surplus during monsoonal months, with surplus/deficit magnitudes >100mm only calculated/estimated for June through September. In other months values are much closer to equilibrium (0mm) as both rainfall and ET_o are smaller in magnitude. CWB is generally, but not uniformly, positive in January/February and similarly negative in April, May and October. Local agricultural practices (near Mukteshwar) generally have two cropping seasons per year with planting timings (~Nov-Jan and ~June-July) coinciding with/immediately preceding moisture surplus periods and harvest timings (~May-June and ~Oct-Nov) coinciding with peak moisture deficit. The range of interannual variability in CWB -- illustrated in Figure 4 by the 10^{th} and 90^{th} percentiles – indicates that some years moisture deficits during the 'maturity' phase will be more severe than others. The impacts of CWB variability on small-scale agriculture in the Mukteshwar area are subjects of on-going research.

[3.2.2] *CWB estimates derived from meteorological reanalyses*

Comparisons of ET_0 estimates from individual reanalyses to ET_0 estimates from local observations of secondary (T_{avg}, DTR) variables and (reanalyses) ensemble mean estimates of tertiary variables (radiation, humidity, windspeed)) show firstly that reanalysis ensemble members either closely match (JRA55, ERA5) or substantially overestimate (ERA-Interim, MERRA2) ET_0 in absolute terms. The overestimation cases appear to be correspond to the absolute bias in T_{avg}. Secondly, the normalisation procedure used for the primary and secondary variables (precipitation, T_{avg} and DTR) shows that despite absolute biases there is strong agreement amongst all data sources regarding the shape/form of the ET_0 annual cycle. Interestingly because the individual reanalyses tend to overestimate (in absolute terms) both precipitation and ET_0, resultant absolute CWB biases are smaller in magnitude. Logically, the normalisation procedure again shows very strong agreement on the shape/form of the CWB annual cycle. Comparing Figures 3 and 4 reveals a notable similarity between the normalised forms of the annual cycles of RH and CWB.

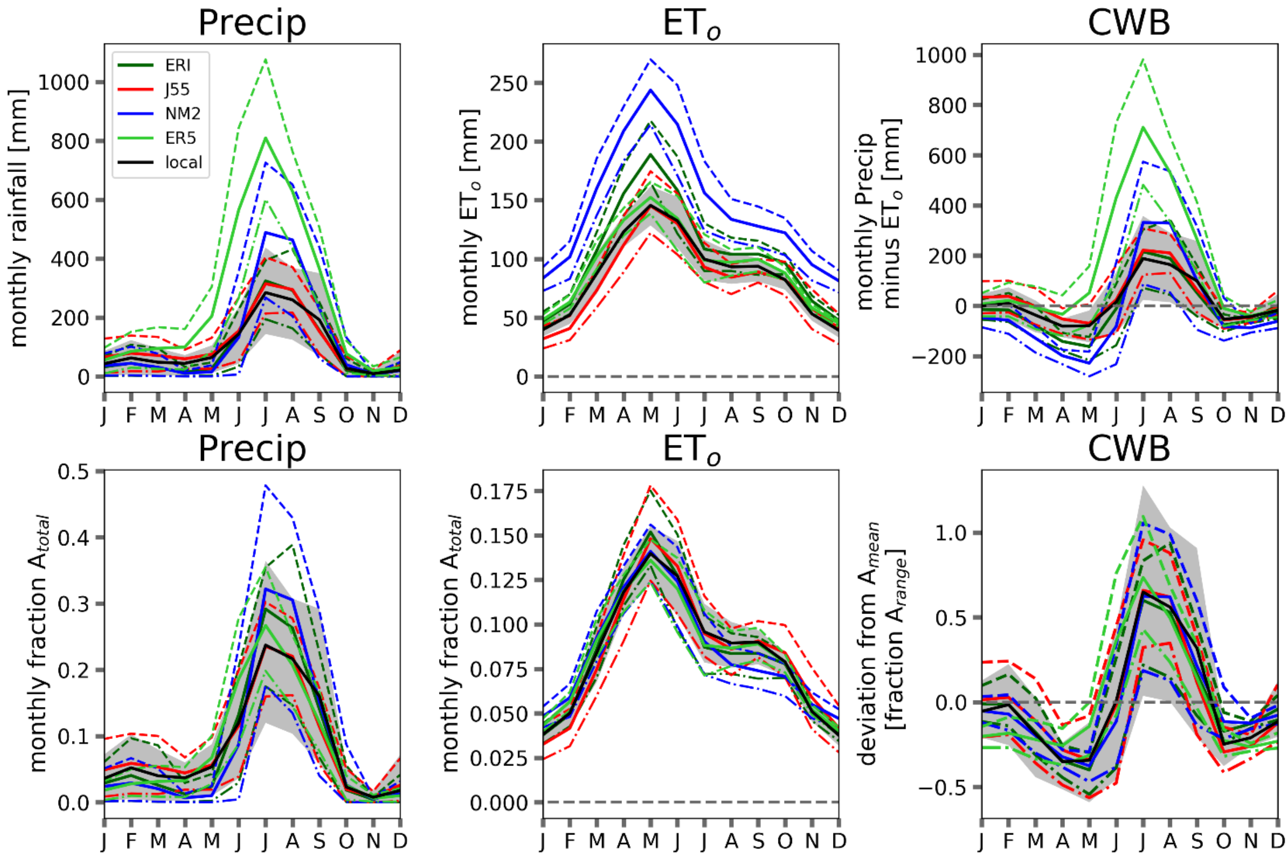

Figure 4: Climatologies of contributing components, i.e. precipitation and reference evapotranspiration (ET_0), along with the climatic water balance (CWB) for the Mukteshwar site from local observations and global meteorological reanalyses. Solid lines indicate period mean values. Areas bounded by grey shading and dashed lines denote interannual variability quantified as ranges of 10th to 90th percentiles respectively for local observations and reanalyses. notes: ERI = ERA-Interim, J55=JRA-55, NM2=NASA MERRA2, ER5 = ERA5, local = local observations at Mukteshwar IMD

[3.3] Time-series in individual variables

Agricultural practice near Mukteshwar predominantly uses two growing seasons per year. To avoid analysing individual growing seasons spanning more than one calendar year we simplified their representation into two five-month time aggregates: January to May (cold) and June to October (monsoonal). These season definitions were then used to calculate yearly time-series of standardised anomalies for the primary and secondary climate variables (Figure 5) and for ET_0 and CWB (Figure 6) from both local observations and large-scale reanalyses. Figures 5 and 6 show that for all variables in both seasons, agreement is reasonably strong both by reanalyses with local observations and between individual reanalyses. Nevertheless, consensus on sign and magnitude of anomalies is visibly closer for the cold season (JFMAM) than during monsoonal months (JJASO). The sequencing of CWB anomalies (Figure 6) in both seasons strongly resembles the corresponding sequencing of precipitation anomalies (Figure 5) thus underlining how precipitation dominates the CWB at Mukteshwar. Meanwhile, the sequencing of ET_0 anomalies (Figure 6) visually resemble Tavg anomalies (Figure 5) in respective seasons, thus providing further evidence for the strong role of incoming shortwave (solar) radiation in driving atmospheric moisture demand.

In terms of emerging patterns of change, none of the variable-season combinations (individual panels in Figures 5 & 6) appear to follow a linear trend. Nevertheless there are substantially fewer negative anomalies in the latter half of the time period for T_{avg} in both seasons, which might indicate local warming. The opposite is true for DTR, with fewer positive anomalies later in the time period during both seasons. This indicates a narrowing of differences between daily maximum and minimum temperatures, possibly due to increasing cloud-cover and/or near surface water vapour. In contrast, precipitation anomalies are highly variable in both seasons. ET_0 anomalies in the cold season appear to increase in line with T_{avg} warming. Evidence of ET_0 change during the monsoonal season is less clear, with negative anomalies at both the beginning and end of the period and maximum values during the 1990s and early 2000s. The distributions of CWB anomalies throughout the time period in both seasons show similar levels of 'noise' (apparent randomness) to those in Precip, albeit with weak indications of a decreasing pattern in the cold season (JFMAM) contrasting with equally weak indications of increases during the monsoon (JJASO).

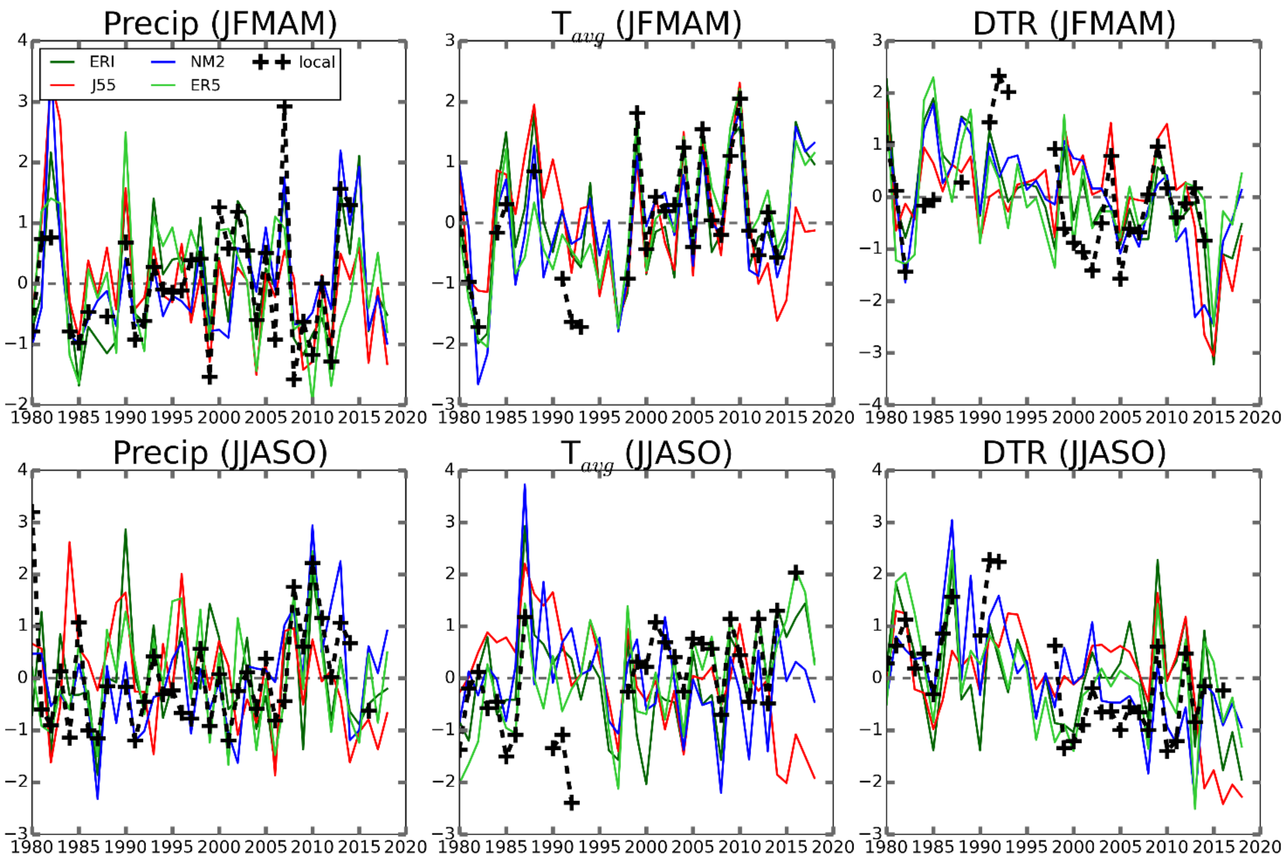

Figure 5: Standardised anomaly (units of 'standard deviation') times series of seasonal aggregates of primary (Precip) and secondary (T_{avg}, DTR) variables. Cold season (JFMAM) is January through May. Warm season (JJASO) is June through October. ERI = ERA-Interim, J55=JRA-55, NM2=NASA MERRA2, ER5 = ERA5, local = local observations at Mukteshwar IMD.

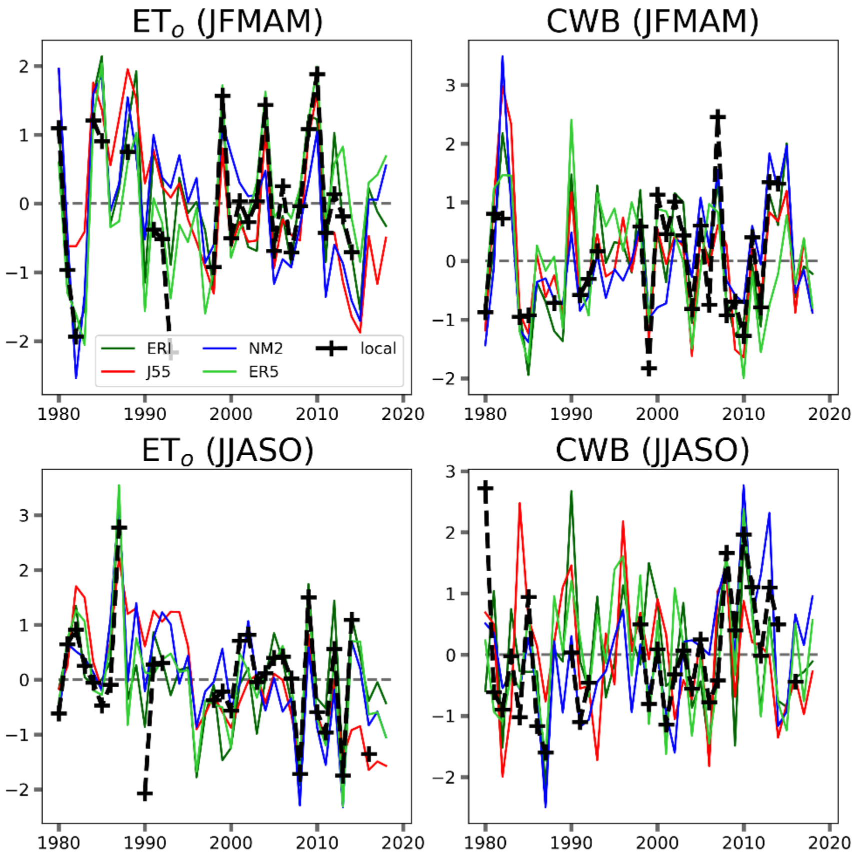

Figure 6: Standardised anomaly (units of 'standard deviation') time-series of seasonal aggregates of extrapolated reference evapotranspiration (ET_0) and climatic water balance (CWB). Season definitions and figure symbology as in Figure 5.

[3.4] Correlations of hydroclimate variables to time (trend precursors)

The underlying variability ("noise") exhibited by the time-series of the hydroclimate variables presented (Figs 5 & 6) shows that attempting to fit linear trend rates to observed historical anomaly patterns would not appear entirely appropriate. Nevertheless, while investigating on-going water cycle change, insight can be gained through assessing the correlation, e.g. Kendal 'tau', of individual variables with time, i.e. series of yearly values for individual calendar months. Results of this procedure for the Mukteshwar site data are shown in Figure 7. Strong positive (negative) correlations to the time index are of course indicative of increasing (decreasing) tendencies in the variable values.

Precipitation is globally recognised as highly variable, and in the Mukteshwar site time-series analyses, noise – correlation values found through random shuffling of observations as described in the Methods section – largely exceeds signal. In contrast, mean temperature (T_{avg}) shows consistently positive correlation throughout the annual cycle, with several months above the 95th percentile – as well as four months above the 99th and even two months exceeding the 99.9th percentile – of results expected from simple random sequencing. Estimated DTR correlation with time, however, shows mixed results across the annual cycle in terms of both strength and sign. In the first 5 months of the year DTR correlation to time is within the random variability or 'noise' range. From June through November there are notable decreasing tendencies (negative correlations), with the monsoonal months in particular exceeding values expected from random sequencing. The near identical patterns of correlation of CWB and precipitation to time further illustrate how Mukteshwar CWB is dominated by moisture inputs rather than potential evaporative demand. Reference evapotranspiration (ET_0) for its part shows a mixed pattern, with the late winter and spring seemingly dominated by mean temperature (thus increasing), but with tendencies in monsoonal months driven by DTR (thus decreasing). If these temporal tendencies continue, the increases during the middle of the 'cold season' cropping cycle could lead to more damaging moisture stresses in dry years. A general remark, applying to all variables shown in Figure 7, is that correlations between variable estimates from three of the reanalyses – ERA-Interim, JRA55 and NASA MERRA2 – and time generally track those for local observations relatively well in cooler months (~Nov to Feb) but often diverge widely in warmer months (March to October). This may well result from generally strong skill of these reanalyses to represent conditions of climates dominated by large-scale/frontal precipitation and weakness at representing moisture and radiation fluxes in convection-dominated conditions. Time-variable correlations from ERA5, however, track noticeably closer to the time-variable correlations in the local observations, with the exception of ET_0. This is despite ERA5 having broadly similar skill to the other reanalyses – albeit with a very strong wet bias in precipitation – in climatological representation of the key variables. ERA5 is the newest of the reanalyses and it will be of scientific interest to explore if this pattern of performance is repeated in through other locations in South Asia and beyond.

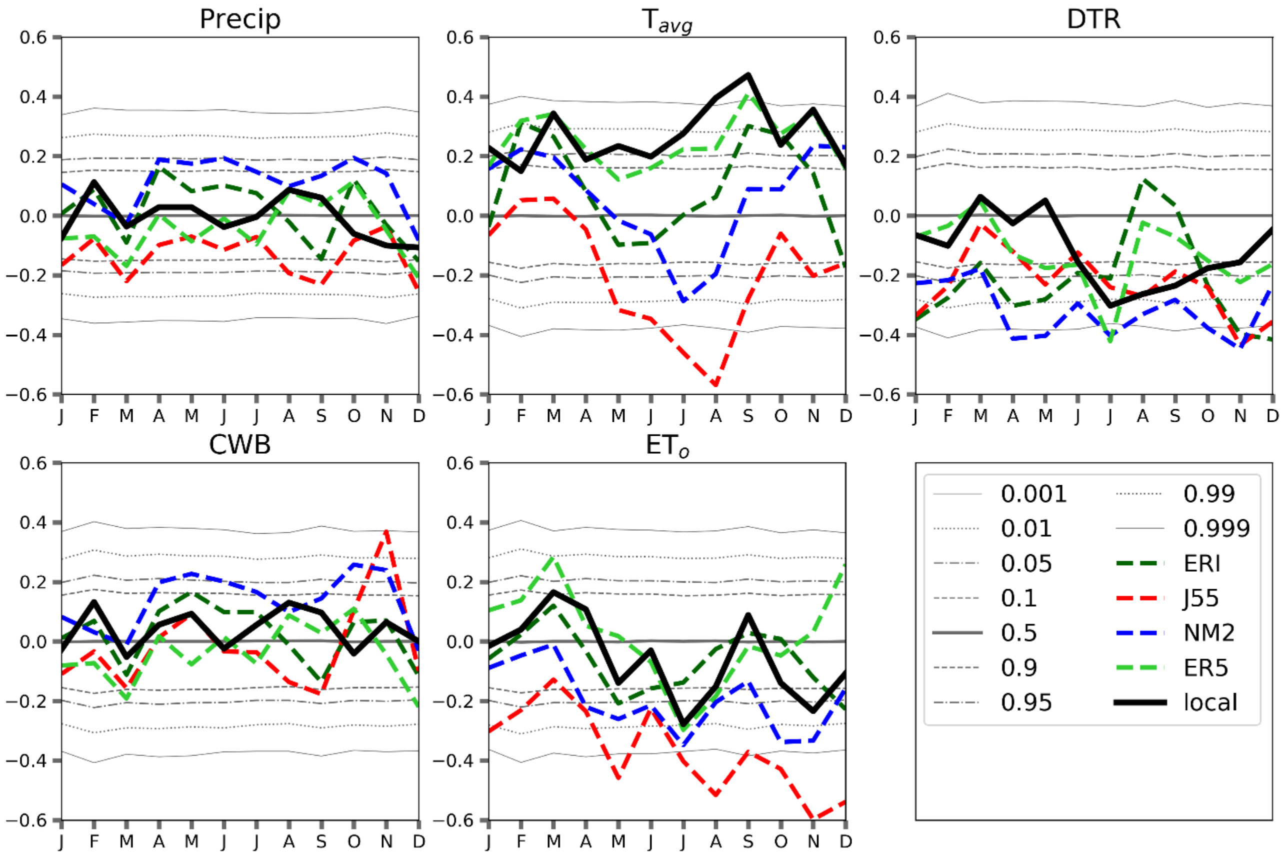

Figure 7: Kendall Tau correlation of hydroclimate variables to time for individual calendar months (totals/means). Grey lines indicate statistical distribution of correlation values resulting through randomisation of observation order/sequencing; ERI=ERA-Interim, NM2=NASA MERRA2, J55=JRA-55, ER5=ERA5, local = local observations at Mukteshwar IMD.

In light of the clearly dominant impact of precipitation on CWB, it is worthwhile to further explore how precipitation might be changing at Mukteshwar, specifically in terms of the frequency of daily rainfall accumulations exceeding specific totals. Before potential changes – assessed as correlations to time – in event intensity can be considered, the (annual cycle) climatology of rainfall accumulations must first be examined (Figure 8). The defining influence of the monsoon on frequency of rainfall events is unmistakeable regardless of whether 1mm, 10mm or 50mm daily accumulation thresholds are utilised. The monthly frequency of events >1mm and >10mm (daily) are both strongly proportional to monthly rainfall totals. With the exception of very rare winter storms (in particular in February), events with daily totals >50mm occur during the monsoonal period from June through September. In comparison with the local observations, all four meteorological reanalyses exhibit characteristic "drizzle biases" (Hong *et al*, 2006; Piani *et al*, 2010) during at least part of the annual cycle in that low intensity events are estimated to occur with excessive frequency. For the high intensity events, exemplified in Figure 8 by daily accumulation >50mm, there are clear differences between the individual reanalyses. ERA-Interim and JRA-55 largely underestimate the absolute frequency of these events. Both NASA MERRA2 and ERA5 in contrast strongly overestimate (absolute) frequencies in June through August but match observed frequencies in September. As with the climatologies of key meteorological variables and CWB components, the normalisation of (observed and) estimated frequency of rainfall events exceeding specified accumulation thresholds shows substantially greater agreement/consensus than the absolute values. This shows that meaningful information content on precipitation event characteristics, including extremes, can be derived from the reanalyses despite biases in absolute values.

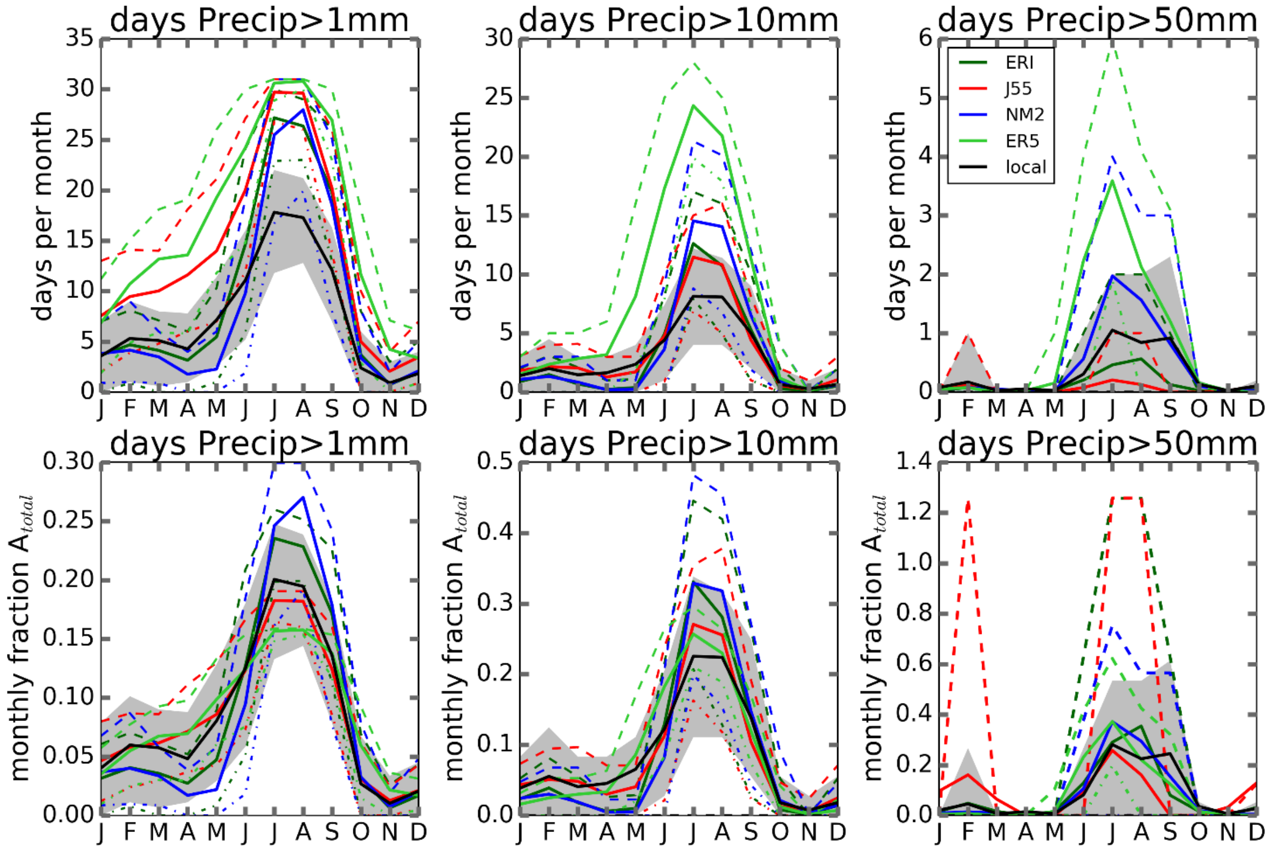

Figure 8: Climatology of frequency of daily precipitation surpassing thresholds. Note: These results are not 'binned', hence lower thresholds include all larger events, i.e. Precip>50mm is a subset of Precip>10mm which is itself a subset of Precip>1mm. Solid lines indicate period mean values. Areas bounded by grey shading and dashed lines denote ranges of 10^{th} to 90^{th} percentiles respectively for local observations and reanalyses. Data source are abbreviations as follows: ERI=ERA-Interim, J55=JRA-55, NM2=NASA MERRA2, ER5 = ERA5, local = local observations at Mukteshwar IMD.

In the context of a globally warming climate there is both scientific expectation and substantial observational evidence for increases in the accumulation of precipitation from individual storm events

– from either increases in intensity, duration or both (Trenberth et al, 2003). At Mukteshwar, however, over the common period (1980 to 2018) covered by local observations and the four meteorological reanalyses, there is an absence of consistency in sign and strength of correlation of precipitation indicators to time and relatively little in way of consistency/consensus between the independent data sources (Figure 9). With specific regard to the local observations, the sequencing of measured monthly precipitation amounts and event (greater than threshold) frequency rarely show correlation strengths greater than that found through <10% of randomisation sequence cases. Even so, one noteworthy aspect is that correlation of precipitation amounts to time appears strongly influenced by correlation of medium to large accumulation events (a mixture of >10mm and >50mm). None of the meteorological reanalyses consistently match the sign and strength of correlations of local observations to time, although ERA5 is marginally closer than the others. There is some indication, however, that agreement is better in colder months (October to April) than in warmer months (May to November). In terms of changes which could be deemed significant, the clearest signals (from local observations) appear to be increases in frequency of >50mm (daily) events in February and August. These specific increases in frequency of large events are counterbalanced by decreases in large event frequency in March and November. It remains to be established whether these apparent changes (shifts in seasonality?) are underpinned by evolving physical mechanisms or are simply indicative of the vast range of inherent variability (‘noise’) in the local precipitation regime.

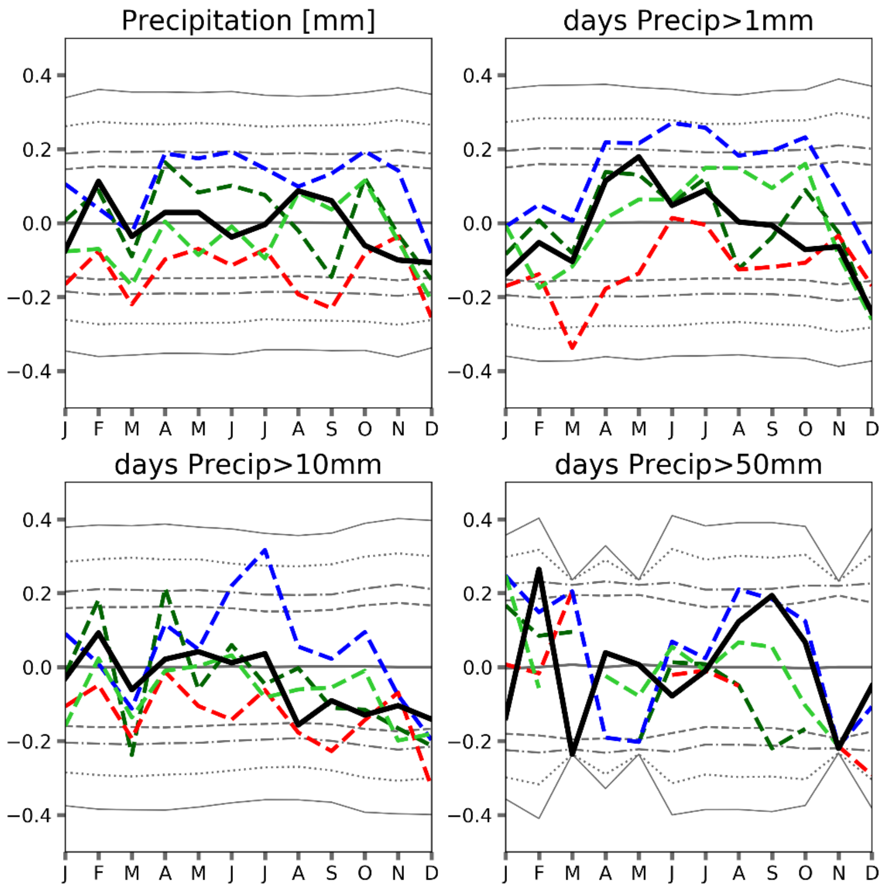

Figure 9: Kendall Tau correlation of frequency of daily precipitation surpassing thresholds to time for individual calendar months. Grey lines indicate statistical distribution of correlation values resulting through randomisation of observation order/sequencing, as per Figure 7; ERI=ERA-Interim, J55=JRA-55, NM2=NASA MERRA2, ER5 = ERA5, local = local observations at Mukteshwar IMD.

[3.5] Atmospheric processes as candidate determinants of CWB change

Simple evaluation of change over recent decades provides little insight into likely future evolution of the climate system unless those changes can be linked to driving physical mechanisms whose behaviour can be anticipated with strong confidence. As an illustrative example the potential influence of (local) cloud cover is examined here to provide context for the historical tendencies reported in the preceding section. Future evolution of cloud cover may be quite complex due to dependency of formation on presence of 'seed particles' (e.g. aerosols) but can nevertheless be interpreted through fundamental aspects of climate science relating changes in evaporation and condensation of water vapour to temperature change. Atmospheric circulation may play a role in evolution of cloud cover through variations in the paths or "tracks" of large-scale storm systems, including those linked to the monsoon.

The potential of (local) cloud cover influence in driving interannual near surface climate variability is examined here as an illustrative example of a causal mechanism. Correlations shown in Figure 10 are essentially monotonic (uniformly signed) and exhibit strength levels which are highly unlikely to exist by chance. There is relatively strong consensus between the correlations found using local observations of near surface climate and those found using reanalyses estimates. These factors underpin relatively straightforward physical interpretations. Precipitation shows strong positive correlation to cloud cover which is logical since rain rarely falls under clear skies. For Mukteshwar there are consistent negative correlations between cloud cover and mean temperature (T_{avg}) although these are weaker in cold months (October to February) and during the late monsoon (August and September) when the cooling influence of clouds through shortwave (solar) radiative forcing is tempered by a warming influence of longwave (thermal) forcing. These findings are in line with a previous study (Forsythe *et al.*, 2015) of cloud influence on temperature elsewhere in the Himalayan arc. DTR also shows consistent negative correlations, although values are perhaps less strong and less consistent in magnitude than could be expected given the presumed relationship between clear (cloudy) skies and amplified (suppressed) DTR. This may either point to limitations in cloud representation by meteorological reanalyses and/or substantial roles for other radiative influences, e.g. water vapour, in modulating DTR.

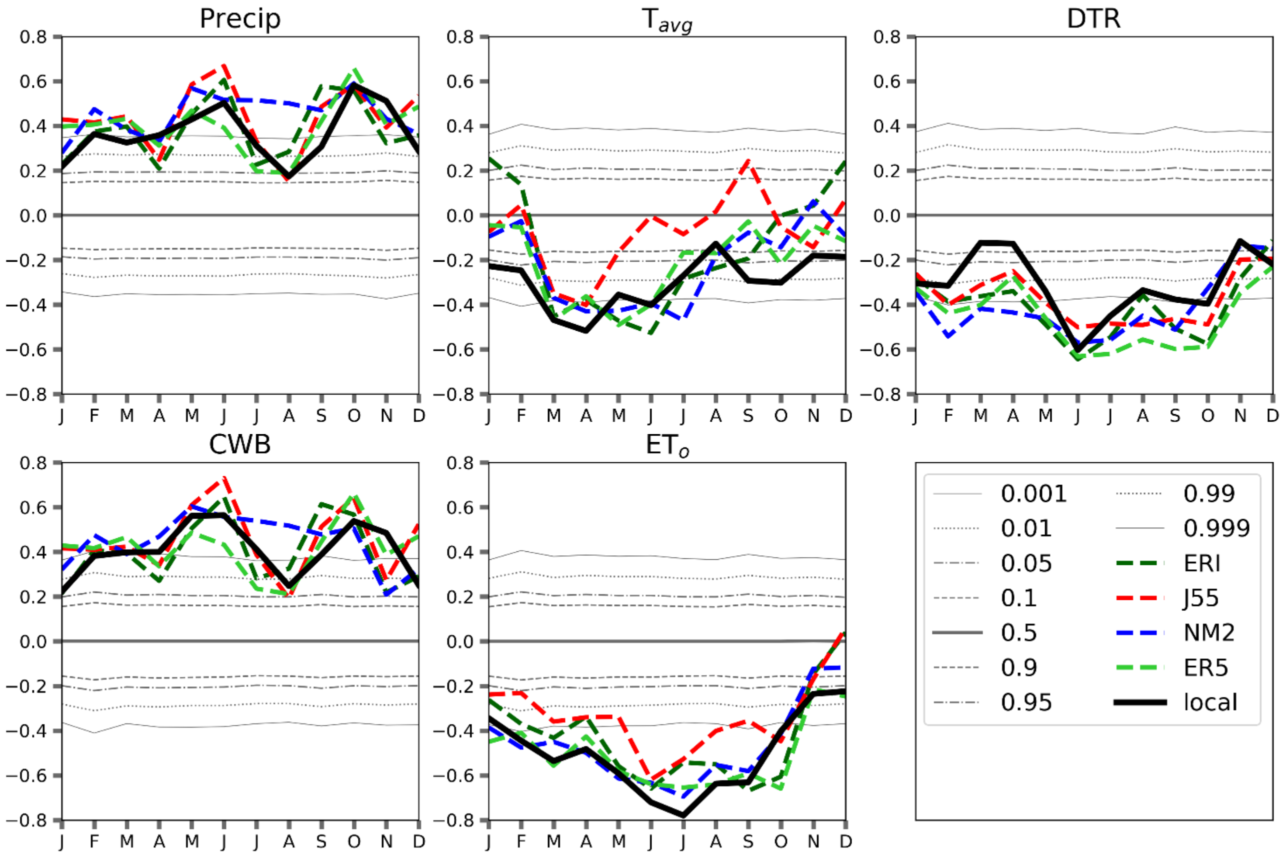

Figure 10 Kendall Tau correlation of near surface climate variables to total cloud fraction for individual calendar months. The individual reanalysis correlations are calculated between variable estimates by that data source. Local observations correlations are calculated against the ensemble mean of cloud cover estimates from the four reanalyses. Grey lines indicate statistical distribution of correlation values resulting through randomisation of observation order/sequencing; ERI=ERA-Interim, J55=JRA-55, NM2=NASA MERRA2, ER5 = ERA5, local = local observations at Mukteshwar IMD.

[4] <u>Discussion and future perspectives</u>

[4.1] Descriptive hydroclimatology

From an objective standpoint, CWB is an imperfect and admittedly oversimplified aggregate metric of water availability. This shortcoming is due to its neglect of the role of soil as reservoir storing moisture between precipitation events. Soil characteristics, along with precipitation intensity/event magnitude, play a critical role in modulating the partitioning of rainfall between direct/surface runoff and (subsurface) infiltration. Nevertheless, CWB provides an accessible/feasible, meaningful indicator of moisture availability without necessitating soil characteristics data (whose acquisition would be cost prohibitive). It would be possible to substitute in-situ subsurface information with sensitivity studies and probabilistic estimation of potential soil moisture balance, but such an approach would inherently entail such broad uncertainty bounds that resulting information content would be questionable.

This study has drawn substantially on climate variable estimates from global meteorological reanalyses. These data sources do have important limitations, particularly in areas where the influence of steep topographic gradients greatly exceeds the level of detail enabled through their relatively coarse spatial resolution. This coarseness along with methodological limitations of the data assimilation and forecasting systems which drive the reanalyses can (often) lead to strong biases in absolute value estimates of key climate variables, particularly precipitation. With relevance to this study specifically, the Mukteshwar area is situated in a (heterogeneous) transition zone at the margin between lowlands/plains and high mountains. The reanalyses do nevertheless have strong advantages, principally that they provide spatially and temporally continuous (internally consistent) estimates of a wide range of climate variables. Much of the aforementioned general biases can be overcome through simple normalisation/standardisation procedures as shown in Figure 2. It must be recognized, however, that these normalisation/standardisation procedures may not be effective at the transitions between climate regimes where different physical processes and seasonalities (timing of annual maxima and minima) intersect.

From a more general scientific standpoint, exploratory data analysis can provide a pathway to improved understanding underlying physical mechanisms driving variability and change in natural systems. In order to attain this goal, temporal aggregation, whether monthly or seasonal, should reflect prevailing climate patterns such as precipitation regimes. The robustness of preliminary results can be assessed based on their independence (i.e. lack of sensitivity) to the choice of 'analytical time-window', i.e. the start & end years for correlation and trend calculations. In the case of Mukteshwar specifically, using the CWB framework, apparent changes in climate over recent decades can be separated based on whether the variables in question influence atmospheric moisture supply or demand. In terms of supply, the dominant aspect of precipitation is arguably underlying/inherent (chaotic) variability although there is tentative evidence for the intensification of the hydrological cycle based on increasing frequency of large (accumulation) rainfall events in key months. This intensification of precipitation events is coherent with theoretical expectations, particularly the Clausius-Clapeyron relationship (Guerreiro *et al*, 2018), of climate evolution driven by anthropogenic global warming. Further research could also investigate in greater detail whether shifts in regional atmospheric circulation are changing the frequency with which storm systems pass through/over the Mukteshwar area. Regarding atmospheric moisture demand, evidence from local observations seems to robustly demonstrate year-round increases in daily mean temperature (T_{avg}) and corresponding decreases (except during spring) in diurnal temperature range (DTR). Additional investigation would be required to determine if the proximate mechanisms driving these changes, and in particular the strong T_{avg} increases in March, are predominantly attributable to cloud radiative influences, changes in regional atmospheric circulation or other underlying factors. On this point, i.e. with respect to water vapour (humidity), in light of visible similarity between the (normalised) annual cycles of RH and CWB, it may be worthwhile to consider the potential role of RH in influencing CWB components and the key climate variables. When using data from meteorological reanalyses, however, it is unlikely there would be substantial additional 'information content' in exploring correlations of (other) near surface climate variables to RH because RH and cloud cover will be highly correlated in these datasets.

These results have potential implications for regional applications of (physically-based) "emergent constraint" approaches for validation/evaluation of climate models (Knutti et al, 2017; Cox et al 2018; Eyring et al, 2019) since accurate representation of moisture fluxes – whether as RH or CWB – near the land surface are central to the plausibility and relevance of simulated future conditions which will modulate the impacts of anthropogenic climate change.

[4.2] Promising avenues and critical pathways:

While the findings of this study are of greatest interest for the Mukteshwar area, adjacent sections of the Kumaun Lesser Himalaya (KLH) and similar areas of the Ganges basin headwaters, the methodology employed has much broader potential relevance/transferability.

[4.2.1] *Validation of simulated historical climatologies and downscaling of projected future conditions*

In addition to driving biases in mean temperature, the precise location of grid cell boundaries can also influence the characterised precipitation regime. In this specific case, both ERA-Interim and NASA MERRA2 appear to somewhat overemphasise the monsoonal character of Mukteshwar precipitation with too large a fraction of annual rainfall found from July to August and too small from January to May. ERA5 is distinct in that its absolute wet bias is severe but its representation of the (normalised) annual distribution of precipitation is relatively skilful albeit with both onset and recession of the monsoon occurring earlier than in local observations. Despite its coarse spatial resolution, JRA55 estimates (relatively) accurately both the magnitude and timing of precipitation. These issues of magnitude and timing (seasonality) may further influence subsequent elements of study/analyses, as it implies differing relative contributions of distinct rainfall generating mechanisms (frontal/stratiform versus convective). Precipitation frequencies and amounts resulting from these mechanisms may follow divergent trajectories as a result of anthropogenic climate change. While large-scale meteorological reanalyses generally represent the shape of the annual cycle well, they struggle nevertheless to adequately capture the magnitude of interannual variability, even in relative terms. This may be linked to aggregation/homogenisation of conditions across large “grid cells” thus smoothing substantial local (“sub-grid”) variability. These limitations, particularly evidenced in the biases shown in the relatively finer resolution ERA5, support the need for high resolution dynamical downscaling of global meteorological reanalyses. Previous studies in North America have found that spatial resolutions finer than 10km are necessary to capture the influence of topography on precipitation gradients (Rasmussen et al, 2011). Separately, in regions with predominantly warm rainfall regimes, precipitation should be simulated using models run at convection-permitting spatial resolutions, i.e. less than 4km (Kendon et al, 2012; Prein et al, 2015).

Looking beyond the evaluation of global meteorological reanalyses as potential sources of historical data in observation void/gap areas, the approach utilised here could equally be applied as a framework for site-based validation of climate model outputs (CORDEX, CMIP, etc). Validation and bias assessment efforts to quantify climate model performance often focus on spatial patterns within the modelled domain or on annual cycles of large spatial aggregates, e.g. along longitudinal bands or over major river basins. Such broad aggregation can easily obscure whether the simulated climatologies are realistic at the scale of natural resource management. By relating – both in absolute and normalised/standardised terms -- climate model outputs to the CWB (derived from local observations) a meaningful assessment of hydro-climatological ‘fidelity’ or skill can be made. Repeating CWB ‘point’ assessments for multiple locations with quality multi-decadal observational records can provide much greater insight into model performance than simple gridded or spatially aggregated assessments would yield. These site-based bias assessments can also provide the foundation for downscaling – if a ‘delta change’/perturbation type approach is adopted -- of future climate projections. This is because it is necessary to relate the incremental (multiplicative for precipitation, additive for temperature) changes between projected future and simulated historical climate conditions to the local observational record in order to minimise ‘contamination’ of impact assessments with model biases. This is, however, an imperfect approach because the underlying climate model errors in representing physical processes will still be present in the projected ‘change factors' (Ehret *et al*, 2012) albeit reduced through exclusion of the most unrealistic models. This fact provides further impetus in the drive toward high-resolution dynamical downscaling capable of accurately simulating physical processes including orographic and convective precipitation.

[4.2.2] *Attaining field-scale representation of CWB and beyond*

Along similar lines, the full suite of meteorological variables utilised to calculate (FAO Penman Monteith) reference evapotranspiration are rarely observed at individual locations particularly in countries with emerging or developing economies (i.e. the 'Global South'). The three key variables -- precipitation, T_{avg} and DTR – can, however, be observed accurately and at low cost around the globe. As such the number of meteorological stations with multi-decadal observational records of these variables is substantial. Even where longstanding measurements have not been conducted, observational systems can quickly be established and, within a few years of operation, results can be compared to national monitoring systems and/or gridded data sources. Supplemental low-cost in-situ measurements of additional variables, such as relative humidity (RH), can further reduce uncertainty in deriving reference evapotranspiration and CWB from these primary climate observations. The role of RH is of high potential interest as it is possible to directly observe RH (in additional to Precip and T_{avg}/DTR) locally using low cost sensors. Expanding the availability of local RH observations could thus provide a promising avenue for highly-scalable additional 'ground truthing' of gridded/global datasets – both meteorological reanalyses and climate models – as well to reduce uncertainty in CWB estimates calculated using estimates of 'tertiary' variables extracted from these datasets. Furthermore, at local spatial scales, meaningful investigations of soil characteristics (depth, texture) become feasible. Such field campaigns thus enable progress from the relatively simple CWB estimates to extrapolation of full water balance, i.e. including actual evapotranspiration (AET), effective precipitation (precip minus AET), direct runoff and percolation/baseflow.

[5] Conclusions

[5.1] Specific findings regarding the climatic water balance in proximity to Mukteshwar

In order to characterise the evolving hydroclimate of a case study within the middle mountains in the transition zone between the Indo-Gangetic plain and the Greater Himalaya, we have utilised meteorological observations from the Mukteshwar station (Nainital district, Uttarakhand state) of the India Meteorological Department (IMD) to quantify the local climatic water balance (CWB) – along with the variables which determine it – in terms of both annual cycles and interannual variability. The observed patterns of year-to-year variability in time-series of seasonal aggregates for the variables of interest do not show linear progression. We have nevertheless investigated the time-dependency of these patterns through correlation analyses (Figures 8 and 9).

In order to corroborate the conditions described by local (IMD) observations, we have also characterised the CWB, and its contributing variables, using data from four global meteorological reanalyses: ERA-Interim, JRA-55, NASA MERRA2 and ERA5. Comparison of climatologies from the four reanalyses to local observations show that although large absolute biases exist in the gridded data sources, simple normalisation (corrective) procedures yield accurate representation of Mukteshwar climatology. This relative skill extends to reasonable estimation of interannual (standardised) seasonal anomaly patterns. Even limited discrepancies between local observations and reanalyses for individual time-steps, however, yield substantial discrepancies in results of the more sensitive procedure of assessing time-dependency.

The CWB component variable characterisation demonstrates that Mukteshwar and the adjacent Kumaun Lesser Himalaya (KLH) have a monsoonal precipitation regime. The annual temperature cycle has a larger amplitude than might otherwise be expected at its latitude (~29.5°N), owing to the high elevation (>2000m asl). Examination of both time-series of seasonally aggregated anomalies and the correlation analyses of the time-dependency of monthly variables show that at Mukteshwar, and the adjacent KLH, CWB variability is driven predominantly by precipitation, i.e. the supply side of the moisture balance equation. Variability in reference evapotranspiration (ET_0), i.e. the demand side of the equation, reflects a combination of the variability in daily mean temperature (T_{avg}) and diurnal temperature range (DTR). In light of the dominant role of precipitation in the CWB, we further investigated the climatology and time-dependency (correlation) of daily precipitation exceeding specific thresholds. These analyses showed that correlations of precipitation to time appear to follow that of medium and heavy wet days (24-hour accumulation of ≥10mm and ≥ 50mm). This dominance of large precipitation events has potentially worrying implications for local resource management and hazard mitigation if the distribution of rainfall shifts toward more large events and fewer gentle/sustained showers. At the local scale, soil is unlikely to be able to infiltrate large precipitation amounts in a short time period. If concentration of precipitation in intense events is coupled with prolonged dry spells between rainfall episodes, the capacity of soil to store sufficient moisture to meet uptake needs by vegetation – both crops and forests – will likely be exceeded. While particularly heavy precipitation can cause crop damage, general intensification of rainfall rates in the uplands will likely result in increased soil erosion and higher peak river discharge. This will complicate infrastructure operation downstream, in the Terai and lowland segments of the Ganges basin, as reservoir storage capacity and flood defences may not provide adequate buffers to intensification of the hydrological cycle.

[5.2] Relevance of CWB methodology for informing adaptive resource management more broadly

The CWB, as a metric of the equilibrium – or lack thereof – between atmospheric moisture supply (precipitation) and demand (potential or reference evapotranspiration) to and from the land surface, provides a very meaningful descriptor of hydroclimate conditions. Quantitative identification of alternating phases of CWB surplus and deficit within the annual cycle contextualises seasonality of local plant growth and water-dependent economic activities in moisture-limited (rather than energy-limited) cases. Time-series analyses of CWB anomalies provide insight on the magnitude, frequency and duration over which near surface atmospheric moisture availability is observed to deviate from mean conditions. Taken together the climatological and 'anomaly-space' approaches usefully frame the time-varying need for local moisture storage either within the natural subsurface – i.e. in soil and aquifers – or in engineered structures ranging from household-level tanks and ponds to regional networks of surface reservoirs and/or groundwater pumping.

In light of the findings regarding the dominance of precipitation and particularly large rainfall events in driving variability and evolution of CWB (as illustrated through the Mukteshwar observational record), it is pragmatic to suggest that local and regional initiatives to develop adaptive resource management should focus on increasing buffering capacity to attenuate moisture supply-demand imbalances. This could be pursued not only through the construction of surface water storage (tanks, reservoirs) and distribution systems, but also through land management activities and interventions to enhance infiltration (e.g. bunds) and soil moisture retention (e.g. increasing topsoil organic content) and to limit evapotranspiration (e.g. mulches). In the context of this study, such initiatives could be tested within the Ramgad and Dhokane watersheds (Figure 1) which lie within the Ramgarh Development Block in the Nainital district of Utttarakhand state, India. Developing systems and methods capable of coping with already high levels of interannual variability would represent an important step toward resilience to future climate change impacts on the water cycle.

These systems could be scalable in terms of both spatial service area and temporal buffering. In the most modest configuration, tanks and subsurface storage would be destined to bridge moisture supply shortfalls over a few days or weeks for the fields of individual smallholder farming families. More ambitious schemes could be designed to store ‘surplus’ monsoonal precipitation to meet moisture demands for the following several months for substantial sections of individual villages (panchayats).

Independent of the scale at which it is applied, the CWB approach, as demonstrated in this study, provides a scientifically robust approach to characterising near surface atmospheric moisture availability. Because it is conceptualised through supply and demand terms analogous to simple accounting principles, its broad strokes should also be accessible to lay-person decision makers who could draw upon its findings to guide adaptive resource management efforts.

Data Availability Statement

The local historical observations meteorological observations from Mukteshwar were obtained via agreement with the India Meteorological Department (IMD). IMD's permission must be obtained for the authors to (re)share this data. All global meteorological reanalyses data used in this study are available from public repositories maintained by their producers, e.g. the European Centre for Medium range Weather Forecasting.

Software availability statement

The software used in this study are simple implementations in Python of standard statistical functions and the FAO56 Penman Monteith method for calculating reference evapotranspiration along with (matplotlib) scripts to visual the results, i.e. generate figures. These fragments have not yet been aggregated into a formal repository. Reasonable requests for specific (sample) elements of the code will be satisfied by the corresponding author.

REFERENCES

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M., 1998. FAO Irrigation and Drainage Paper No. 56. FAO, Rome, Italy. 300pp. http://www.fao.org/docrep/X0490E/X0490E00.htm

Bosilovich, M.G., Chen, J., Robertson, F.R. and Adler, R.F., 2008. Evaluation of global precipitation in reanalyses. Journal of Applied Meteorology and Climatology, 47 (9), 2279-2299. DOI: 10.1175/2008JAMC1921.1

Burton, A., Kilsby, C.G., Fowler, H.J., Cowpertwait, P.S.P. and O’Connell, P.E. 2008. RainSim: A spatial temporal stochastic rainfall modelling system. Environmental Modelling and Software, 23 (12), 1356-1369. DOI:10.1016/j.envsoft.2008.04.003

Churchill, D.J., Larson, A.J., Dahlgreen, M.C., Franklin, J.F., Hessburg, P.F. and Lutz, J.A., 2013. Restoring forest resilience: From reference spatial patterns to silvicultural prescriptions and monitoring. Forest Ecology and Management, 291, 442-457. http://dx.doi.org/10.1016/j.foreco.2012.11.007

Cox, P.M., Huntingford, C. and Williamson, M.S., 2018. Emergent constraint on equilibrium climate sensitivity from global temperature variability. Nature, 553 (7688), 319-322. DOI: 10.1038/nature25450

Crimmins, S.M., Dobrowski, S.Z., Greenberg, J.A., Abatzoglou, J.T. and Mynsberge, A.R., 2011. Changes in climatic water balance drive downhill shifts in plant species' optimum elevations. Science, 331 (6015), 324-327. http://dx.doi.org/10.1126/science.1199040

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Holm, E.V., Isaksen, L., Kallberg, P., Kohler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.-N. and Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 137 (656), 553–597. http://dx.doi.org/10.1002/qj.828

Droogers, P. and Allen, R.G., 2002. Estimating reference evapotranspiration under inaccurate data conditions. Irrigation and Drainage Systems, 16 (1), 33-45. DOI: 10.1023/A:1015508322413

Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., Harada, Y., Yasui, S., Miyaoka, K., Takahashi, K., Kamahori, H., Kobayashi, C., Endo, H., Soma, M., Oikawa, Y. and Ishimizu, T.,, 2011. The Japanese 55-year Reanalysis "JRA-55": An Interim Report. Scientific Online Letters on the Atmosphere, 7, 149-152. http://dx.doi.org/10.2151/sola.2011-038

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J., 2012. (HESS Opinions) "Should we apply bias correction to global and regional climate model data?", Hydrol. Earth Syst. Sci., 16, 3391-3404, https://doi.org/10.5194/hess-16-3391-2012

Eyring, V., Cox, P.M., Flato, G.M., Gleckler, P.J., Abramowitz, G., Caldwell, P., Collins, W.D., Gier, B.K., Hall, A.D., Hoffman, F.M., Hurtt, G.C., Jahn, A., Jones, C.D., Klein, S.A., Krasting, J.P., Kwiatkowski, L., Lorenz, R., Maloney, E., Meehl, G.A., Pendergrass, A.G., Pincus, R., Ruane, A.C., Russell, J.L., Sanderson, B.M., Santer, B.D., Sherwood, S.C., Simpson, I.R., Stouffer, R.J. and Williamson, M.S., 2019. Taking climate model evaluation to the next level. Nature Climate Change, 9 (2), 102-110. DOI: 10.1038/s41558-018-0355-y

Forsythe, N., Hardy, A.J., Fowler, H.J. Blenkinsop, S., Kilsby, C.G., Archer, D.R. and Hashmi, M.Z., 2015. A detailed cloud fraction climatology of the Upper Indus Basin and its implications for near surface air temperature. Journal of Climate, 28 (9), 3537–3556. DOI:10.1175/JCLI-D-14-00505.1

Grunwald, S., 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. Geoderma, 152(3), 195-207. https://doi.org/10.1016/j.geoderma.2009.06.003

Guerreiro, S.B., Fowler, H.J., Barbero, R., Westra, S., Lenderink, G., Blenkinsop, S., Lewis, E., and Li, X.F., 2018. Nature Climate Change 8, 803–807. https://doi.org/10.1038/s41558-018-0245-3

Hargreaves, G.H., 1994. Defining and using reference evapotranspiration. Journal of Irrigation and Drainage Engineering, 120 (6), 1132-1139. DOI: 10.1061/(ASCE)0733-9437(1994)120:6(1132)

Hargreaves, G.H. and Allen, R.G., 2003. History and evaluation of Hargreaves evapotranspiration equation. Journal of Irrigation and Drainage Engineering, 129 (1), 53-63. DOI: 10.1061/(ASCE)0733-9437(2003)129:1(53)

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.-N., 2020. The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 146 (730), 1999-2049. DOI:10.1002/qj.3803.

Hong, S., Y. Noh, and J. Dudhia, 2006: A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. Mon. Wea. Rev., 134, 2318–2341, https://doi.org/10.1175/MWR3199.1

Huntington, T.G., 2006. Evidence for intensification of the global water cycle: Review and synthesis. Journal of Hydrology, 319 (1-4), 83-95. DOI: 10.1016/j.jhydrol.2005.07.003

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S.I., Sheffield, J., Goulden, M.L., Bonan, G., Cescatti, A., Chen, J., De Jeu, R., Dolman, A.J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B.E., Montagnani, L., Mu, Q., Mueller, B., Oleson, K., Papale, D., Richardson, A.D., Roupsard, O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S. and Zhang, K., 2010. Recent decline in the global land evapotranspiration trend due to limited moisture supply. Nature, 467 (7318), 951-954. DOI: 10.1038/nature09396

Kendon, E.J., Roberts, N.M., Senior, C.A. and Roberts, M.J., 2012. Realism of Rainfall in a Very High-Resolution Regional Climate Model. J. Climate, 25, 5791–5806, https://doi.org/10.1175/JCLI-D-11-00562.1

Kilsby, C.G., Jones, P.D., Burton, A., Ford, A.C., Fowler, H.J., Harpham, C., James, P., Smith, A. and Wilby, R.L., 2007. A daily weather generator for use in climate change studies. Environmental Modelling and Software, 22 (12), 1705-1719. DOI:10.1016/j.envsoft.2007.02.005

Knutti, R., Sedláček, J., Sanderson, B.M., Lorenz, R., Fischer, E.M. and Eyring, V., 2017. A climate model projection weighting scheme accounting for performance and interdependence. Geophysical Research Letters, 44 (4), 1909-1918. DOI: 10.1002/2016GL072012

Lawrimore, J.H., Menne, M.J., Gleason, B.E., Williams, C.N., Wuertz, D.B., Vose, R.S., and Rennie, J., 2011. An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3, J. Geophys. Res., 116, D19121, http://dx.doi.org/10.1029/2011JD016187.

Li, X.-F., Fowler, H.J., Forsythe, N., Blenkinsop, S., Pritchard, D, 2018. The Karakoram/Western Tibetan vortex: seasonal and year-to-year variability. Climate Dynamics, 51 (9-10), 883-3906. DOI:10.1007/s00382-018-4118-2

Lorenz, C. and Kunstmann, H., 2012. The hydrological cycle in three state-of-the-art reanalyses: Intercomparison and performance analysis. Journal of Hydrometeorology, 13 (5), 1397-1420. DOI: 10.1175/JHM-D-11-088.1

Milly, P.C.D., Dunne, K.A. and Vecchia, A.V., 2005. Global pattern of trends in streamflow and water availability in a changing climate. Nature, 438 (7066), 347-350. DOI: 10.1038/nature04312

Monteith, J.L., 1965. Evaporation and environment. Symposia of the Society for Experimental Biology, 19, 205-234.

Oki, T. and Kanae, S., 2006. Global hydrological cycles and world water resources. Science, 313 (5790), 1068-1072. DOI: 10.1126/science.1128845

Piani, C., Haerter, J.O. & Coppola, E., 2010. Statistical bias correction for daily precipitation in regional climate models over Europe. Theor Appl Climatol, 99: 187. https://doi.org/10.1007/s00704-009-0134-9

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O. and Feser, F., 2015. A review on regional convection permitting climate modeling: Demonstrations, prospects, and challenges. Rev. Geophys., 53, 323– 361. http://dx.doi.org/10.1002/2014RG000475.

Rahaman, M.M., 2009. Integrated Ganges basin management: conflict and hope for regional development. Water Policy, 11 (2), 168-190. http:/dx.doi.org/10.2166/wp.2009.012

Rasmussen, R., C. Liu, K. Ikeda, D. Gochis, D. Yates, F. Chen, M. Tewari, M. Barlage, J. Dudhia, W. Yu, K. Miller, K. Arsenault, V. Grubišić, G. Thompson, and E. Gutmann, 2011. High-Resolution Coupled Climate Runoff Simulations of Seasonal Snowfall over Colorado: A Process Study of Current and Warmer Climate. J. Climate, 24, 3015–3048, https://doi.org/10.1175/2010JCLI3985.1

Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R.D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C.R., Reichle, R., Robertson, F.R., Ruddick, A.G., Sienkiewicz, M. and Woollen, J., 2011. MERRA: NASA's modern-era retrospective analysis for research and applications. Journal of Climate, 24 (14), 3624-3648. http://dx.doi.org/10.1175/JCLI-D-11-00015.1

Serinaldi, F., Kilsby, C.G. and Lombardo, F., 2018. Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology. Advances in Water Resources, 111, 132-155. DOI: 10.1016/j.advwatres.2017.10.015

Sharma, B.R., Rao, K.V., Vittal, K.P.R., Ramakrishna, Y.S. and Amarasinghe, U., 2010. Estimating the potential of rainfed agriculture in India: Prospects for water productivity improvements. Agricultural Water Management, 97 (1), 23-30. http://dx.doi.org/10.1016/j.agwat.2009.08.002

Sheffield, J. and Wood, E.F., 2008. Global trends and variability in soil moisture and drought characteristics, 1950-2000, from observation-driven simulations of the terrestrial hydrologic cycle. Journal of Climate, 21 (3), 432-458. DOI: 10.1175/2007JCLI1822.1

Thornthwaite, C.W., 1948. An approach toward a rational classification of climate. Geographical Review, 38 (1), 55-94. DOI:10.2307/210739

Trenberth, K.E., Dai, A., Rasmussen, R.M. and Parsons, D.B., 2003. The changing character of precipitation. Bulletin of the American Meteorological Society, 84 (9), 1205-1217+1161. DOI: 10.1175/BAMS-84-9-1205

Trenberth, K.E., Dai, A., Van Der Schrier, G., Jones, P.D., Barichivich, J., Briffa, K.R. and Sheffield, J., 2014. Global warming and changes in drought. Nature Climate Change, 4 (1), 17-22. DOI: 10.1038/nclimate2067

Vose, R.S., Applequist, S., Menne, M.J., Williams, C.N., Thorne, P. 2012. An intercomparison of temperature trends in the U.S. Historical Climatology Network and recent atmospheric reanalyses. Geophysical Research Letters, 39 (10), art. no. L10703. DOI: 10.1029/2012GL051387

ACKNOWLEDGEMENTS

This research was initially funded by a grant (DST-UKIERI-2014-15-DST-122) from the British Council. Subsequent work was enabled by Global Challenges Research Fund (GCRF) grants administer by Royal Society for the CSAICLAWPS (CH160148) and PAPPADAAM (CHG\R1\170057) projects. Additionally: Nathan Forsythe, David Pritchard and Hayley Fowler were supported by the GCRF FutureDAMS project (grant ES/P011373/1) administered by the UK ESRC. Hayley Fowler was also funded by the Wolfson Foundation and the Royal Society as a Royal Society Wolfson Research Merit Award holder (grant WM140025). Professor Prakash C. Tiwari and Dr Bhagwati Joshi acknowledge the generous financial support provided by the Department of Science and Technology, Government of India New Delhi for carrying out the research included in the paper

SUPPLEMENTARY INFORMATION

Additional information on evaluation of reanalyses estimates of local conditions

[1] Time correlations between local observations and reanalyses estimates of key variables (Precip, T_{avg}, DTR)

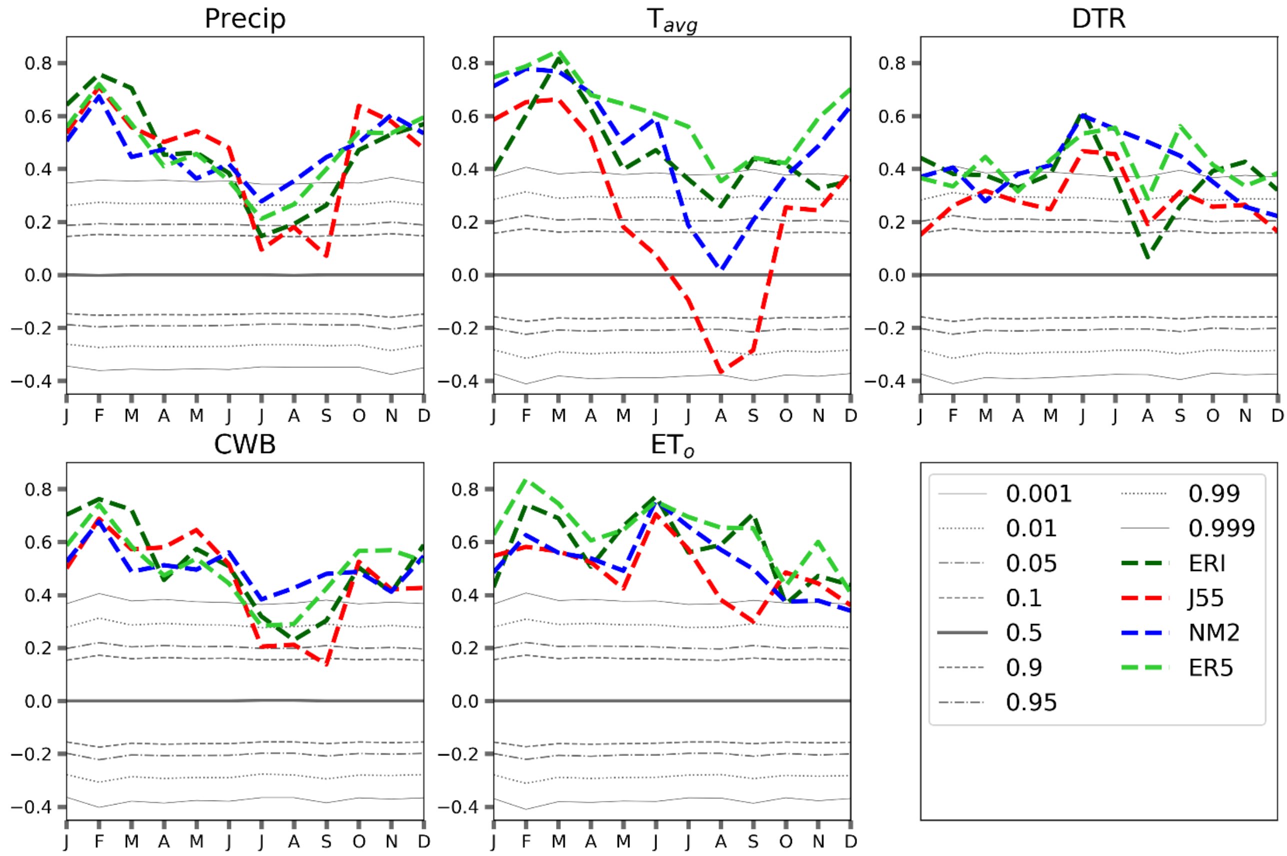

Figure S1 Kendall Tau correlation of reanalyses estimates of near surface climate variables to local observations (from Mukteshwar IMD). These correlations are based on monthly aggregated values. Grey lines indicate statistical distribution of correlation values resulting through randomisation of observation order/sequencing; ERI=ERA-interim, NM2=NASA MERRA2, J55=JRA-55, ER5=ERA5.