Supplementary Information

A Tri-Approach for Diagnosing Gridded Precipitation Datasets for Watershed Glacio-Hydrological Simulation in Mountain Regions

Muhammad Shafeeque[1,2], Yi Luo[1,2,3*]

[1]Key Lab of Ecosystem Network Observation and Modelling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, 100101 Beijing, China

[2]University of Chinese Academy of Sciences, 100049 Beijing, China

[3]Research Center for Ecology and Environment of Central Asia, Urumqi, China

[*]Corresponding author: Yi Luo (Email: luoyi@igsnrr.ac.cn)

1 Meteorological station data

Table S1. Details of meteorological stations in UIB.

Sr. No.	Station	Agency	Latitude (°)	Longitude (°)	Altitude (m)	Period	No. of years
1	Astore	PMD	35.366	74.865	2168	1954-2010	57
2	Bunji	PMD	35.646	74.629	1372	1953-2010	58
3	Burzil	WAPDA	34.899	75.079	4239	1995-2015	21
4	Deosai	WAPDA	35.004	75.592	4149	2000-2015	16
5	Gilgit	PMD	35.921	74.327	1460	1951-2010	60
6	Gupis	PMD	36.179	73.439	2156	1955-2010	56
7	Hushey	WAPDA	35.342	76.139	3075	1994-2007	11
8	Khunjrab	WAPDA	36.812	75.332	4730	2000-2015	16
9	Naltar	WAPDA	36.168	74.175	2898	1999-2011	13
10	Rama	WAPDA	35.455	74.776	3179	1999-2008	10
11	Rattu	WAPDA	35.161	74.785	2718	1995-2015	21
12	Shendure	WAPDA	36.088	72.547	3712	1995-2015	19
13	Skardu	PMD	35.286	75.563	2317	1952-2010	59
14	Ushkore	WAPDA	36.027	73.415	3051	1999-2015	17
15	Yasin	WAPDA	36.451	73.294	3280	1999-2015	16
16	Zani Pass	WAPDA	36.352	72.169	3839	1999-2007	13
17	Ziarat	WAPDA	36.829	74.418	3020	1999-2015	15
18	Qinghe	CMDSN	32.500	80.080	4279	1962-2012	51

2 Hydrological station data

Table S2. Details of hydrological stations and data in UIB.

Sr. No.	River	Station	Latitude	Longitude	Altitude	Period	No. of years
1	Indus	Kharamong	34.93	76.21	2542	1983-2010	28
2	Shyok	Yugo	35.18	76.10	2469	1974-2010	37
3	Shigar	Shigar	35.33	75.75	2438	1985-2002	18
4	Indus	Kachura	35.45	75.41	2341	1970-2010	41
5	Hunza	Dainyor Br	35.92	74.37	1370	1966-2010	45
6	Gilgit	Gilgit	35.92	74.30	1430	1970-2010	41
7	Gilgit	Alam Br	35.76	74.59	1280	1970-2010	41
8	Indus	Partab/Bunji Br	35.73	74.62	1250	1962-2010	49
9	Astore	Doyian	35.54	74.70	1583	1974-2010	37
10	Indus	Shatial Br	35.53	73.56	1040	1984-2010	27
11	UIB	Besham Qila	34.92	72.88	580	1969-2010	42

3 Temperature data

The average annual minimum and maximum temperature was -0.98 ºC and 11.58 ºC in UIB, respectively (Figure 1). A significant increasing trend (+0.06 ºC yr^{-1}) was observed in T_{avg} for 1985-2014. Increasing trends of T_{min} (+0.07 ºC yr^{-1}) was greater than that of T_{max} (+0.05 ºC yr^{-1}). It highlighted that minimum temperature was increasing at higher rates than maximum temperature.

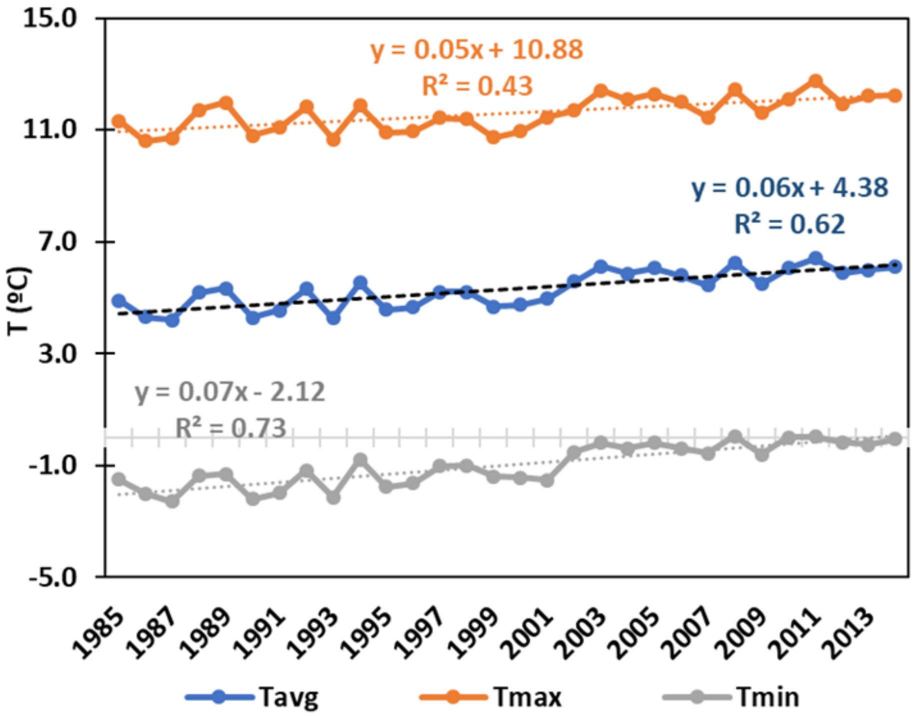

Figure S1. Annual average, maximum, and minimum temperature in entire UIB (average) for 1985-2014.

4 ET_p and ET_a data

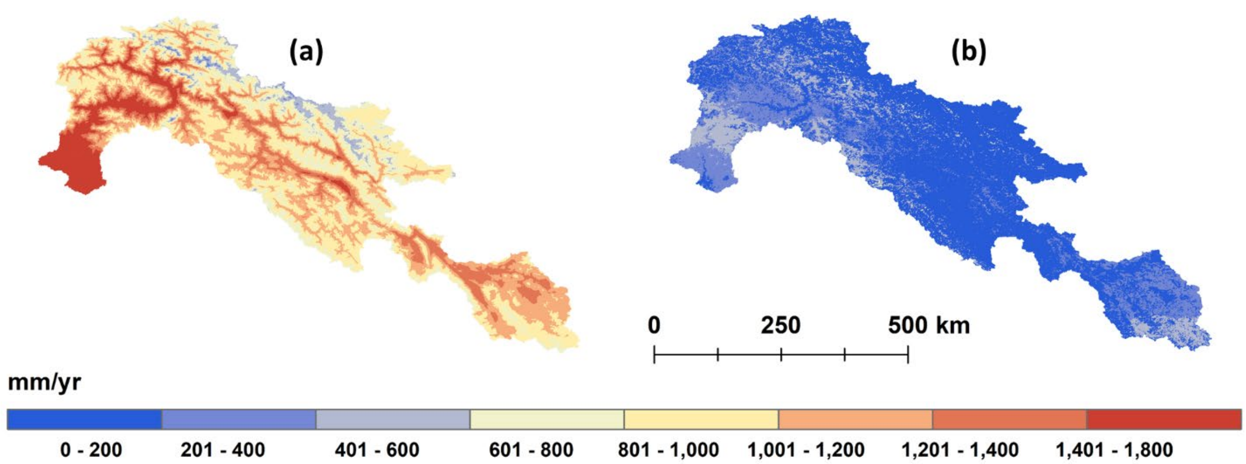

Figure S2. ETp extracted from Global Reference Evapo-Transpiration (Global-ET0), (b) ETa extracted from Esri_hydro "average annual actual evapotranspiration", Land use and land cover data.

5 Land use/land cover data

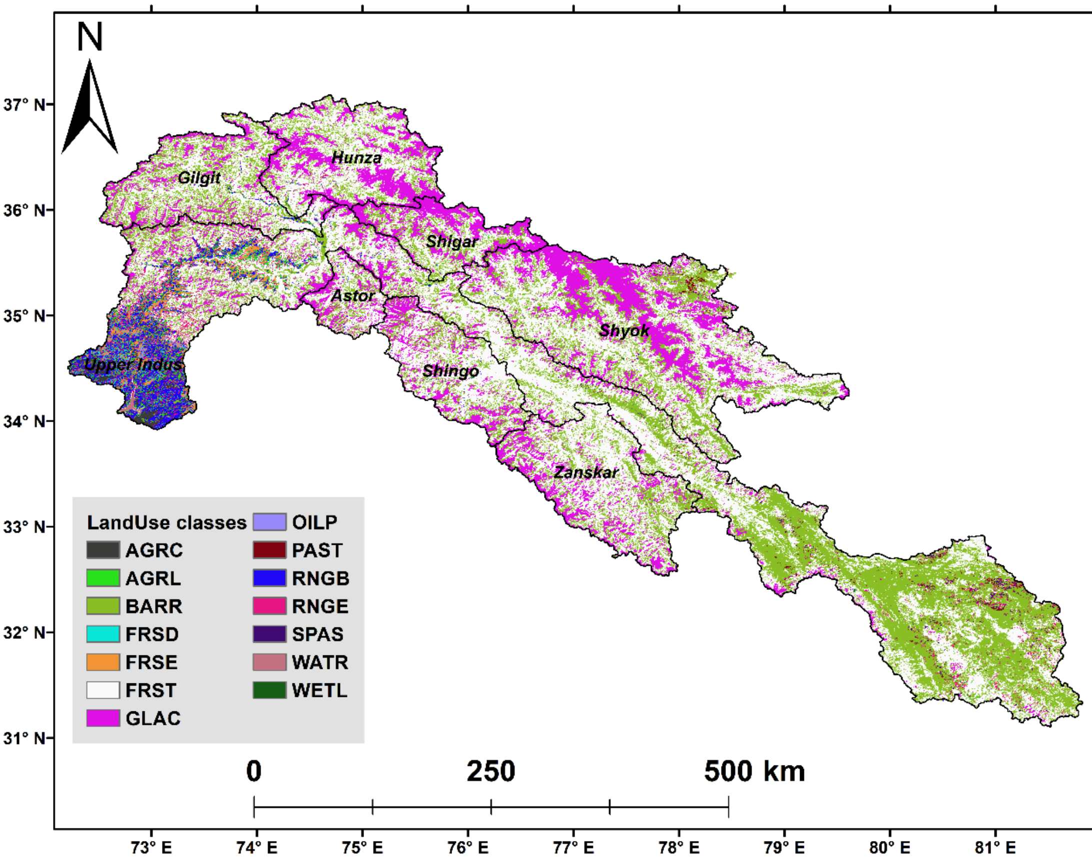

Figure S3. Land use classification in UIB based on Globecover data.

6 Soil data

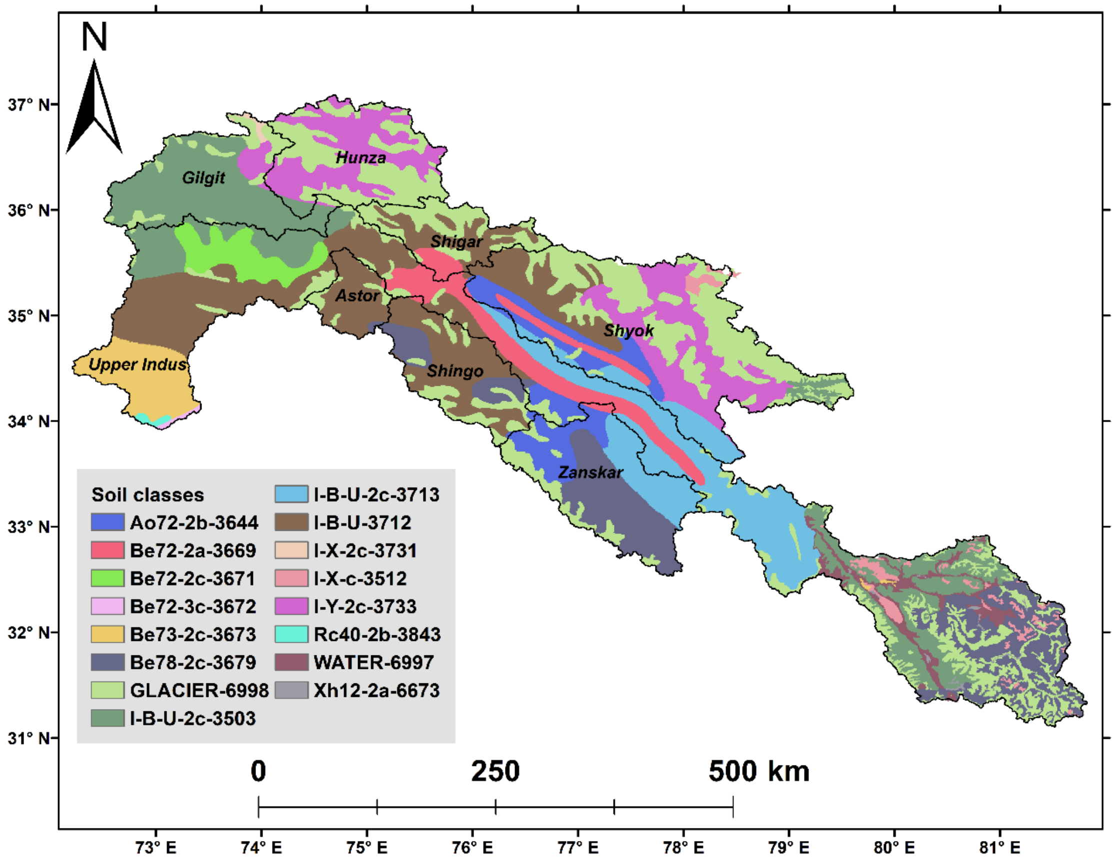

Figure S4. Soil map of UIB based on FAO HWSD V1.2.

7 Trend analysis

The Mann-Kendall (MK) test was applied for trend analysis of precipitation data series. The MK test is widely used for non-parametric analysis in hydrometeorological studies (Hirsch et al., 1991). This test makes no assumptions regarding data distribution and can be used for incomplete seasonal data with serial dependence and linear or non-linear trends. The MK statistic (*S*) was computed as follows:

$$S = \sum_{k=1}^{n-1} \times \sum_{j=k+1}^{n} Sgn(x_j - x_k) \qquad \textit{Eq. S1}$$

$$Sgn(x_j - x_k) = \begin{cases} +1, & (x_j - x_k) > 0 \\ 0, & (x_j - x_k) > 0 \\ -1, & (x_j - x_k) > 0 \end{cases} \qquad Eq.\ S2$$

where $x_1 \dots x_n$ is the climate variables ordered chronologically, and n is the number of points for analysis. For large datasets $n > 40$, the Z-test statistics are given as:

$$Z = \begin{cases} \dfrac{S-1}{\sqrt{Var(S)}}, & S > 0 \\ 0, & S = 1 \\ \dfrac{S-1}{\sqrt{Var(S)}}, & S < 0 \end{cases} \qquad Eq.\ S3$$

$$Var(S) = \frac{1}{18}[n(n-1)(2n+5)] \qquad Eq.\ S4$$

where $Var(S)$ is the variance of the data series.

The variance correction (VC) approach is the modified method of removing serial correlation, which is based on the fact that *N* serially correlated data contain the same information as N^* (< *N*) uncorrelated data. Yue et al. (2002) demonstrated through extensive Monte Carlo simulations that the presence of serial correlation in a time series does not alter the asymptotic normality of the MK test statistic S, nor does it modify the mean of S but it does change the dispersion of the distribution of *S*. The existence of positive (negative) serial correlation increases (decreases) the variance of S. Based on the work of Hamed and Rao (1998) and Yue and Wang (2004) proposed correcting the variance of the MK test statistic S by using an effective sample size that reflects the effect of serial correlation on the variance of S. The modified variance of the MK test statistic is given by;

$$Var(S)^* = CF \times Var(S) \tag{Eq. S5}$$

where $Var(S)$ is the variance of the MK test statistic S for the original sample data, and CF is a correction factor. The correction factors proposed by Hamed and Rao (1998) (denoted $CF1$) and Yue and Wang (2004) (denoted $CF2$) are;

$$CF1 = 1 + \frac{2}{N(N-1)(N-2)} \sum_{k=1}^{N-1} (N-k)(N-k-1)(N-k-2){r_k}^R \tag{Eq. S6}$$

$$CF2 = 1 + 2 \sum_{k=1}^{N-1} (1 - k/N) r_k \tag{Eq. S7}$$

$$r_k = \frac{\frac{1}{n-k} \sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{Eq. S8}$$

where r_k and ${r_k}^R$ are the lag-k serial correlation coefficient of data and ranks of data, respectively. $\bar{x}$ represents the mean of time series. To judge if the series data are serially correlated, the significance of r_k at the significance level of $\alpha = 0.01$ of the two-tailed test is assessed using the following approximation:

$$\frac{-1 - 1.645\sqrt{n-k-1}}{n-k} \leq r_k \leq \frac{-1 + 1.645\sqrt{n-k-1}}{n-k} \tag{Eq. S9}$$

If the lag-k serial correlation computed by Eq. S8 falls within the confidence interval given by Eq. S9, the data are assumed to be serially independent. Otherwise, the data are considered to be significantly serially correlated.

Using these equations, the modified test statistic can be obtained. By considering only r_1 and ${r_1}^R$, both approaches will work under the assumption of an AR (1) process. The MK test with the $CF1$ and $CF2$ correction is referred to as MK-CF1 and MK-CF2, respectively. In this study, these tests are applied with the AR (1) assumption, e.g., by taking $r_k = {r_1}^{|k|}$ for the MK-CF2 test, which used serial correlation coefficients of the data.

A robust estimate of the slope can be calculated using Sen's non-parametric method (Sen, 1968). In this case, the Sen's slope estimator is the median of the $(n-1)/2$ slopes of the pairs $(x_j : x_k)$ where $j > k$. The equation is as following:

$$\beta = Median\left(\frac{x_j - x_k}{j - k}\right), \forall k < j \qquad \textit{Eq. S10}$$

Table S3. Statistics for Mann-Kendall test results and Sen's Slope for precipitation datasets over varying periods. Z-value greater than 1.96 and less than -1.96 represents significant positive and negative trends, respectively.

Dataset	Index	HMLA	HNDKSH	KRKRM	UIB
OBS (1954-2015)	S	322	-58	382	326
	Var(S)	25823	27099	25823	25823
	Z-value	2	-0.35	2.37	2.02
	Sen's Slope	5.66	-0.41	6.42	4.19
APHRO (1951-2007)	S	-176	150	22	92
	Var(S)	21103	20020	21103	20020
	Z-value	-1.2	-1.05	0.14	0.64
	Sen's Slope	-0.46	1.8	0.13	0.4
CFSR (1979-2010)	S	1	-10	18	16
	Var(S)	3803	3803	3803	3803
	Z-value	0.01	-0.15	0.28	0.24
	Sen's Slope	0.02	-1.41	3.37	1.95
HAR (2001-2013)	S	20	30	8	26
	Var(S)	269	269	269	269
	Z-value	1.16	1.77	0.43	1.53
	Sen's Slope	13.57	39.92	6.53	20.89
PGMFD (1960-2016)	S	170	12	132	88
	Var(S)	21103	21103	21103	21103
	Z-value	1.16	0.08	0.9	0.6
	Sen's Slope	0.8	0.18	0.44	0.38
TRMM (1998-2017)	S	16	-18	58	10
	Var(S)	950	950	950	950
	Z-value	0.49	-0.55	1.85	0.29
	Sen's Slope	1.57	-6.42	5.24	1.59

8 Monthly performance evaluation-Taylor's diagrams

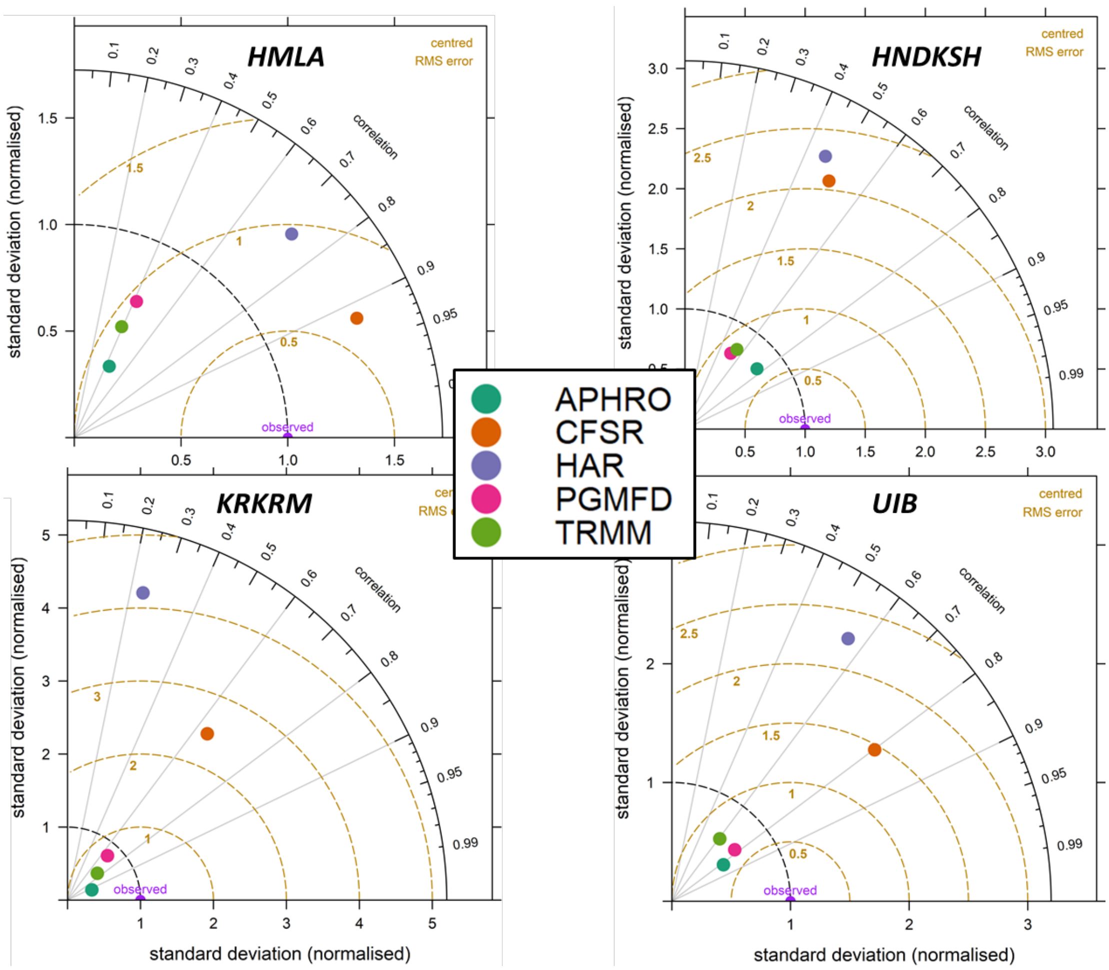

Figure S5. Taylor's diagrams to depict the monthly performance of GPDs.

9 References

Hamed, K.H., Rao, R.A., 1998. A modified Mann-Kendall trend test for autocorrelated data. Journal of Hydrology, 204(1-4): 182-196. DOI:10.1016/S0022-1694(97)00125-X

Hirsch, R.M., Alexander, R.B., Smith, R.A., 1991. Selection of methods for the detection and estimation of trends in water quality. Water Resources Research, 27(5): 803-813. DOI:10.1029/91WR00259

Sen, P.K., 1968. Estimates of the Regression Coefficient Based on Kendall's Tau. Journal of the American Statistical Association, 63(324): 1379-1389. DOI:10.1080/01621459.1968.10480934

Yue, S., Pilon, P., Phinney, B., Cavadias, G., 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. Hydrological Processes, 16(9): 1807-1829. DOI:10.1002/hyp.1095

Yue, S., Wang, C., 2004. The Mann-Kendall Test Modified by Effective Sample Size to Detect Trend in Serially Correlated Hydrological Series. Water Resources Management, 18(3): 201-218. DOI:10.1023/B:WARM.0000043140.61082.60