HI-AWARE

Research Component 1

Reference Climate Dataset for the Indus, Ganges and Brahmaputra River Basins

July 2015

Authors

A.F. Lutz (MSc.) Dr. W.W. Immerzeel

Report FutureWater: 146

FutureWater Costerweg 1V 6702 AA Wageningen The Netherlands

+31 (0)317 460050

info@futurewater.nl

www.futurewater.nl

Research





Table of contents

1	Introduction	5
2	Baseline reference climate data	6
2.1	Selection of baseline climate data	6
2.2	Upstream domain	9
2.3	Downstream domain	12
3	Correction of reference climate data	13
3.1	Upstream air temperature	13
3.2	Upstream precipitation	15
	3.2.1 Concept: precipitation lapse rates	15
	3.2.2 Region-wide glacier mass balance	17
	3.2.3 Model implementation and uncertainty analysis	19
3.3	Downstream climate	21
4	Corrected reference climate data	22
4.1	Upstream dataset	22
	4.1.1 Air temperature	22
	4.1.2 Precipitation	23
4.2	Total IGB domain	26
	4.2.1 Air temperature	27
	4.2.2 Precipitation	28
	4.2.3 Reference evapotranspiration	30
4.3	Subregional summaries and trends in time	31
5	Dataset metadata	33
5.1	Upstream IGB dataset	33
5.2	Total IGB dataset	33
6	References	34





Tables

. 7
. 7
11
12
13
18
20
31
1 1 2 3

Figures

Figure 1: Average annual precipitation sum 1998-2007 according to WFDEI corrected with CRU (upper panel), WFDEI corrected with GPCC (middle panel) and raw ERA-Interim (lower panel). Resolutions are the products' nominal resolutions
and vertical difference between the two DEMs at 1x1 km resolution (lower panel)
Figure 5: Uncorrected mean air temperature 1981-2010 (left) and uncorrected annual precipitation 1981-2010 (right) at 1x1 km resolution
Figure 7: Locations of meteorological stations in the upper IGB basins
Figure 9: Spatially interpolated bias between station observations and uncorrected air temperature grids 1981-2010 (left) and spatially interpolated bias derived from maximum ablation rate over glaciers 2000-2010.
Figure 10: Conceptual model of vertical and horizontal meteorological and cryospheric regimes in the Karakoram [<i>Hewitt</i> , 2007b]
Figure 11:Sites in the HKH region where geodetic mass balance data has been analysed by [<i>Gardelle et al.</i> , 2013]. Figure source: [<i>Gardelle et al.</i> , 2013]
indicated with black dots (left). Interpolated error in mass balance (right)
corrected precipitation (upper right). Corrected precipitation divided over uncorrected precipitation 2001-2010 (lower left). Mean precipitation gradient (lower right)





Figure 15: Mean air temperature 1981-2010 corrected upstream dataset
Figure 16: Monthly mean air temperature 1981-2010 corrected upstream dataset23
Figure 17: Mean annual precipitation sum 1981-2010 corrected upstream dataset and relative
distribution of precipitation in north-south and west-east direction
Figure 18: Mean monthly precipitation sum 1981-2010 corrected upstream dataset24
Figure 19: Precipitation during monsoon season (June-September) as percentage of total
annual precipitation (left) and precipitation during winter (December-February) as percentage of
total annual precipitation (right)25
Figure 20: Validation of corrected precipitation product for 15 subbasins to observed discharge
and estimated actual evapotranspiration
Figure 21: Air temperature 1981-2010 corrected dataset covering entire IGB river basins. Mean
air temperature (upper left), range between maximum and minimum air temperature (upper
right), maximum air temperature (lower left), minimum air temperature (lower right)27
Figure 22: Monthly mean air temperature 1981-2010 corrected dataset covering entire IGB river
basins
Figure 23: Mean annual precipitation sum 1981-2010 corrected dataset covering entire IGB
river basins
Figure 24: Mean monthly precipitation sum 1981-2010 corrected dataset covering entire 29
Figure 25: Mean annual reference evapotranspiration 1981-2010
Figure 26: Monthly mean reference evapotranspiration 1981-2010
Figure 27: Zonal averages of annual mean air temperature 1981-2010 for the upstream basins
(upper panel) and downstream basins (lower panel)
Figure 28: Zonal averages of annual precipitation sums 1981-2010





1 Introduction

High altitude climates are particularly uncertain and commonly used climate datasets are grossly inaccurate at high altitude. Therefore a novel reference climate dataset covering the Indus, Ganges and Brahmaputra (IGB) river basins is constructed, with particular focus on improved representation of high altitude precipitation.

This document describes the construction of a historical climate dataset for the IGB river basins, which is constructed for widespread use in the HI-AWARE project. The dataset covers the period from 1 January 1981 until 31 December 2010 with a daily time step and covers the entire IGB at 10x10 km spatial resolution. Additionally the upstream parts of the basins are covered at 5x5 km spatial resolution in a separate dataset to account for the larger variability in mountainous terrain. This document describes the methods to generate the datasets and illustrates the dataset contents.

Recently a method was developed that uses the presumed glacier mass balance to infer the high altitude precipitation, e.g. based on the size and mass balance of a glacier it is possible to estimate the amount of precipitation that is required to sustain this mass balance [*Immerzeel et al.*, 2012, 2015]. This approach is adapted and extended to the entire upstream parts of the IGB. For the downstream areas which are less affected by steep topography more straightforward geo-statistical interpolation techniques are used. These corrections resulted into a high quality, high resolution historical reference dataset spanning 30 years. The correction of the upstream domain and downstream domain are done separately, and the resulting products are merged and stored as NetCDF files, which are available for all consortium partners. At the end of the document the technical metadata of the dataset are listed.





2 Baseline reference climate data

2.1 Selection of baseline climate data

Understanding the spatial and temporal variability of precipitation in mountainous areas remains a key challenge. Point measurements are often not sufficient to capture the strong gradients in the multiple local factors that determine the distribution of precipitation. Climatologists have created numerous gridded datasets, based on observations. Since many of the existing gridded data products include precipitation and temperature at near surface level, they can be used to overcome data gaps in observations.

A distinction in two groups can be made regarding gridded datasets for temperature and precipitation: (i) datasets based created using advanced geo-statistical interpolation techniques based on station observations, (ii) datasets based on blending of climate model output and observations (often referred to as reanalysis products), and (iii) datasets based on satellite observations (remote sensing). Apart from differences in the underlying methodology (interpolation of observations or reanalysis) the main differences in the datasets are the spatial resolution, temporal resolution and time span covered.

A thorough comparison on the performance of existing gridded products for the HKH region [*Palazzi et al.*, 2013] highlights the striking differences between the different products. All the analyzed products are subject to limited spatial resolution. They are mostly suitable for large-scale continental studies. However, to analyse climate variations at smaller scales and in orographically complex regions, such as the IGB, they lack accuracy.

Researches who compared the performance of TRMM and APHRODITE over Nepal concluded that the latter is the more accurate dataset [*Duncan and Biggs*, 2012]. Other researchers also concluded that there is large variability in performance between different gridded products by comparing them for multiple transects crossing the Himalayan ranges [*Andermann et al.*, 2011]. They also conclude that APHRODITE, based on ground station data solely, gives the best precipitation estimates. However, they also mention that the lack of stations at high elevations limits the accuracy of this dataset. A study for the Upper Indus basin, also showed that high altitude precipitation in APHRODITE is strongly underestimated [*Lutz et al.*, 2014]. [*Immerzeel et al.*, 2015] compared four precipitation products for the Upper Indus basin and validated them to observed river discharge. According to their analysis, ERA-Interim provides the best estimate of precipitation in terms of annual totals, however the relatively coarse resolution limits its usability.

In 2014, the Refined High Asian Reanalysis (HAR) was released [*Maussion et al.*, 2014]. HAR is based on WRF model runs with an hourly time step, which are bounded daily to the ERA-INTERIM dataset. Although the product has a high spatial (10 km) and temporal (1 h) resolution, it covers a relatively short time range (2000-2012), and does not cover the entire IGB, since the western part of the upper Indus basin is not included.

The Watch Forcing ERA-Interim (WFDEI) dataset [*Weedon et al.*, 2014], is based on the WATCH methodology [*Weedon et al.*, 2011], integrated with the ERA-INTERIM dataset [*Dee et*



al., 2011]. Precipitation in the WFDEI dataset is bias-corrected using either the GPCC [*Schneider et al.*, 2013] dataset or the CRU dataset [*Harris et al.*, 2013].

Because ERA-Interim showed the most realistic precipitation totals as mentioned above, the decision was made to use an ERA-Interim based dataset as basis. Comparison of the ERA-Interim based WFDEI dataset to the raw ERA-Interim dataset showed that WFDEI has a higher spatial resolution than ERA-Interim and that WFDEI precipitation data that is bias-corrected using GPCC [*Schneider et al.*, 2013] shows more realistic spatial patterns, due to the correction with station data (Figure 1). It is desirable to use temperature data from the same dataset to ensure physical consistency between the two climatic variables (e.g. lower temperatures on rainy days), thus the WFDEI temperature data is used as basis.

Product	P Upper basins 1998-2007 (mm/yr)	P UIB 1998- 2007 (mm/yr)	P UGB 1998- 2007 (mm/yr)	P UBB 1998- 2007 (mm/yr)
WFDEI CRU	809	565	1004	1043
WFDEI GPCC	925	611	1488	1117
ERA-INT	1441	967	1704	1888

Table 1: Average annual precipiation sum 1998-2007.

Figure 2 indicates that the differences between the different air temperature datasets are very large. They are in the order of several degrees Celsius, with maximum differences around 10 degrees Celsius. Averaged over the upstream basins, the differences between the different datasets are also significant (Table 2).

Table 2: Mean air temperature 1998-2007	for the upstream	IGB according to	different
gridded air temperature products.			

Product	Mean T upstream IGB 1998-2007 (°C)
ERA-Interim	1.38
Aphrodite	5.34
Princeton	4.93
WFDEI	4.42





Figure 1: Average annual precipitation sum 1998-2007 according to WFDEI corrected with CRU (upper panel), WFDEI corrected with GPCC (middle panel) and raw ERA-Interim (lower panel). Resolutions are the products' nominal resolutions.





2.2 Upstream domain

Raw daily mean air temperature from the WFDEI dataset is spatially interpolated from 0.5°x0.5° (~50x50 km) to 1x1 km spatial resolution using a cubic spline interpolation and subsequently downscaled using a DEM at 1x1 km resolution (Figure 3, upper left) and vertical temperature lapse rates. Elevation differences (Figure 3, lower panel) between the DEM at 1x1 km resolution and the DEM used in WFDEI at 0.5°x0.5° resolution (Figure 3, upper right) determine the vertical distance over which the air temperature data is lapsed.







Figure 3: High-resolution 1x1 km DEM (upper left), WFDEI nominal 0.5x0.5 DEM (upper right), and vertical difference between the two DEMs at 1x1 km resolution (lower panel).

Temperature lapse rates vary locally, as under high and dry conditions the lapse rates are generally more steep than for humid conditions [*Kattel et al.*, 2012; *Immerzeel et al.*, 2014]. Therefore, the vertical lapse rate is determined locally at a monthly time scale. For this lapse rate derivation, lapse rates are determined at the $0.5x0.5^{\circ}$ grid cell level by doing a neighborhood operation for the grid cell under consideration and its 8 neighboring grid cells. A linear temperature-elevation relation is fitted using the 9 pairs of grid cell elevation and air temperature. This is done at a daily time step. Outliers are removed from the daily grids that are constructed this way. Values outside the range $\mu \pm 2\delta$ are considered as outliers. The resulting grid is spatially smoothed by averaging values over a 3 x 3 grid cells moving window. Daily grids are averaged over a month and the resulting monthly grids are used to downscale daily air temperature as:

$$T_{DS,d} = T_{WFDEI,d} + (DEM_{1km} - DEM_{WFDEI}) * \gamma_m$$

where $T_{DS,d}$ is daily downscaled air temperature, $T_{WFDEI,d}$ is daily temperature in WFDEI and γ_m is the monthly spatial grid with the vertical temperature lapse rate. Figure 4 shows the average vertical temperature lapse rates for January and December. The differences are generally as





expected, with steeper lapse rates observed during the dryer month of January and shallower lapse rates during July, when the monsoon occurs.



Figure 4: Average vertical temperature lapse rate for January (left) and July (right) 1981-2010.

Maximum and minimum air temperature are preprocessed in a very similar way. Since these data are not readily available in the WFDEI dataset at a daily time step, they are derived from the 3-hourly data, with the daily maximum air temperature being the maximum value of the eight 3-hourly values during the day and the minimum air temperature being the minimum value of the eight 3-hourly values. These temperature fields are downscaled using the same vertical lapse rates as used for the average air temperature. Although the lapse rates may actually be different for maximum, minimum and average air temperature in reality, the same lapse rates are used for each of these variables to ensure the consistency of the data. Otherwise the situation can occur that the average air temperature. Similarly the maximum air temperature could become lower than the minimum air temperature.

Precipitation data in WFDEI is only available differentiated as rain and snow. Therefore both fields are summed during preprocessing. Subsequently the data are interpolated from the $0.5^{\circ} \times 0.5^{\circ}$ to $1 \times 1 \text{ km}$ by a cubic spline interpolation.



Figure 5: Uncorrected mean air temperature 1981-2010 (left) and uncorrected annual precipitation 1981-2010 (right) at 1x1 km resolution.

Table 3: Basin-averaged air tem	perature and precipitatio	on 1981-2010 uncorrected	data.
Table 5. Dasin-averaged an term	iperature and precipitatio		uata.

	Upper Indus	Upper Ganges	Upper Brahmaputra
Mean T 1981-2010 (°C)	4.39	6.31	3.59
Mean annual P 1981-2010 (mm)	617	1497	1119





2.3 Downstream domain

Air temperature for the downstream domain is downscaled from $0.5^{\circ}x0.5^{\circ}$ (~50x50 km) resolution to 10x10 km resolution using a DEM at 10x10 km resolution in a similar way as for the upstream domain. In this case the vertical lapse rates are not derived from the data itself, but fixed lapse rate of -0.0065 °C m⁻¹ is assumed. Precipitation data are interpolated to 10x10 km resolution using a cubic spline interpolation, similar as for the upstream domain.



Figure 6: Mean air temperature (left) and annual precipitation sum (right) 1981-2010 for IGB domain

Table 4: Mean air temperature and mean precipitation 1981-2010 per basin.

	Mean T 1981-	Annual P 1981-
	2010 (°C)	2010 (mm)
Upper Indus	4.48	654
Lower Indus	23.97	372
Upper Ganges	4.96	1460
Lower Ganges	25.19	1139
Upper Brahmaputra	2.80	1051
Lower Brahmaputra	22.56	2842





3 Correction of reference climate data

3.1 Upstream air temperature

Air temperature data in the upstream domain is bias-corrected to data from station observations. Station observations in the upstream parts of the IGB basins are sparse. Figure 7 and Table 5 list the stations and station metadata, including record length, as used in this project. As evident from the map, the stations are very unequally distributed over the basin and mostly located in the valleys. As can be seen in the table, eight out of forty stations are located above 4000 m a.s.l., with the highest being located at an elevation of 4730 m a.s.l. In addition, many stations have rather short records available.



Figure 7: Locations of meteorological stations in the upper IGB basins.

ID	Name	Source	Lon (dd)	Lat (dd)	Elevation (m asl)	StartDate	EndDate
1	Burzil	WAPDA	75.088	34.911	4030	01/01/2000	31/12/2008
2	Khunjerab	WAPDA	75.400	36.850	4730	01/01/2000	31/12/2008
3	Naltar	WAPDA	74.189	36.158	2810	01/01/2000	31/12/2008
4	Rama	WAPDA	74.817	35.367	3000	01/01/2000	31/12/2008
5	Rattu	WAPDA	74.871	36.515	2570	01/01/2000	31/12/2008
6	Yasin	WAPDA	73.300	36.450	3150	01/01/2000	31/12/2008
7	Ziarat	WAPDA	74.276	36.836	3669	01/01/2000	31/12/2008
8	Astore	PMD	74.857	35.329	2168	01/01/2000	31/12/2005
9	Bunji	PMD	74.633	35.667	1470	01/01/2000	31/12/2005
10	Chilas	PMD	74.100	35.417	1251	01/01/2000	31/12/2005
11	Gilgit	PMD	74.333	35.917	1459	01/01/2000	31/12/2005
12	Gupis	PMD	73.400	36.230	2156	01/01/2000	31/12/2005

Table 5: Meteorological ground station records in the upper IGB used in HIAWARE.





13	Skardu	PMD	75.680	35.300	2210	01/01/2000	31/12/2005
14	Askole	PMD	75.815	35.681	3015	10/08/2005	31/12/2007
15	Urdukas	PMD	76.286	35.728	3927	06/17/2004	31/12/2007
16	Chitral	PMD	71.780	35.839	1500	01/01/2000	01/01/2005
17	Kotli	PMD	73.900	33.520	2017	01/01/2000	01/01/2005
18	Parachinar	PMD	70.083	33.867	1726	01/01/2000	01/01/2005
19	Khunjerab	Winiger/ICIMOD	74.417	36.850	4700	01/01/2000	12/31/2012
20	Bomi	ICIMOD	95.76	29.86	2736	01/01/2000	12/31/2006
21	Chayu	ICIMOD	97.46	28.65	2327.6	01/01/2000	12/31/2006
22	Cuona	ICIMOD	91.95	27.98	4280	01/01/2000	12/31/2006
23	Dangxiong	ICIMOD	91.1	30.48	4200	01/01/2000	12/31/2006
24	Jiacha	ICIMOD	92.58	29.15	3260	01/01/2000	12/31/2006
25	Jiali	ICIMOD	93.28	30.66	4488.8	01/01/2000	12/31/2006
26	Langkazi	ICIMOD	90.4	28.96	4431.7	01/01/2000	12/31/2006
27	Lazi	ICIMOD	87.63	29.08	4000	01/01/2000	12/30/2005
28	Lhasa	ICIMOD	91.13	29.67	3648.7	01/01/2000	12/31/2006
29	Linzhi	ICIMOD	94.47	29.57	3000	01/01/2000	12/31/2006
30	Namulin	ICIMOD	89.1	29.68	4000	01/01/2000	12/31/2006
31	Pali	ICIMOD	89.08	27.73	4300	01/01/2000	12/31/2006
32	Dungkhar	ICIMOD	91.1	27.82	2010	01/01/2000	12/31/2006
33	Phobijekha	ICIMOD	90.18	27.47	2860	01/01/2000	12/31/2006
34	Sunkosh	ICIMOD	90.07	27.02	410	01/01/2000	12/31/2006
35	Wamrong	ICIMOD	91.57	27.13	2180	01/01/2000	12/31/2006
36	Kakani	ICIMOD	85.25	27.8	2064	01/01/2000	12/31/2009
37	Taplejung	ICIMOD	87.66667	27.35	1732	01/01/2000	12/31/2010
38	Nielamu	ICIMOD	85.96	28.18	3310	01/01/2000	12/31/2006
39	Pulan	ICIMOD	81.25	30.28	3900	01/01/2000	12/31/2006
40	Shiquanhe	ICIMOD	80.08	32.5	4278	01/01/2000	12/31/2006

In a preceding project implemented by FutureWater for ICIMOD, station temperature data were corrected using a linear relationship that was found between the temperature bias in the Aphrodite dataset and elevation in the upper Indus basin [Lutz et al., 2014; Immerzeel et al., 2015]. In this case with WFDEI for the upstream IGB no such correlation could be established (Figure 8). Therefore the average biases at the station's locations were interpolated spatially to generate a spatial correction grid, which was applied to the uncorrected temperature fields. An additional bias-correction is done by using the degree-day glacier melt simulation component in SPHY [Terink et al., 2015] to simulate the distributed amount of melt over the glaciers using the temperature field that are bias-corrected to the station observations. These temperature fields are corrected downwards (to cooler temperatures) for unrealistic high amounts of melt. This is done by assuming a maximum annual ablation rate of 1.35 m we yr⁻¹ based on findings in scientific literature and field data from the Khumbu area [Immerzeel et al., 2011; Ragettli et al., 2015, Wagnon, personal communication]. Uncertainty in this assumption is taken into account in a Monte Carlo analysis of 100 runs with the uncertainty in the maximum annual ablation rate assumed to be Gaussian distributed with mean value 1.35 m we yr⁻¹ and 0.50 m standard deviation.





Figure 8: Average monthly bias between uncorrected temperature data and stations 1981-2010 plotted versus station elevation.



Figure 9: Spatially interpolated bias between station observations and uncorrected air temperature grids 1981-2010 (left) and spatially interpolated bias derived from maximum ablation rate over glaciers 2000-2010.

3.2 Upstream precipitation

3.2.1 Concept: precipitation lapse rates

The precipitation data is corrected using observed glacier mass balance data, according to the methodology developed in [*Immerzeel et al.*, 2012, 2015]. Based on geodetic measurements of glacier mass balance [*Gardelle et al.*, 2012, 2013], precipitation gradients are calculated to improve and downscale the uncorrected WFDEI precipitation fields.





Since the amounts of precipitation in the ground station data and gridded product are underestimated it is very likely that the precipitation necessary to supply the observed amount of discharged water is occurring at high altitudes. Research in this area [*"Batura Investigations Group,"* 1979; *Hewitt*, 2005, 2007a, 2011; *Winiger et al.*, 2005] suggests that precipitation increases up to 5000 to 6000 m a.s.l., where it is at its maximum, and decreases at higher altitudes (Figure 10, right panel).



Figure 10: Conceptual model of vertical and horizontal meteorological and cryospheric regimes in the Karakoram [*Hewitt*, 2007b].

In the construction of an improved gridded meteorological dataset for the upper IGB basins we implement this conceptual model to infer vertical precipitation lapse rates based on a linear increase of precipitation from a certain reference elevation (HREF) up to an elevation of maximum precipitation (HMAX) and decreasing linearly at higher altitudes with the same lapse rate [*Immerzeel et al.*, 2012, 2015].

In summary, the methodology to improve the data for precipitation is as follows:

- Observed geodetic mass balance data is used to construct a spatial mass balance grid covering the upstream IGB
- Using the downscaled and bias-corrected temperature fields (section 3.1), a distributed ablation model is applied to the glaciers in the IGB
- Local precipitation lapse rates are derived at glacier level to correct uncorrected precipitation data such that the observed mass balance can be sustained taking into account the simulated ablation
- Local precipitation lapse rates are spatially interpolated to correct precipitation for the entire upstream IGB.
- Corrected precipitation is aggregated from 1x1 km resolution to 5x5 km resolution

We assume that precipitation increases linearly with elevation up to an elevation with maximum precipitation and decreases with the same lapse rate above that elevation:





 $P_{COR}(x, y) = P_{WFDEI}(x, y) \cdot (1 + \left(\left(h(x, y) - h_{ref}(x, y)\right) \cdot \gamma \cdot 0.01\right)$

for $h < H_{MAX}$, and:

$$P_{COR}(x,y) = P_{WFDEI}(x,y) \cdot \left\{ 1 + \left(\left(\left(h_{MAX} - h_{ref}(x,y) \right) + \left(h_{MAX} - h(x,y) \right) \right) \cdot \gamma \cdot 0.01 \right) \right\}$$

for $h \ge H_{MAX}$

where P_{COR} is the corrected precipitation, P_{WFDEI} is the precipitation according to WFEI, h_{ref} is a reference elevation from which precipitation gradients occur, h is the elevation for the grid cell, and γ is the precipitation gradient (% m⁻¹).

3.2.2 Region-wide glacier mass balance

To calculate the precipitation gradients for individual glacier systems, we use geodetic mass balance data for eight sites in the HKH region [*Gardelle et al.*, 2012, 2013] (Figure 11). From these regions we select all glacier systems that have an area > 5 km², which are 346 individual systems in the 8 regions in the IGB (Table 6). Glaciers which are not completely covered by a geodetic mass balance grid are also removed. For each study site, [*Gardelle et al.*, 2013] used the Shuttle Radar Topographic Mission (SRTM) version 4 DEM, acquired mid-February 2000, as the reference topography. The elevation differences between the SRTM DEM and SPOT DEMs acquired between 2008 and 2011, depending on the study site, have been analysed at grid cell level and corrected for several biases except for seasonality (see [*Gardelle et al.*, 2013] for details). The elevation differences are converted to ice mass changes (meters water equivalent) using a recommended density of 850 kg m⁻³ [*Huss*, 2013]. Using glacier outlines from the ICIMOD glacier inventory [*Bajracharya and Shrestha*, 2011], the observed mass balance per glacier is calculated from the geodetic mass balance grids. Within each region, outliers are removed. Glaciers with average geodetic mass balance values deviating 2 or more standard deviations from the region mean are considered as outliers.





Figure 11:Sites in the HKH region where geodetic mass balance data has been analysed by [*Gardelle et al.*, 2013]. Figure source: [*Gardelle et al.*, 2013]

Table 6: Properties of sites in the HKH region where geodetic mass balance data has
been analysed by [<i>Gardelle et al.</i> , 2013].

Site	Name	Date of SPOT5 DEM	No. of glaciers > 5 km ²	Average MB (m we yr ⁻¹) [<i>Gardelle et</i> <i>al.</i> , 2013]	MB error (m we yr ⁻¹) [<i>Gardelle et</i> <i>al.</i> , 2013]	σ between glaciers (MB error * √n)
1	HinduKush	17-21 Oct 2008	24	-0.12	0.16	0.784
2	Karakoram West	3 Dec 2008	52	0.09	0.18	1.298
3	Karakoram East	31 Oct 2010	37	0.11	0.14	0.856
4	Spiti Lahaul	20 Oct 2011	59	-0.45	0.14	1.075
5	West Nepal	3 Jan 2011	27	-0.32	0.14	0.727
6	Everest	4 Jan 2011	43	-0.26	0.14	0.918
7	Bhutan	20 Dec 2010	45	-0.22	0.13	0.872
8	Hengduan Shan	24 Nov 2011	59	-0.33	0.14	1.075

The glacier mass balance per glacier is spatially interpolated by an inverse distance weighted interpolation and additional smoothing using a moving window averaging to obtain a spatial mass balance grid covering the entire IGB (Figure 12). Subsequently the glacier mass balance for glaciers in the IGB that are not included in one of the eight regions for which the geodetic glacier mass balance has been determined is taken from this spatially interpolated grid. The uncertainty in the mass balance data is provided for each region by [*Gardelle et al.*, 2013]. This error is spatially interpolated to get the error of the mass balance at interpolated locations (Figure 12). This error is taken into account in a Monte Carlo uncertainty analysis.







Figure 12: Spatially interpolated grid of glacier mass balance interpolated from glaciers indicated with black dots (left). Interpolated error in mass balance (right).

A glacier's mass balance is determined by the amount of accumulation and the amount of ablation:

$\Delta M = C - A$

where C is the accumulation and A is the ablation. For each of the glacier systems the ablation can be determined using the distributed degree day melt model in SPHY at 1x1 km resolution forced with the corrected gridded temperature fields. Calculating the glacier accumulation is a bit more complex, since the accumulation area of a glacier is often not entirely included in the glacier outlines in a glacier inventory. Especially in the HKH region, the glacier accumulation consists for large part of snow fed to the glacier surface by avalanching. To include this, we assume the accumulation area of a glacier system to include the grid cells covered by the glacier outline from the glacier inventory and in addition the adjacent grid cells that have their "drain" direction to the glacier surface and have a slope steeper than 0.20 m m⁻¹. This slope threshold is estimated from the slope distribution of the glacierised area in the UIB. Uncertainty in this assumption is included in a Monte Carlo uncertainty analysis.

3.2.3 Model implementation and uncertainty analysis

The model is implemented at 1x1 km spatial resolution, running from February 2000, which is the acquisition data of the SRTM DEM, until 31 December 2010, which is the last day included in the reference climate dataset. Depending on the acquisition date of the SPOT5 DEM, the mass balance state on that date is used to calculate the average simulated mass balance for each glacier. One hundred realizations are run, in which the model parameters are variated according to their uncertainties (Table 7). For each realization the model is run twice, with fixed precipitation gradients of 0.3 and 0.6 % m⁻¹. The precipitation gradient is then optimized by a linear regression through these two precipitation gradient values and the associated simulated mass balances to find the precipitation gradient which is required to simulate the observed glacier mass balance. The resulting precipitation gradients over the individual glaciers are spatially interpolated by a kriging operation.





Table 7: Parameter values used in precipitation correction model.

Parameter	Acronym	Distribution	Mean	SD
Reference elevation (m asl)	H _{REF}	log-Gaussian	2500	500
Maximum elevation (m asl)	H _{MAX}	log-Gaussian		
 Karakoram and Hindu Kush 			5500	500
- Other mountain ranges			4500	500
Degree day factor debris covered glacier (mm °C ⁻¹ d ⁻¹)	DDFdc	log-Gaussian	7	2
Degree day factor debris free glacier (mm °C ⁻¹ d ⁻¹)	DDFdf	log-Gaussian	2	2
Slope threshold (m m ⁻¹)	TS	log-Gaussian	0.2	0.05
Mass balance for individual glaciers	MB	Gaussian	Figure	Figure
			12 (left)	12 (right)
Maximum annual ablation (m we yr ⁻¹)	A _{MAX}	Gaussian	1.35	0.50

Figure 13 shows the mean corrected precipitation for 2000-2010 resulting from the 100 realizations. The strongest corrections are made in the eastern part of the Karakoram (Figure 14 lower left and lower right) with up to 7 times more precipitation in the corrected product compared to the original WFDEI. Largest uncertainties are present in the southern ranges of the Himalayas in the Brahmaputra basin (Figure 14 upper right).











Figure 14: Annual uncorrected precipitation 2000-2010 (upper left). Standard deviation of corrected precipitation resulting from 100 realizations Monte Carlo uncertainty analysis (upper right). Corrected precipitation divided over uncorrected precipitation 2001-2010 (lower left). Mean precipitation gradient (lower right).

The correction grid (Figure 14 lower left) is subsequently multiplied with the uncorrected daily precipitation grids from 1981-2010, which are then aggregated to 5x5 km to be used as model forcing for the upstream SPHY model in a later stage of the project.

3.3 Downstream climate

For the downstream parts of the IGB river basins a more straightforward correction can be applied since the biases in the WFDEI dataset are much smaller here because of the higher station density and less complex climate. Air temperature data is downscaled from $0.5^{\circ} \times 0.5^{\circ}$ (~50x50 km) to 10x10 km spatial resolution similarly as for the upstream domain by lapsing air temperature over the vertical difference between a DEM at 10x10 km resolution and a DEM at 0.5° x 0.5°. A fixed vertical lapse rate (-0.0065 °C m⁻¹) is applied. Precipitation fields in WFDEI which are already corrected to the GPCC are spatially interpolated from 0.5° x 0.5° (~50x50 km) to 10x10 km spatial resolution using a cubic spline interpolation.





4 Corrected reference climate data

This chapter summarizes the corrected reference climate datasets with figures illustrating the datasets' properties at grid cell level. Two datasets are delivered: one at 5x5 km spatial resolution for the upstream domain, and one at 10x10 km for the total domain. The 5x5 km upstream dataset is aggregated to 10x10 km resolution and combined with the 10x10 km downstream data to generate the dataset for the total domain.

4.1 Upstream dataset

4.1.1 Air temperature

The corrected air temperature dataset is on average colder than the uncorrected data. Strongest negative corrections are made for the upper Indus basin and the upper Brahmapatura basin, whereas the corrections are mostly slightly positive or neurtal for large parts of the upper Ganges basin.



Figure 15: Mean air temperature 1981-2010 corrected upstream dataset.







Figure 16: Monthly mean air temperature 1981-2010 corrected upstream dataset.

4.1.2 Precipitation

Strongest changes are present in the precipitation data, with strongest corrections being made in the Indus basin, indicating that precipitation there is being underestimated most severely in the uncorrected WFDEI dataset. The mean annual precipitation sum clearly shows the south to north and east to west gradients in the strength of the monsoon. Most precipitation falls in the southern and eastern ranges of the upstream domain. Looking at the intra-annual patterns in the data (Figure 18), shows that the seasonal patterns are well captured in the dataset. Most precipitation fall during the monsoon season (June-September) on the southern and eastern ranges. Also the changes in intensity of the monsoon during the monsoon season are well represented in the data. Precipitation in the most upstream parts of the Indus basin (Hindu Kush and Karakoram ranges) falls during the winter months, which is also well represented in the data.

This is also nicely illustrated when looking at the precipitation during the different season as percentage of the total annual precipitation (Figure 19). This clearly shows the southeast to northwest trend of decreasing monsoon dominance for the precipitation regime. The correct representation of the fluctuations in precipitation intensity in space and time is of utmost importance for the forcing of hydrological models, for which the forcing data is often the most uncertain component in the modeling assessment for the HKH region.







Figure 17: Mean annual precipitation sum 1981-2010 corrected upstream dataset and relative distribution of precipitation in north-south and west-east direction.



Figure 18: Mean monthly precipitation sum 1981-2010 corrected upstream dataset.







Figure 19: Precipitation during monsoon season (June-September) as percentage of total annual precipitation (left) and precipitation during winter (December-February) as percentage of total annual precipitation (right).

4.1.2.1 Validation to observed discharge and actual evapotranspiration

The corrected precipitation dataset is subjected to a first order validation with observed or estimated other components of the water balance at subbasin level. It is assumed that the catchment's corrected precipitation should equal the sum of the catchment's discharge and actual evapotranspiration and eventual increases in the stored water volume as glacier ice. When mass balance is negative, then the decrease in the volume of water stored as ice is considered as outward flux. We estimate average discharge from multiple multi-year discharge records. Note however, that in all cases the period of the climatic dataset (1981-2010) and the period of the discharge records are only partly overlapping. Actual evapotranspiration is estimated from PCRGLOBWB model output for 2003-2007 [*Wada et al.*, 2011]. Glacier mass balance is estimated with the trends for 2000-2010 used in this study (Figure 12).



Catchments







Figure 20: Validation of corrected precipitation product for 15 subbasins to observed discharge and estimated actual evapotranspiration.

Figure 20 shows that in most catchments, the uncorrected precipitation is by far not enough to have a closed water balance, whereas in most cases the water balance can be closed with the corrected precipitation data. This provides confidence that the corrected precipitation is close to reality in most cases, although there are cases where the corrected precipitation input is still too small (e.g. catchment ID's 6, 13, 18, 25, 35). On the other hand, it seems that the corrected precipitation overestimates the precipitation on the most northern parts of the basins, located on the Tibetan Plateau (e.g. catchment ID's 9, 30). However, an important remark to be made here is that fluxes such as infiltration to deep groundwater and sublimation, which can be significant in this area are not included in this validation, due to the lack of observations. Besides, the comparison with fluxes that only overlap a part of the reference climate dataset makes that this validation exercise we conclude that the corrected dataset is a significant improvement in the representation of precipitation in the upstream basins compared to the uncorrected data.

4.2 Total IGB domain

The dataset covering the total Indus, Ganges and Brahmaputra basins has 10x10 km spatial resolution and contains the upstream dataset as well, which has been aggregated from 5x5 km to 10x10 km spatial resolution. As visible in in Figure 21 and Figure 23, the datasets are merged seamlessly and provide a consistent corrected dataset for the entire IGB domain.





4.2.1 Air temperature

Figure 21 shows the mean temperature grids for the 30-years reference period. Means over the entire period are shown for mean air temperature, maximum air temperature, minimum air temperature and the diurnal range between the maximum air temperature and the minimum air temperature. As expected, this diurnal range is largest for the driest regions: the lower Indus and the upstream parts located on the Tibetan Plateau. The smallest ranges are in the wettest areas in the lower Brahmaputra basin (see also Figure 23). This shows that the temperature and precipitation datasets are physically consistent.



Figure 21: Air temperature 1981-2010 corrected dataset covering entire IGB river basins. Mean air temperature (upper left), range between maximum and minimum air temperature (upper right), maximum air temperature (lower left), minimum air temperature (lower right).

The monthly averages of the mean air temperature (Figure 22) show the expected seasonal patterns with colder winters and warmer summers. Strong gradients are present with cold temperatures in the upstream parts of the basins and high temperatures in the downstream parts.







Figure 22: Monthly mean air temperature 1981-2010 corrected dataset covering entire IGB river basins.

4.2.2 Precipitation

Figure 23 nicely illustrates the important effect that altitude has on precipitation patterns. The traverse from the Indo-Gangetic plain to the Himalaya, Karakoram and Hindu-Kush mountain ranges clearly shows the increasing precipitation with altitude. The spatial differences in precipitation quantities are clearly visible. South to north and and east to west gradients in the intensity of the monsoon are well captured during the monsoon season (June-September, Figure 24). As mentioned before in section 4.1.2, the high altitude winter precipitation in the upper Indus basin is also well represented in this corrected dataset.







Figure 23: Mean annual precipitation sum 1981-2010 corrected dataset covering entire IGB river basins.



Mean Monthly Precipitation Sum 1981-2010 (mm)

Figure 24: Mean monthly precipitation sum 1981-2010 corrected dataset covering entire IGB river basins.





4.2.3 Reference evapotranspiration

The reference evapotranspiration (ET_{ref}) is calculated for the total IGB at 10x10 km spatial resolution and daily time step by applying the Modified-Hargreaves equation [*Droogers and Allen*, 2002]. The Modified-Hargreaves equation (MH-equation) has the advantage over the widely used Penman-Monteith equation that it can be applied in low-data situations. The MH-equation requires extraterrestrial radiation, maximum air temperature and minimum air temperature:

ET_{ref} =0.0023 * 0.408 * Ra(Tavg+17.8) * TD^{0.5}

where Ra $(MJm^{-2} day^{-1})$ is the extraterrestrial radiation, Tavg (°C) the average daily air temperature, and TD (°C) the daily temperature range, defined as the difference between the daily maximum and minimum air temperature. The constant 0.408 is required to convert the units to mm, and Ra can be obtained from equations using the day of the year and the latitude of the grid cell. The reference evapotranspiration data is included in the upstream dataset as well as the downstream dataset.



Figure 25: Mean annual reference evapotranspiration 1981-2010.







Figure 26: Monthly mean reference evapotranspiration 1981-2010.

4.3 Subregional summaries and trends in time

Table 8 shows the mean air temperature and mean annual precipitation for six subregions in the IGB domain, averaged over the entire reference period (1981-2010). The differences between the corrected data and the uncorrected data for the upstream basins are striking. Precipitation amounts are much higher (compare Table 3 and Table 8), with the largest correction being made for the upper Indus basin. Air temperatures have been corrected to cooler temperatures. In the upstream basins.

•		• •
	Mean T	Annual P
	1981-2010 (°C)	1981-2010 (mm)
Upper Indus	-1.6	1052
Lower Indus	23.8	373
Upper Ganges	2.4	1810
Lower Ganges	25.1	1139
Upper Brahmaputra	-1.1	1424
Lower Brahmaputra	21.8	2847

Table 8: Zonal averages of corrected climatic forcing per subbasin.





Mean air temperature shows increasing trends in all basins during the reference period (Figure 27). The dataset also captures the observed elevation dependent warming [*Rangwala and Miller*, 2012; *Pepin et al.*, 2015], with steeper increasing trends for temperature in the upstream basins, compared to the downstream parts of the basins.

Linear regression of precipitation trends show slightly decreasing or neutral trends for the subregions (Figure 28). These trends must be interpreted with care, since the significance of these trends is questionable. Precipitation is characterized by a larger inter-annual variability, which makes it more difficult to infer significant trends.



Figure 27: Zonal averages of annual mean air temperature 1981-2010 for the upstream basins (upper panel) and downstream basins (lower panel).









5 Dataset metadata

5.1 Upstream IGB dataset

Projection	WGS 84 UTM Zone 45N (EPSG:32645)			
Extent	xmin: -1330000		ymin: 2940000	
	xmax: 1	570000	ymax: 419000	
Spatial resolution	5000 x 5000 meter			
	(580 columns, 250 rows, 145000 grid cells)			
Temporal resolution and timespan	Daily time step, 1 Jan 1981 – 31 Dec 2010			
Variables and units	prec	Daily precipitation sum n		mm
	tavg	Daily mean air temperature °C		°C
	tmax	Daily maximum air temperature °C		°C
	tmin	Daily minimum air temperature °C		°C
	eref	f Reference evapotranspiration mm		
Data format	NetCDF (1 file per year and per variable)			
Format of filenames	HIAWARE_IGBupstr_variable_year.nc			

5.2 Total IGB dataset

Projection	WGS 84 UTM Zone 45N (EPSG:32645)			
Extent	xmin: -1	600000	ymin: 2300000	
	xmax: 1	600000	ymax: 420000	
Spatial resolution	10000 x 10000 meter			
	(320 columns, 190 rows, 60800 grid cells)			
Temporal resolution and timespan	Daily time step, 1 Jan 1981 – 31 Dec 2010			
Variables and units	prec	Daily precipitation sum		mm
	tavg	Daily mean air temperature°C		°C
	tmax	Daily maximum air temperature °C		°C
	tmin	Daily minimum air temperature °C		°C
	eref	Reference evapotranspiration mm		
Data format	NetCDF (1 file per year and per variable)			
Format of filenames	HIAWARE_IGBtotal_variable_year.nc			





6 References

- Andermann, C., S. Bonnet, and R. Gloaguen (2011), Evaluation of precipitation data sets along the Himalayan front, *Geochemistry, Geophys. Geosystems*, *12*(7), 1–16, doi:10.1029/2011GC003513.
- Bajracharya, S. R., and A. B. Shrestha (2011), The Status of Glaciers in the Hindu Kush-Himalayan Region.
- "Batura Investigations Group" (1979), The Batura Glacier in the Karakoram Mountains and its variations., *Sci. Sin.*, 22(8), 958–974.
- Dee, D. P. et al. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*(656), 553–597, doi:10.1002/qj.828.
- Droogers, P., and R. G. Allen (2002), Estimating reference evapotranspiration under inaccurate data conditions, *Irrig. Drain. Syst.*, *16*, 33–45.
- Duncan, J. M. A., and E. M. Biggs (2012), Assessing the accuracy and applied use of satellitederived precipitation estimates over Nepal, *Appl. Geogr.*, 34, 626–638, doi:10.1016/j.apgeog.2012.04.001.
- Gardelle, J., E. Berthier, and Y. Arnaud (2012), Slight mass gain of Karakoram glaciers in the early twenty-first century, *Nat. Geosci.*, *5*(5), 322–325, doi:10.1038/ngeo1450.

Gardelle, J., E. Berthier, Y. Arnaud, and a. Kääb (2013), Region-wide glacier mass balances over the Pamir-Karakoram-Himalaya during 1999–2011, *Cryosph.*, 7(4), 1263–1286, doi:10.5194/tc-7-1263-2013.

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister (2013), Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, *Int. J. Climatol.*, doi:10.1002/joc.3711.

Hewitt, K. (2005), The Karakoram Anomaly? Glacier Expansion and the "Elevation Effect," Karakoram Himalaya, *Mt. Res. Dev.*, *25*(4), 332 – 340.

Hewitt, K. (2007a), Tributary glacier surges: an exceptional concentration at Panmah Glacier, Karakoram Himalaya, *J. Glaciol.*, *53*(181), 181–188, doi:10.3189/172756507782202829.

Hewitt, K. (2007b), Tributary glacier surges: an exceptional concentration at Panmah Glacier, Karakoram Himalaya, *J. Glaciol.*, *53*(181), 181–188, doi:10.3189/172756507782202829.

- Hewitt, K. (2011), Glacier Change, Concentration, and Elevation Effects in the Karakoram Himalaya, Upper Indus Basin, *Mt. Res. Dev.*, *31*(3), 188–200, doi:10.1659/MRD-JOURNAL-D-11-00020.1.
- Huss, M. (2013), Density assumptions for converting geodetic glacier volume change to mass change, *Cryosph.*, 7(3), 877–887, doi:10.5194/tc-7-877-2013.
- Immerzeel, W. W., L. P. H. van Beek, M. Konz, A. B. Shrestha, and M. F. P. Bierkens (2011), Hydrological response to climate change in a glacierized catchment in the Himalayas, *Clim. Change*, *110*, 721–736, doi:10.1007/s10584-011-0143-4.





- Immerzeel, W. W., F. Pellicciotti, and A. B. Shrestha (2012), Glaciers as a Proxy to Quantify the Spatial Distribution of Precipitation in the Hunza Basin, *Mt. Res. Dev.*, *32*(1), 30–38, doi:10.1659/MRD-JOURNAL-D-11-00097.1.
- Immerzeel, W. W., L. Petersen, S. Ragettli, and F. Pellicciotti (2014), The importance of observed gradients of air temperature and precipitation for modeling runoff from a glacierized watershed in the Nepalese Himalayas, *Water Resour. Res.*, *50*, WR014506, doi:10.1002/2013WR014506.Received.
- Immerzeel, W. W., N. Wanders, A. F. Lutz, J. M. Shea, and M. F. P. Bierkens (2015), Reconciling high altitude precipitation with glacier mass balances and runoff, *Hydrol. Earth Syst. Sci. Discuss.*, *12*, 4755–4784, doi:10.5194/hessd-12-4755-2015.
- Kattel, D. B., T. Yao, K. Yang, L. Tian, G. Yang, and D. Joswiak (2012), Temperature lapse rate in complex mountain terrain on the southern slope of the central Himalayas, *Theor. Appl. Climatol.*, *113*(3-4), 671–682, doi:10.1007/s00704-012-0816-6.
- Lutz, A. F., W. W. Immerzeel, and P. D. A. Kraaijenbrink (2014), *Gridded Meteorological Datasets and Hydrological Modelling in the Upper Indus Basin. FutureWater Report 130*, Wageningen, The Netherlands.
- Maussion, F., D. Scherer, T. Mölg, E. Collier, J. Curio, and R. Finkelnburg (2014), Precipitation Seasonality and Variability over the Tibetan Plateau as Resolved by the High Asia Reanalysis, *J. Clim.*, *27*(5), 1910–1927, doi:10.1175/JCLI-D-13-00282.1.
- Palazzi, E., J. Von Hardenberg, and A. Provenzale (2013), Precipitation in the Hindu-Kush Karakoram Himalaya : Observations and future scenarios, *J. Geophys. Res. Atmos.*, *118*, 85–100, doi:10.1029/2012JD018697.
- Pepin, N. et al. (2015), Elevation-dependent warming in mountain regions of the world, *Nat. Clim. Chang.*, *5*, 424–430, doi:10.1038/nclimate2563.
- Ragettli, S., F. Pellicciotti, W. W. Immerzeel, E. Miles, L. Petersen, M. Heynen, J. M. Shea, D. Stumm, S. Joshi, and A. B. Shrestha (2015), Unraveling the hydrology of a Himalayan watershed through integration of high resolution in- situ data and remote sensing with an advanced simulation model, *Adv. Water Resour.*, *78*, 94–111, doi:10.1016/j.advwatres.2015.01.013.
- Rangwala, I., and J. R. Miller (2012), Climate change in mountains: A review of elevationdependent warming and its possible causes, *Clim. Change*, *114*, 527–547, doi:10.1007/s10584-012-0419-3.
- Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, and B. Rudolf (2013), GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, *Theor. Appl. Climatol.*, 26, doi:10.1007/s00704-013-0860-x.
- Terink, W., A. F. Lutz, G. W. H. Simons, W. W. Immerzeel, and P. Droogers (2015), SPHY v2.0: Spatial Processes in HYdrology, *Geosci. Model Dev.*, *8*, 2009–2034, doi:10.5194/gmd-8-2009-2015.
- Wada, Y., L. P. H. van Beek, and M. F. P. Bierkens (2011), Modelling global water stress of the recent past: on the relative importance of trends in water demand and climate variability, *Hydrol. Earth Syst. Sci.*, 15(12), 3785–3808, doi:10.5194/hess-15-3785-2011.





- Weedon, G. P., S. Gomes, P. Viterbo, W. J. Shuttleworth, E. Blyth, H. Österle, J. C. Adam, N. Bellouin, O. Boucher, and M. Best (2011), Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century, *J. Hydrometeorol.*, *12*(5), 823–848, doi:10.1175/2011JHM1369.1.
- Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo (2014), The WFDEI meteorological forcing data set:WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, *Water Resour. Res.*, 7505–7514, doi:10.1002/2014WR015638.Received.
- Winiger, M., M. Gumpert, and H. Yamout (2005), Karakorum-Hindukush-western Himalaya: assessing high-altitude water resources, *Hydrol. Process.*, *19*(12), 2329–2338, doi:10.1002/hyp.5887.

