

Point by point response to the comments

Referee No. 1

Scenario approach for the seasonal forecast of Kharif flows from Upper Indus Basin

Muhammad Fraz Ismail and Wolfgang Bogacki

bogacki@hs-koblenz.de

Dear Referee No. 1

We thank you very much for your valuable comments. These will definitely improve our paper. We have tried to answer your main remarks by splitting them into separate comments, which as we hope give a correct interpretation of your major points.

COMMENT: The authors do not review different methods that have been proposed or are in operational use for developing seasonal forecasts. I would expect some discussion on statistical methods as opposed to methods that use numerical weather prediction models. At the minimum the Ensemble Streamflow Prediction (ESP) approach, which has been widely used for seasonal water resources predictions should be considered.

RESPONSE: In the revised manuscript in section 2.6 *Scenario Approach of forecasting*, we have discussed the differences and similarities of Ensemble Stream Flow Prediction (ESP) method and the scenario approach of forecasting.

COMMENT: I could not identify in the results and discussion section the resulting ensembles or evaluation of the results of the scenarios selected. In assessing the skill of the forecasts I think the authors could significantly extend the current discussion and results presented. (...) A more elaborate verification of forecasts would be insightful.

RESPONSE:

In the new section 3.4 *Evaluation of forecast skills*, we have included a more detailed discussion on the forecast skills and give a yearly comparison of the forecasts (Table 6) as well some verification metrics (Table 7). We also have included examples of plume diagrams of ensemble member trace have been given in the Figure 14 and 15 in the revised manuscript.

COMMENT: It is not that clear in the manuscript how these [model parameters] have been derived during the calibration. Was this done as a pure calibration exercise, or was some physical basis used to estimate parameter values?

RESPONSE: In section 2.5 *Model parameters* are explained as well as in Table 3 of the revised manuscript all the parameters and their ranges are given.

Table 3: SRM+G Model Parameters for both Upper and Lower UIB

Parameters	Symbol	Value	Units	Remarks
Temperature Lapse-Rate	γ	6.0	$^{\circ}\text{C km}^{-1}$	
Recession Coefficient	k_x	1.193	–	October-February
		1.060		March – September
	k_y	0.029		October-February
		0.020		March – September

Parameters	Symbol	Value	Units	Remarks
Critical Precipitation	P_{crit}	1	cm	constant
Lag Time	L	54	h	2.5 days delay between melt and runoff at Tarbela
Critical Temperature	T_{crit}	0.5 – 3.0	°C	variable
Rainfall Contributing Area	RCA	0 1	–	November – March April – October
Runoff Coefficient - Snow	c_S	0.80	–	constant
Runoff Coefficient - Glacier	c_G	0.70	–	constant
Runoff Coefficient - Rain	c_R	0.25-0.75	–	
Degree-Day Factor - Snow	α	0.15-0.80	cm °C ⁻¹ d ⁻¹	
Degree-Day Factor - Glacier	a_G	0.70	cm °C ⁻¹ d ⁻¹	constant

COMMENT: What is interesting is the linear relationship, with [degree-day] factors increasing during the season. I would like some more discussion on how realistic this is. Is there a physical reason that these melt-factors increase to such an extent, or is this the result of the calibration procedure?

RESPONSE: A detailed discussion for the linear increase of the degree-day factors is given in section 2.5 *Model parameters*. To summarise, the increase of the degree-day factors with the passage of time is because the snow absorbs energy due to its physical condition, in terms of increasing temperatures and solar radiations intensities. This process of energy storage plays a pivotal role in the ripening of the snowpack, which melts rapidly as the snow melting season progresses. The extent to which degree-day factors increase is related to the calibration procedure because it was observed during the model calibration that in a certain elevation zone when the degree-day factors attain the value e.g (0.80 cm °C⁻¹ d⁻¹), the snow cover area in that very elevation zone has almost completely faded away so there is no need to further increase the values of degree-day factors. The limit to what extent the degree-day factors increase at a certain spatio-temporal region depends upon various physiographic and climatic parameters and a research is on-going for evaluating the trend of degree-day factors in response to the aforementioned parameters.

COMMENT: While RFE is chosen as an input, which given the scarcity of data is to my mind a reasonable choice, it is not so clear if there is an under-prediction of precipitation input. Comparison to the few ground stations may not help as these are, again as the authors note, likely underestimating the precipitation.

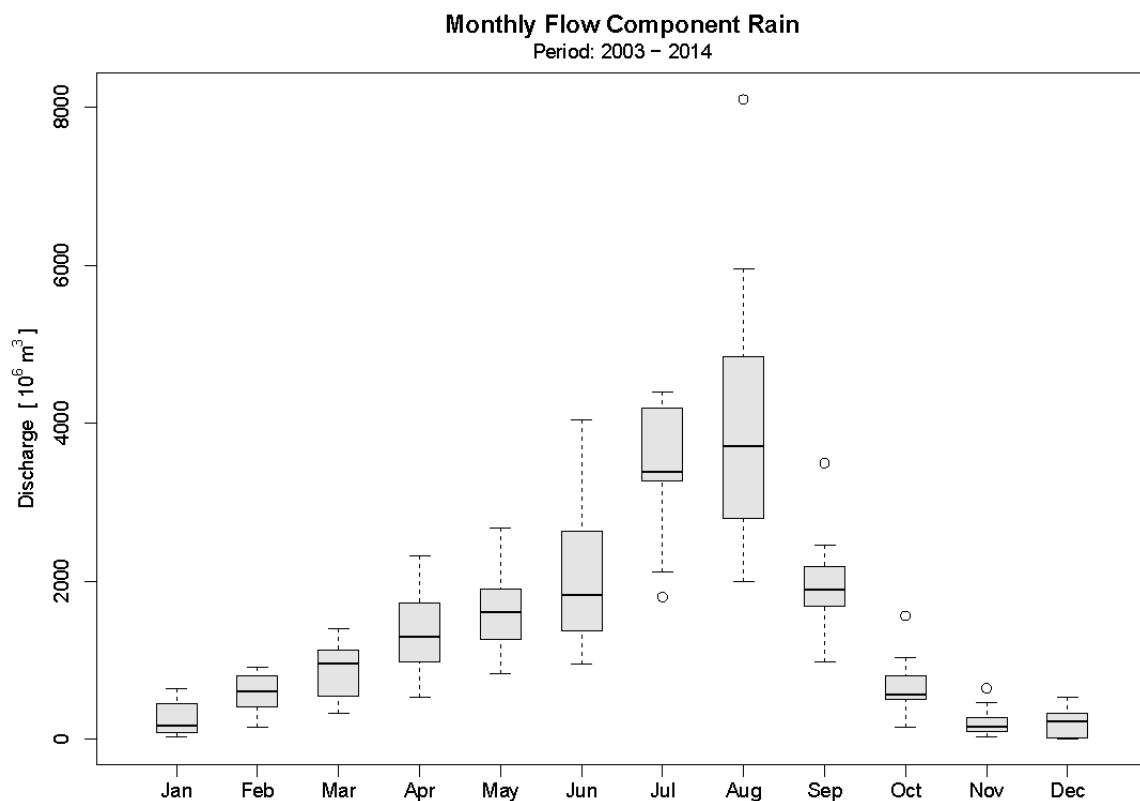
RESPONSE: We fully agree that, as others also have noted, ground stations due to their location at the valley floors tend to underestimate the actual precipitation. We therefore have referred in section 2.4 *Data sources* to a paper by Reggiani and Rientjnes (2015) who have compared a number of different studies with own calculations and estimate the mean annual precipitation in UIB to 675 ±100 mm/a, which is higher than in most of the other studies. The RFE basin-wide annual mean for the period 2003 – 2015 is 701 mm/a, which corresponds to the ERA-Interim (681 mm/a) and NCEP/NCAR (705 mm/a) reanalysis means calculated by Reggiani and Rientjnes.

COMMENT: The hydrographs, the authors present in figures 7&8 do not clearly show the contribution of the different sources of runoff that are considered in the model; direct runoff from

precipitation; snowmelt; and glacier melt. (...) Although the model itself is not conservative, and does not include a base-flow component such as subsurface flow, the authors could evaluate the water balance across the fourteen-year period they have selected.

RESPONSE: It is not possible to split the calculated daily discharge into the different sources of runoff, as due to SRM's recession flow approach (eq. 1 in the revised manuscript) only a small part (about 10%) contributes directly to the daily discharge while the larger part originates from recession flow Q_n . In this respect the model is not mass-conservative. Annual water balances give an average contribution of 26%, 53%, and 21% from rain, snowmelt, and glacier melt respectively, which coincides well with figures given by Immerzeel et al. (2010) or Charles et al. (2017).

However, we have now included the Figure 7 in the revised manuscript to show the monthly distribution of flow component in the UIB from 2003 - 2014. We have also updated the Figures 10 and 11 in order to make them more elaborative by showing the validation of SRM+G.



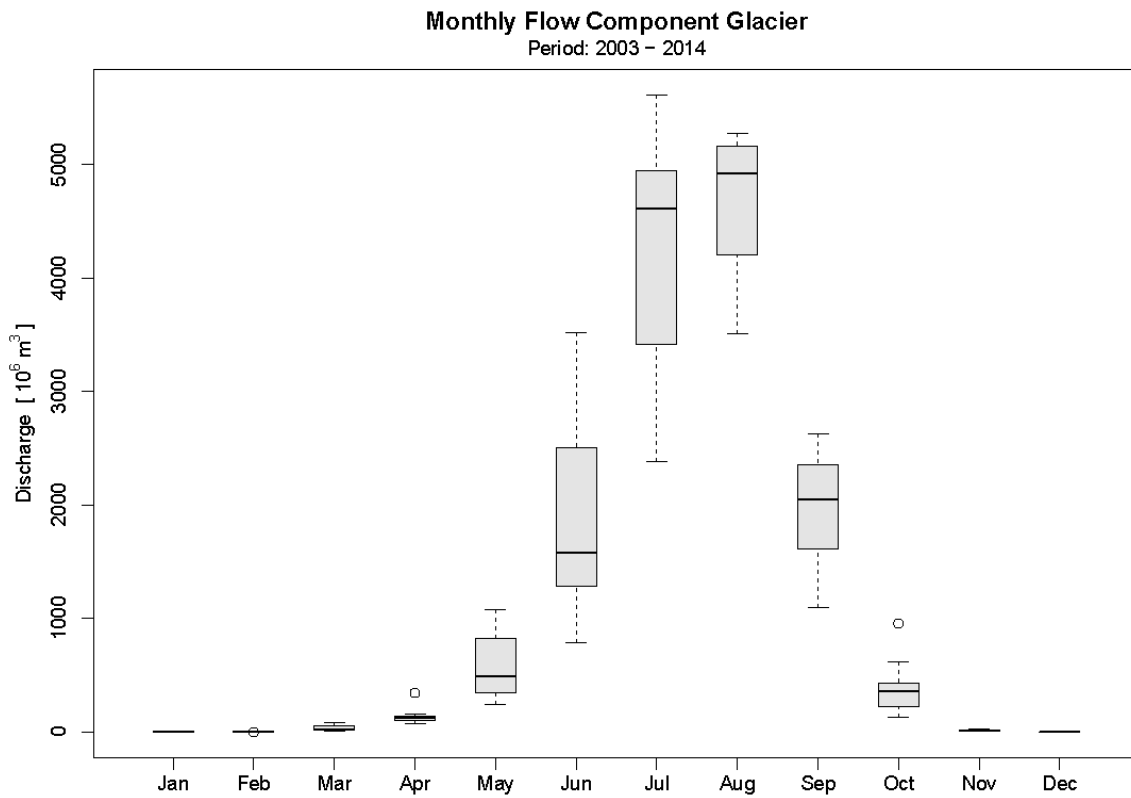
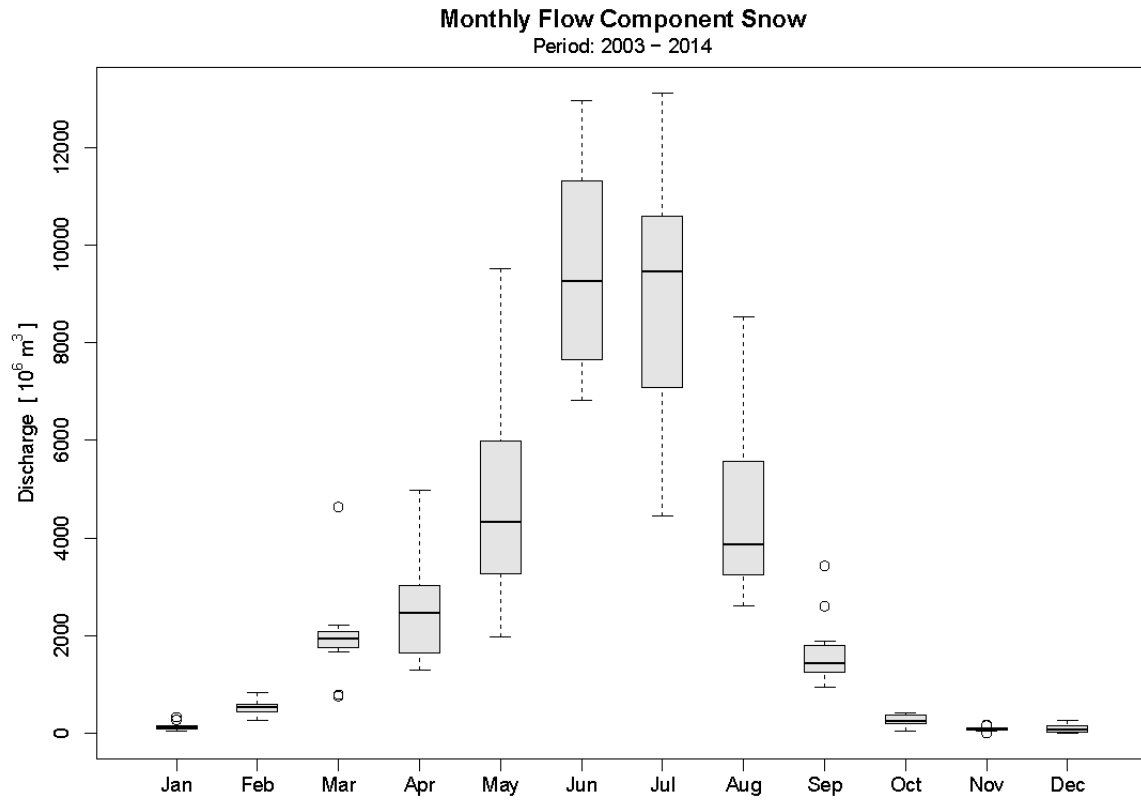


Figure 7: Monthly distribution of flow components in UIB

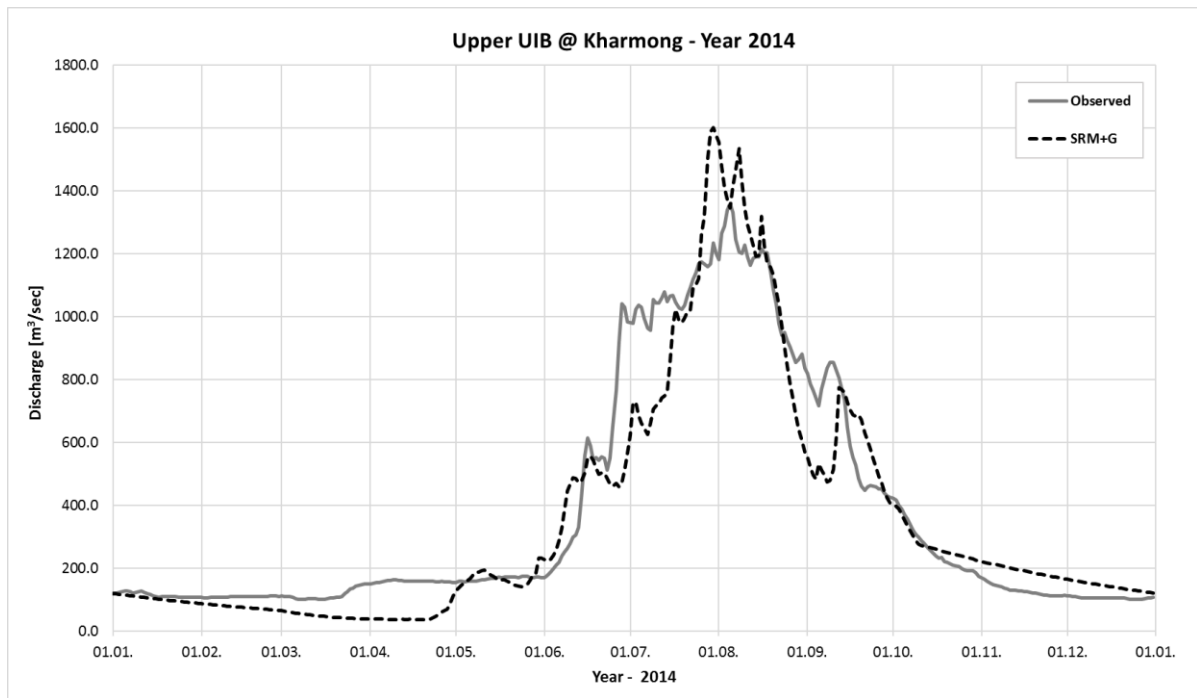


Figure 10: Results of validation of final Upper UIB flow forecast model (dashed line) compared to observed flows at Kharmong (solid line) for the year 2014.

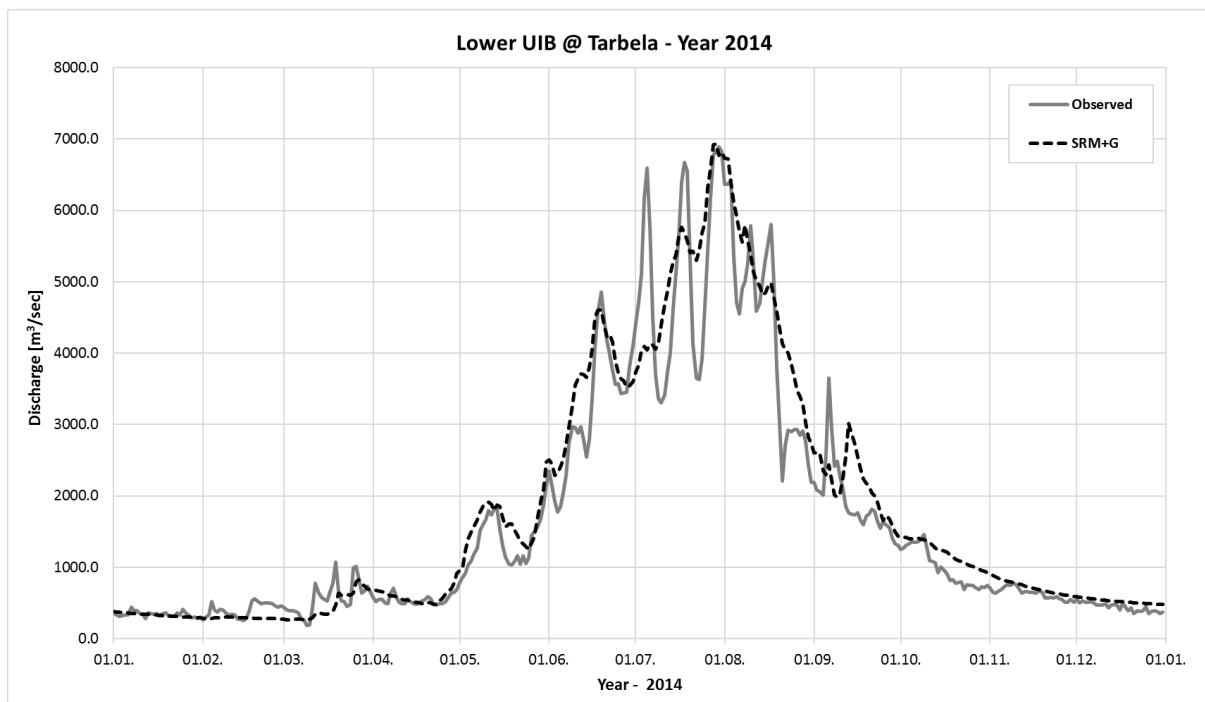


Figure 11: Results of validation of final Lower UIB flow forecast model (dashed line) compared to observed inflows at Tarbela (solid line) for the year 2014.

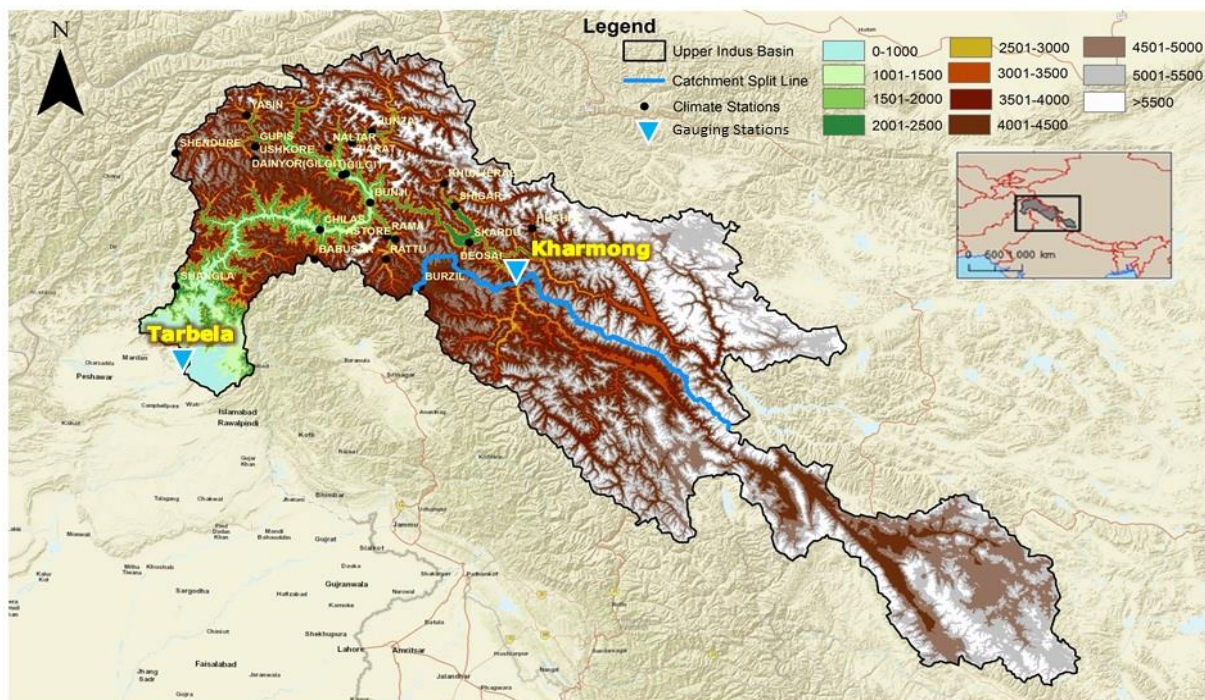
COMMENT: Overall the presentation of the results and discussion is weak. Results of some of the methods described are not presented, such as the results of the ensemble scenario approach. Also, the hydrographs presented at the two gauging stations are for 2008, and seem not to represent real forecasts, but rather simulations. (...) Forecasts have been developed for 2015 and 2016, but the results of these are not really presented, other than stating the estimated error.

RESPONSE: We have tried to improve the results and discussion by separating it into different parts including, the splitting of the UIB, incorporation of glacier component in the SRM model, simulation model verification and evaluation of forecasting skills.

We have also substituted Table 6 in the paper providing a yearly comparison of forecast results with all three models. Regarding SRM+G, figures for the years 2003 – 2014 are from hind-casts using the developed forecasting procedures, while 2015 and 2016 are genuine forecasts. Figures 10 and 11 (see previous response) indeed show the results of a model simulation aimed to validate the forecasting parameter sets and rules

COMMENT: The Figures included are not clear and would need to be improved. Figure 3 is to a large extent redundant, and with the division of the basin easily displayed on Figure 1. The figures that include the hydrographs are also not easy to read.

RESPONSE: We have combined the figure 1 and figure 3 of the old manuscript into one figure as shown in Figure 1 of the revised manuscript. This Figure 1 shows the UIB boundary with different elevation zones as well as the catchment split line.



Background Map Source: ESRI Open Street Map

Figure 1: Map of the Upper Indus Basin showing different elevations and splitting of UIB at the Kharhong gauging station into Upper and Lower UIB

COMMENT: Finally, the manuscript lacks clear conclusions or take-home messages, as well as an outlook to scientific challenges that have been identified.

RESPONSE: The authors think that they have given some conclusions, e.g. on improving the UIB SRM+G model (inclusion of glacier melt, splitting of the catchment), as well as an outlook on further fields of research, e.g. selection of scenarios respectively weighting of ensemble members according to climate signals. Nevertheless, we appreciate the comment as it is indicating, that this is not communicated clearly and we have now given a separate and extended chapter of 'Conclusions' in a revised version of manuscript.

Point by point response to the comments

Referee No. 2

Scenario approach for the seasonal forecast of Kharif flows from Upper Indus Basin

Muhammad Fraz Ismail and Wolfgang Bogacki

bogacki@hs-kobelnz.de

Dear Referee No. 2

Thank you for your specific comments and suggestion, we have tried to address your concerns on the points you have mentioned below. We appreciate your precise comments and detailed suggestion for the improvement of our paper.

COMMENT: Section 2.2 (Line 11-14). You should provide enough quantitative evidences to explain how to divide the Upper and Lower parts within UIB.

RESPONSE: In section 2.3 *Splitting the UIB into two sub-catchments*, we have tried to explain the reasons behind the splitting in more detail. Figure 6 as well as Table 4 also go in line with the physical reasoning we have explained that section.

Ideally the split between Upper and Lower UIB model should be done downstream of the Tibetan Plateau. Due to the lack of discharge data in this part of the UIB, Khariong as the closest gauging station where the daily discharge data is available was selected.

Table 4: Depletion of snow cover area for Upper and Lower UIB during March 2003

Elevation (m asl)	3500	4000	4500	5000	5500	>5500
1 st March						
Lower UIB	66%	82%	88%	87%	83%	94%
Upper UIB	58%	79%	58%	51%	58%	71%
1 st April						
Lower UIB	42%	71%	84%	84%	78%	92%
Upper UIB	24%	50%	48%	43%	51%	73%

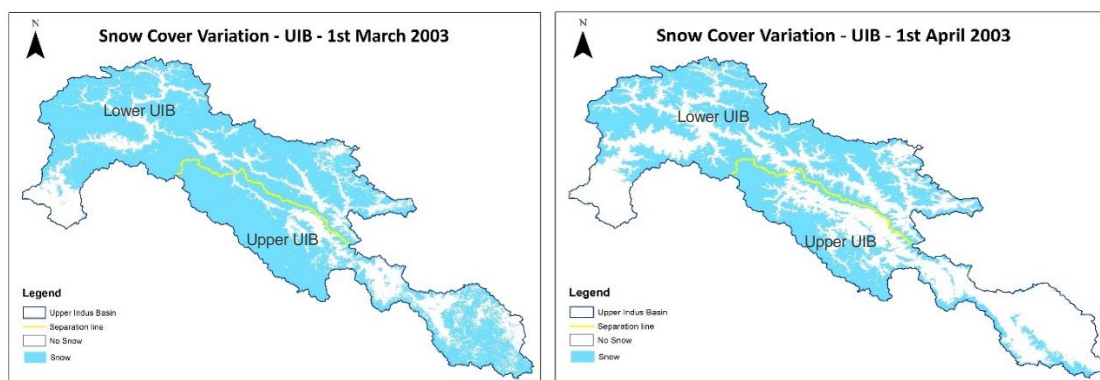


Figure 6: Snow cover variation in the month of March and April 2003 in UIB

COMMENT: Section 2.2 (Line 15-16). There are little details about the Kirpich travel-time equation that is used to determine the 3-day time lag between Kharhong and Tarbela. Please add the relevant information in the revised version.

RESPONSE: We have now explained in the revised manuscript under section 2.3 *Splitting the UIB into two sub-catchments* that how the Kirpich travel time equation was applied.

Kirpich equation (5) (Kirpich, 1940; USAD, 2010)

$$t = 0.00195 L^{0.77} S^{-0.385} \quad (5)$$

as in this empirical equation the time of concentration t [min] is only related to the length of the main channel L [m] and the slope of the longest hydraulic length S [-]. Given the altitudes of Kharhong and Darband (upstream Tarbela reservoir) gauging stations as 2,542 and 436 m a.s.l respectively and a channel length¹ of about 617 km, the approximated time-lag of 5000 min was finally rounded to 3 days.

COMMENT: Section 3 (Line 1-5). This part (forecast skill metrics) belongs to method description. Please move it to Section 2.

RESPONSE: You are right. In our revised paper, we have now made a section 2.7 *Verification methods* as you have pointed out.

COMMENT: The authors only focused on the forecast performance of median (50%) volume values. Actually, the extreme volumes, like “dry” (20%) and “wet” (80%) conditions, may be of greater importance for the downstream regions. It should add a few additional skill assessments in terms of predicting extreme conditions.

RESPONSE: We have now included Table 7 for the comparison of forecast skill with different available methods. We have made a “dry” (20%) and “wet” (80%) comparison for SRM+G for the Kharif flow volumes as shown in Figure 12 and for calculating the PSS and the RPSS. We have also prepared a Taylor diagram (see Figure 13) for the model performance.

Table 7: Comparison of forecast skills between IRSA, UBC, and SRM+G

Model	MAE	RMSE	MPE	MAPE	R	AC _u	PSS
	km ³	km ³	%	%	-	-	-
IRSA	6.5	8.0	5.8	10.9	0.107	0.085	-0.070
UBC	6.9	7.7	6.3	11.4	0.318	0.260	0.096
SRM+G	6.0	7.0	-2.0	9.5	0.223	0.168	-0.079

¹ digitised from Esri's World Imagery. Source: Esri, DigitalGlobe, GeoEye, i-cubed, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community

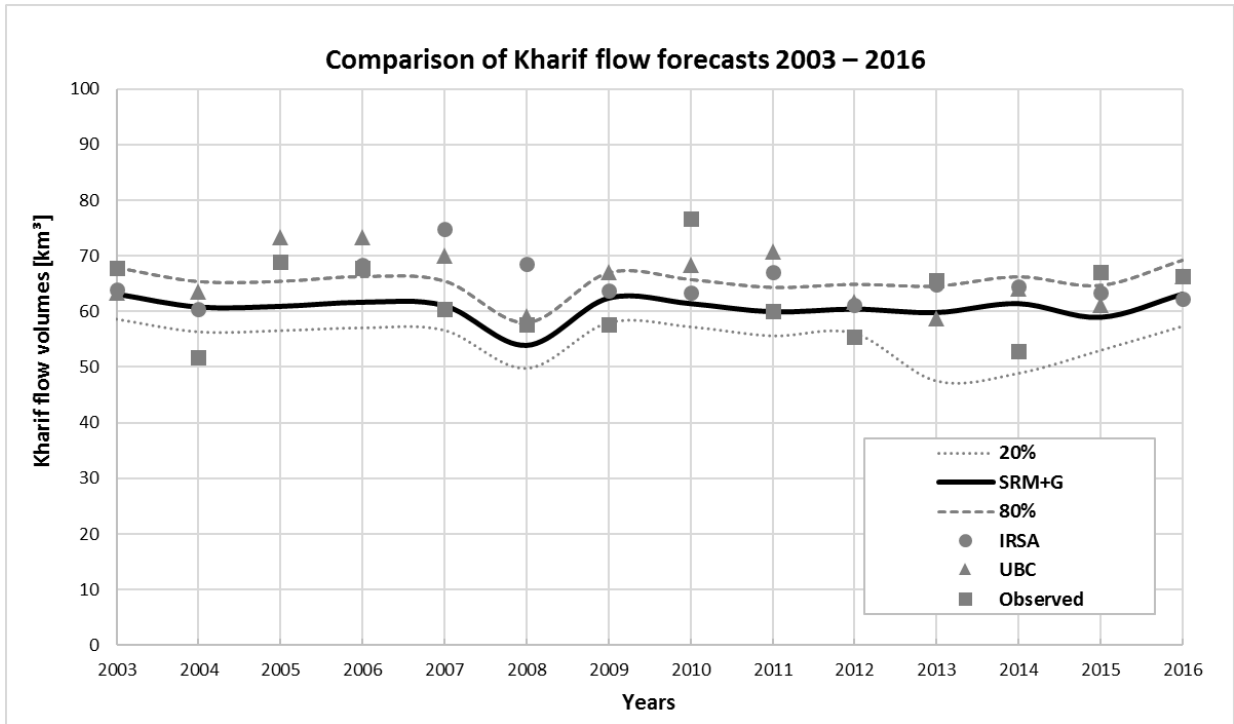


Figure 12: Comparison of Kharif flow forecast with 20% and 80% quantiles of SRM+G scenario ensembles years 2003 - 2016

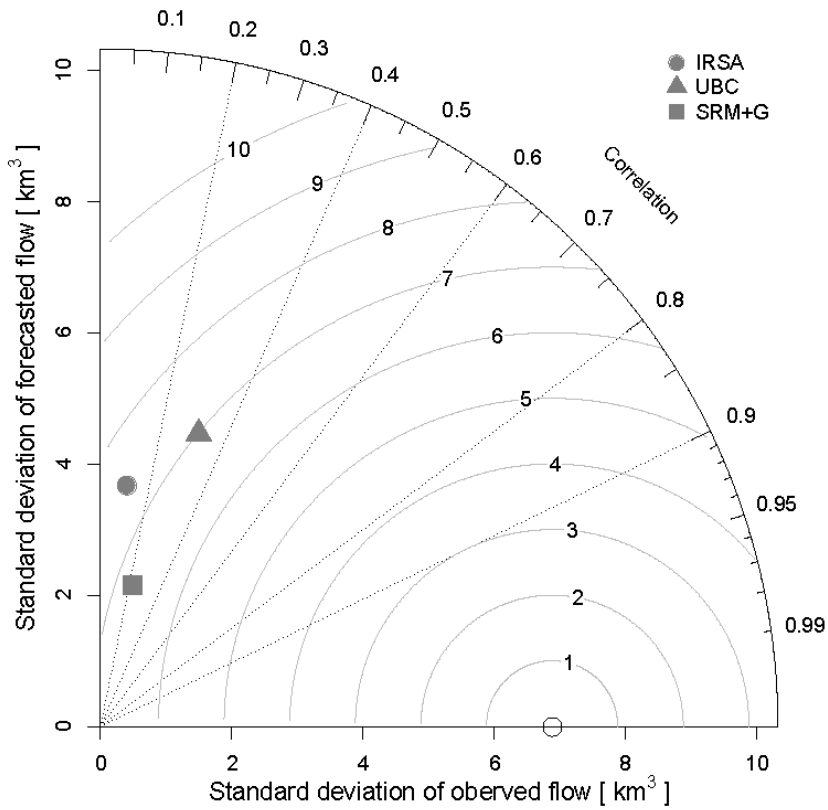


Figure 13: Taylor diagram of IRSA, UBC, and SRM+G model performance

COMMENT: This study assessed the forecast skill only by examining the volume difference of determined values. In the revised version, a few probabilistic quantitative metrics, like anomaly correlation (AC), Brier Score (BS), the false alarm rate (FAR), hit rate (HR), and Equitable Threat Score (ETS), should be employed to assess its skill in probabilistic forecasting.

RESPONSE: See response of the previous comment. Where we have given a table for the comparison of forecast skills.

COMMENT: Section 3. Please add 1-2 figures to illustrate the comparison of SRM+G model-based scenario approach with other forecast models.

RESPONSE: We have provided a comparison of all three operational forecast models as given in Table 6.

Table 6: Comparison of Kharif flow volumes [km³] 2003 - 2016

Year	Observed	IRSA	UBC	SRM+G
2003	67.8	64.0	63.5	63.1
2004	51.8	60.5	63.6	60.8
2005	68.9	69.0	73.3	60.9
2006	67.8	68.4	73.3	61.6
2007	60.5	74.9	70.1	61.0
2008	57.7	68.5	59.2	53.9
2009	57.6	63.7	67.2	62.4
2010	76.6	63.3	68.4	61.4
2011	60.0	67.2	70.8	59.9
2012	55.4	61.3	61.7	60.4
2013	65.6	64.9	58.8	59.8
2014	52.9	64.6	64.2	61.4
2015	67.2	63.3	61.3	58.9
2016	66.4	62.4	66.5	63.1

COMMENT: You are suggested to add a comparison between SRM+G and SRM, to highlight the superiority of SRM+G in terms of incorporating glacier component.

RESPONSE: We have added a comparison of SRM with and without glacier component in order to see what is the possible impact of glaciers melting component in the late summer in the revised manuscript (See Figure 8 and 9).

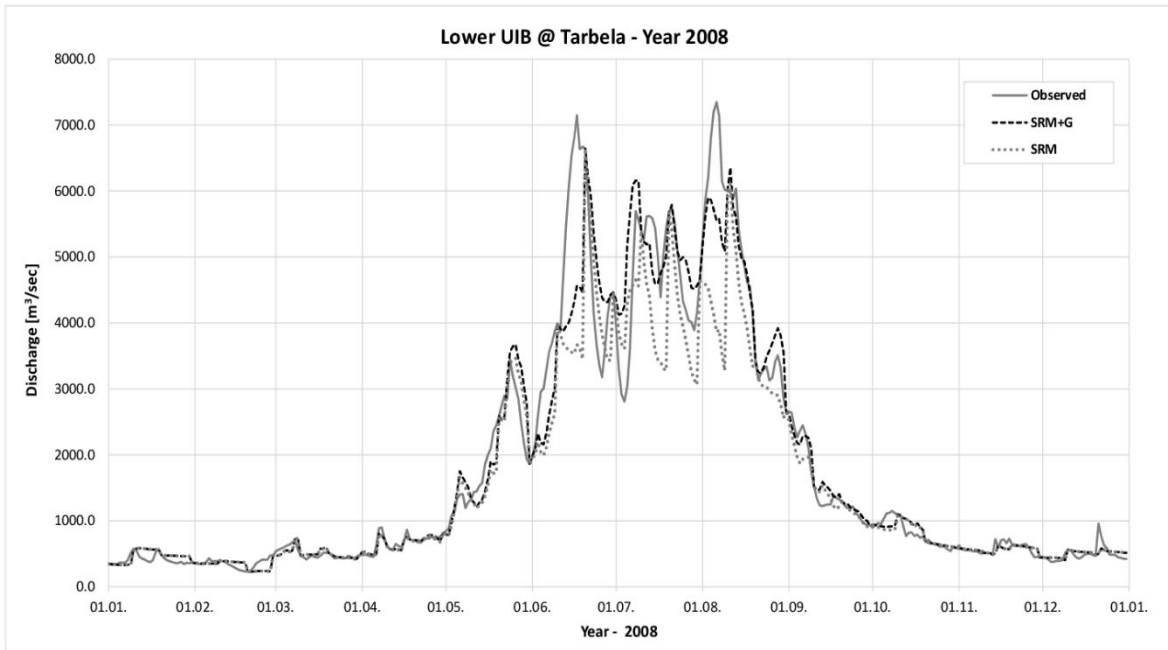


Figure 8: Comparison of SRM+G (with glaciers) and SRM (without glaciers) for Lower UIB – 2008

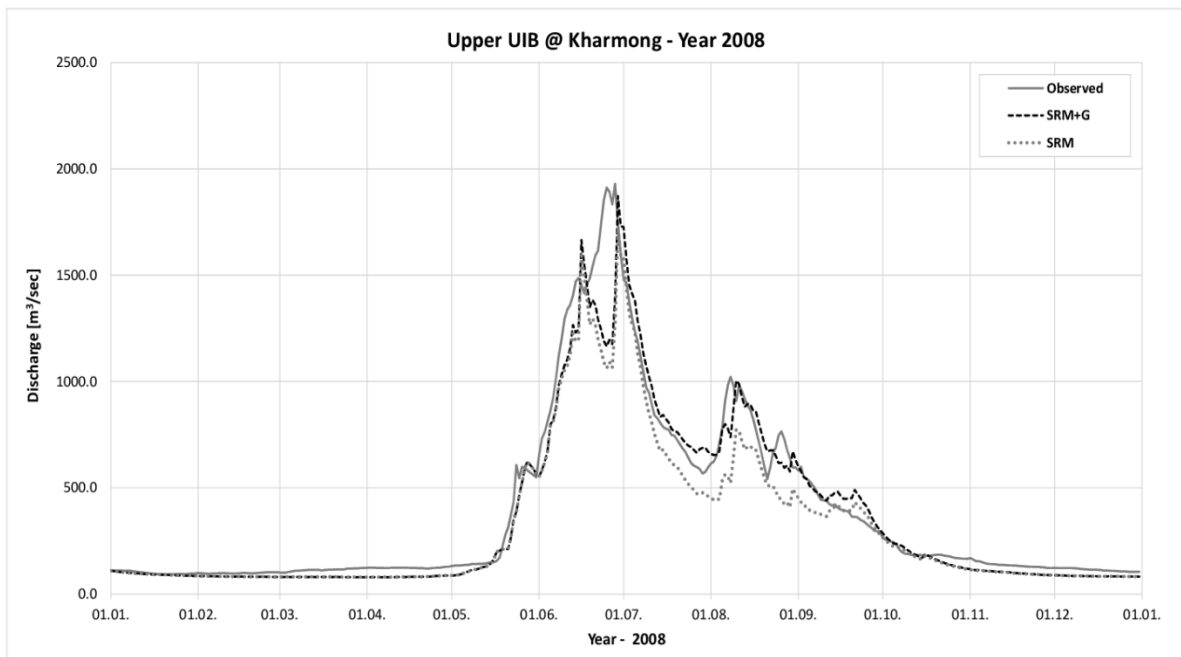


Figure 9: Comparison of SRM+G (with glaciers) and SRM (without glaciers) for Upper UIB - 2008

The impact of glacier in the Lower UIB is much more prominent than in the Upper UIB. The reason is that in the Lower UIB the glaciated area is about 10.5% of the catchment area while in the Upper UIB the glaciated area is only 1.7%.

COMMENT: Adding a comparison of SRM+G estimates between with and without consideration of divisions (Upper and Lower parts) should be inserted and discussed in the revision.

RESPONSE: A detailed discussion has been added in section 3.1 *Splitting of the UIB catchment*. Though the whole UIB showed an acceptable agreement between simulated and observed flows in terms of R^2 and D_v , initial hindcast results proved to be not satisfactory especially for the Early Kharif (1st April – 10th June) season, having a mean absolute percentage error MAPE of 25.9%. The MODIS snow cover analysis showed that in the Tibetan plateau already in March the snow cover is fading away rapidly while on the other hand, in the north-western part of the catchment the same elevation zone is still widely covered with snow (Figure 6). With splitting the MAPE could be reduced to 15.8%.

COMMENT: Section 2.3 (Line 30). If I understand correctly, R indicates the daily runoff depth, not precipitation depth.

RESPONSE: Yes, you are right, precipitation depth relates to the original SRM notation. We have corrected it in the revised version see section 2.2 *Model structure* line 15.

COMMENT: Section 2.5 (Line 33). “TRMM 3B34 product” should be “TRMM 3B34 product”. Please correct it.

RESPONSE: Corrected in revised manuscript.

List of relevant changes

LIST OF CHANGES

Following is the list of all relevant changes made in the revised manuscript:

Figures:

1. Combined Figure 1 & 3 => Now revised manuscript Figure 1
2. Updated Figure 6 and 7 => Now Figure 10 and 11 in revised manuscript
3. Added new Figures 6, 7, 8, 9, 12, 13, 14 and 15 in revised manuscript

Tables:

1. Added new Tables 3, 5, 6 and 7 in revised manuscript

Chapters:

1. Corrected the mistake in units of R in chapter 2.2 Model Structure
2. Incorporated the Kirpich equation and related description in chapter 2.3 Splitting the UIB into two sub-catchments
3. Incorporated the description about the degree-day factors linear increase during the snowmelt season in chapter 2.5 Model parameters
4. Incorporated the discussion about ESP in chapter 2.6
5. Added a new chapter 2.7 verification methods and discussed forecast verification metrics like Anomaly Correlation ACu, Peirce skill score (PSS), Ranked Probability Score (RPS) and Ranked Probability Skill Score (RPSS).
6. Updated the chapter 3 Results and Discussion by incorporation sub-sections like
 - a. Splitting of the UIB,
 - b. Glacier melt component,
 - c. Simulation model verification including Table 5 with R^2 and D_v
 - d. Evaluation of forecasting skills including Table 7 with verification metrics
7. Updated the chapter 4 Conclusion
8. Updated the relevant references

Marked-up manuscript

Scenario approach for the seasonal forecast of Kharif flows from Upper Indus Basin

Muhammad Fraz Ismail and Wolfgang Bogacki

Department of Architectural and Civil Engineering, Koblenz University of Applied Sciences, Germany.

Correspondence to: W. Bogacki (bogacki@hs-koblenz.de)

Abstract. Snow and glacial melt runoff are the major sources of water contribution from the high mountainous terrain of Indus river upstream of the Tarbela reservoir. A reliable forecast of seasonal water availability for the Kharif cropping season (April – September) can pave the way towards the better water management and subsequently boost the agro-economy of Pakistan. The use of degree-day models in conjunction with the satellite based remote sensing data for the forecasting of seasonal snow and ice melt runoff has proved to be a suitable approach for the data scarce regions. In the present research, Snowmelt Runoff Model (SRM) has not only been enhanced by incorporating the “glacier (G)” component but also applied for the forecast of seasonal water availability from the Upper Indus Basin (UIB). Excel based SRM+G takes into account of separate degree-day factors for snow and glacier melt processes. All year simulation runs with SRM+G for the period 2003 – 2014 result in an average flow component distribution of 53%, 21%, and 26% for snow, glacier and rain respectively. The UIB has been divided into Upper and Lower part because of the different climatic conditions in the Tibetan plateau. The scenario approach for seasonal forecasting, which like Ensemble Streamflow Prediction uses historic meteorology as model forcings, has proven to be adequate for long-term water availability forecasts. The accuracy of the forecast with a MAPE of 9.5% could be slightly improved compared to two existing operation forecasts for the UIB and the bias could be reduced to -2.0%. However, the association between forecasts and observations as well as the skill in predicting extreme conditions is rather weak for all three models, which motivates further research on the selection of a subset of ensemble members according to forecasted seasonal anomalies.

Commented [i1]: New text

1 Introduction

Mountains are the water towers of the world. They are the biggest resource of freshwater to half of the world's population fulfilling their needs for irrigation, industry, domestic and hydropower (Viviroli et al., 2007). The Indus River on which Pakistan's socio-economic development depends can be termed as the bread basket of Pakistan (Clarke, 2015). Due to agrarian economy, Pakistan's agriculture share in water usage is about 97%, which is well above the global average of about 70% (Akram, 2009). In Pakistan, Indus River System Authority (IRSA) decides the provincial water shares according to the Water Apportionment Accord (WAA) of 1991 and provincial irrigation departments subsequently determine the seasonal water allocation to the different canal command areas depending upon the water availability forecast carried out at the end of March for the forthcoming Kharif cropping season (April-September). A reliable seasonal forecast of the water availability from snow and glacial melt is therefore of utmost importance for the agricultural production and efficient water management.

But on the other hand snowmelt runoff modelling in mountainous regions faces the challenge of data scarcity as well as the uncertainty in parameter calibration (Pellicciotti et al., 2012). The need of the hour is to not only develop such a hydrological model which has the capability to cater both snow and glacial melt component but also a reliable forecast technique which could help the water managers and policy makers for enhancing the water resources management in future. The present paper focuses on the implementation of a snow and glacial melt runoff model on the Upper Indus Basin (UIB) as more than 70% of the inflows to Tarbela reservoir originate from these two flow components (Tahir et al., 2011). The ExcelSRM version (Bogacki and Hashmi, 2013) of WinSRM (Martinec et al., 2011) has been enhanced by taking into account the glacial melt component based on the methodology proposed by Schaper et al. (1999) and by introducing the possibility to split a watershed into sub-catchments which is not implemented in original WinSRM. For long-term streamflow forecasting, Ensemble Streamflow Prediction (ESP) developed at the U.S. National Weather Service (Day, 1985) is widely used to generate probabilistic flow forecasts. As already successfully applied in the Upper Jhelum basin (Bogacki and Ismail, 2016), a scenario approach is used for seasonal flow forecasting in the UIB, which has much similarity to ESP. It also uses historical meteorology as model forcings, however like the other operational forecast models for UIB, it is mainly focussed on a deterministic forecast of total Kharif inflow to the Tarbela reservoir.

2 Materials and Methods

2.1 Study area

The upper catchments of the Indus River basin (Figure 1) primarily feed the Tarbela reservoir, which is the larger of the only two major reservoirs in Pakistan. The Upper Indus Basin has an area of about 173,345 km² of which approx. 11.5% is covered by perennial glacial ice (Tahir et al., 2011). At the end of most winters nearly the entire UIB above 2,200m a.s.l is covered with snow, resulting in that more than 60% of annual flow in the Upper Indus River is contributed by the snowmelt (Bookhagen and Burbank, 2010). The distribution of monthly inflows to the Tarbela reservoir (see Figure 2) shows that these flows tend to rise progressively as melting temperatures advance into areas of higher snowpack at the higher elevations. Indus River starts rising gradually in March reaching its maximum in July, while peak flood events usually occur during the monsoon season in July – September. When

by the end of July the flows reduce due to diminished snow cover in the lower catchment, the high altitude glacierised basins become important contributors to the flows due to first melting of their seasonal snow cover and when the snow has vanished then melting of the glacier ice. According to Tahir et al. (2011) glacial melt dominates the flows of the largest tributaries of Indus River, i.e. Chitral, Gilgit, Hunza, Braldu and Shyok rivers.

5 2.2 Model Structure

The Snowmelt Runoff Model (SRM) (Martinec, 1975) is a semi-distributed, lumped temperature-index model which is specifically designed to simulate the runoff in snow-dominated catchments that has been successfully applied in more than hundred snow-driven basins around the globe (Martinec et al., 2011). Input variables of SRM are daily values of air temperature, precipitation, and snow covered area. The catchment is usually subdivided into elevation zones of about 500 m each and the input variables are distributed accordingly. The total daily amount of water produced from snowmelt and rainfall in the catchment is superimposed on the calculated recession flow according to the equation (1):

$$Q_{n+1} = \sum_{i=1}^m \left\{ [M_{n,i} + R_{n,i}] \cdot \frac{10000}{86400} A_i \right\} \cdot (1 - k_{n+1}) + Q_n k_{n+1} \quad (1)$$

Where, Q is the average daily discharge [m^3s^{-1}], M and R are the daily runoff depth originating from snowmelt and rainfall [cm d^{-1}], A is the total area of the elevation zone [km^2], k is the recession coefficient [-], n is the index of the simulation day and i, m are the indices and total number of elevation zones respectively. Daily runoff from snowmelt and rainfall is calculated by equations (2) and (3):

$$M_{n,i} = c_{Sn,i} \cdot a_{Sn,i} \cdot T_{n,i} \cdot S_{n,i} \quad (2)$$

$$R_{n,i} = c_{Rn,i} \cdot P_{n,i} \quad (3)$$

Where, c_S and c_R are the runoff coefficients [-] for snowmelt and rain, a_S is the degree-day factor for snow [$\text{cm } ^\circ\text{C}^{-1} \text{d}^{-1}$], T the number of degree-days [$^\circ\text{C d}$] for each elevation zone, S the ratio of the snow covered area to the total area [-] and P the daily precipitation [cm d^{-1}].

Schaper et al., (1999) introduced an enhancement in the original SRM approach by incorporating the separate glacial melt component in the model. In addition to the variables used by SRM it also considers the area covered by exposed, i.e. not snow covered, glaciers. An additional melt component is added to equation (1) that takes into account the specific degree-day factors for glaciers according to equation (4):

$$G_{n,i} = c_{Gn,i} \cdot a_{Gn,i} \cdot T_{n,i} \cdot S_{Gn,i} \quad (4)$$

Where, G is the daily melt [cm d^{-1}] from exposed glaciers in each elevation zone, c_G is the runoff coefficient [-] and a_G is the degree-day factor [$\text{cm } ^\circ\text{C}^{-1} \text{d}^{-1}$] for glaciers, and S_G is the ratio of the exposed glacier area to the total area [-].

Commented [i2]: updated

This model was tested in several basins and found with high accuracy even in basins with 67% glacier areas on three alpine basins Rhine-Felsberg, Rhône-Sion and Ticino-Bellinzona in Switzerland (Schaper and Seidel, 2000). Apart from the improvement of the runoff modelling, the independent computation of glacier melt is an important step towards evaluations of glacier behaviour with regard to climate change.

5 The glacier melt component according to equation (4) was incorporated into the existing ExcelSRM further referred to as SRM+G. This extension requires the glacier exposed area as an additional daily input variable and respective model parameters as given in equation (4).

An additional enhancement is the possibility to split the watershed into different sub-catchments. This feature is realised by adding the pre-calculated outflow of a sub-catchment obtained by a separate simulation to the discharge
10 of the downstream sub-catchment. The travel time can be considered by applying a time-lag to the daily discharge time-series.

2.3 Splitting the UIB into two sub-catchments

In the Karakorum – Western Himalayas region snow accumulates during winter and reaches its maximum extension in February/March. Higher altitudes typically have a 90% – 100% snow cover that stays more or less constant until
15 melting starts in spring. There is however a characteristic bias between the north-western part of the UIB where at altitudes above 4,000 m a.s.l. the snow covered area usually starts gradually decreasing in March, while in the south-eastern part namely the Tibetan plateau at the same altitudes the snow cover is fading away very soon. This bias leads to an inevitable under-estimation in forecasting the snowmelt dominated Early Kharif flows (see Chap. 3.1), which motivates the splitting of the UIB into two sub-catchments.

20 Ideally, the catchment should be split right downstream of the Tibetan plateau. However, as the first gauging station where daily flow data was available is Kharhong gauging station, when the Upper Indus River has entered into Pakistan, this location was chosen to split the UIB in an upstream and a downstream sub-catchment, namely the Lower and Upper UIB (Figure 1). The hypsometric characteristics including the number of elevation zones and their corresponding areas of both the sub-catchments are shown in Figure 3.

25 According to the two sub-catchments, two separate SRM+G models were created. For each simulation, first the Upper UIB model is run in order to simulate flows at Kharhong. These flows are then superimposed to the flows calculated by the Lower UIB model using a time-lag between Kharhong and Tarbela that was estimated by the Kirpich equation (5) (Kirpich, 1940; USAD, 2010)

$$30 \quad t = 0.00195 L^{0.77} S^{-0.385} \quad (5)$$

as in this empirical equation the time of concentration t [min] is only related to the length of the main channel L [m] and the slope of the longest hydraulic length S [-]. Given the altitudes of Kharhong and Darband (upstream Tarbela reservoir) gauging stations as 2,542 and 436 m a.s.l respectively and a channel length¹ of about 617 km,
35 the approximated time-lag of 5000 min was finally rounded to 3 days.

Commented [i3]: New text and Equation

¹ digitised from Esri's World Imagery. Source: Esri, DigitalGlobe, GeoEye, i-cubed, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community

2.4 Data sources

There are a number of high elevation climate stations in the Pakistani part of the Upper Indus Basin operated by WAPDA's² Glacier Monitoring and Research Centre (GMRC) and Pakistan Meteorological Department (PMD). However, they are concentrated on the western part of the UIB and data is not available online. In order to have most recent data for operational flow forecasting, the World Meteorological Organization (WMO) climate station at Srinagar airport located at an altitude of 1,587 m a.s.l. was chosen as temperature base station, which already had proven to give representative temperatures for that region in the SRM model of the Upper Jhelum catchment (Bogacki and Ismail, 2016) and a full set of climatic data can be obtained online from the GSOD³ data-base with a time-lag of about 2 days only. Based on the daily air temperature data, degree-days in each elevation zone were calculated using a constant temperature lapse-rate of $-6^{\circ}\text{C km}^{-1}$.

The MODIS/Terra Snow Cover Daily L3 Global 500 m Grid (MOD10A1) product⁴ has been used to determine the daily snow covered area in the elevation zones. The compatibility of using MODIS data in conjunction with SRM in the Himalayas and its surroundings has already investigated by Immerzeel et al. (2009, 2010). As the MODIS sensor cannot detect snow below clouds, a cloud elimination algorithm is applied using temporal interpolation between two cloud-free days for each pixel. Afterwards the daily percentage of snow cover area in each elevation zone is calculated and smoothed by moving average.

At the beginning of the melting season, glaciers are usually completely covered by fresh snow. As the melting season progresses the snow cover will fade away and glacier exposed area will increase. The actual glacier extent was derived from two data sources. As a major source on global glacier distribution the Global Land Ice Measurements from Space (GLIMS) data archive was used (Raup et al., 2007). This data was complemented by interpretation of Landsat 8 scenes (30 m spatial resolution) from late summer to early fall 2013, in order to identify the maximum of the glacier exposed area. The merged data was mapped on the 500 m MODIS grid. On a daily basis, the glacier exposed area is determined by all pixels that are classified as glacier but not identified as snow by the MODIS sensor.

A spatial interpolation of in-situ (station) precipitation data in mountainous regions is particularly difficult and often biased towards lower values (Archer and Fowler, 2004) as the rain gauge network is usually sparse and mainly located at the valley floors while maximum precipitation occurs on mountains slopes and increases with altitude in general. A promising alternative to station data are gridded, remote sensing based precipitation products. However, regional and temporal pattern as well as multiannual means of these products differ significantly in the Himalayas (Palazzi et al., 2013). In particular, the widely used TRMM data-set is known to underestimate the precipitation in high altitudes as found in the UIB (Forsythe et al., 2011) or the Andes (Ward et al., 2011).

Based on own precipitation product comparisons for the Upper Chenab catchment, the gridded RFE 2.0 Central Asia⁵ daily rainfall product (Xie et al., 2002) is used in the present model. According to SRM's elevation band

² Pakistan Water and Power Development Authority

³ Global Summary Of the Day. Download at: <ftp://ftp.ncdc.noaa.gov/pub/data/g sod/>

⁴ Hall et al. (2006), updated daily. MODIS/Terra Snow Cover Daily L3 Global 500m Grid V005, [Feb. 2000 – Sep. 2016, tiles h23v05 & h24v05]. NSIDC Boulder, Colorado USA. Download at: <ftp://n5eil01u.ecs.nsidc.org/SAN/MOST/MOD10A1.005>

⁵ RainFall Estimates version 2.0 created by the NOAA Climate Prediction Center's FEWS-NET group sponsored by USAID. Download at: <ftp://ftp.cpc.ncep.noaa.gov/fews/afghan>

approach, the gridded data, having a spatial resolution of 0.1° latitude/longitude, is mapped to the respective elevation zones. For the period 2003 – 2015 the product yields a mean annual precipitation of 854 and 482 mm/a for the Lower and the Upper UIB respectively, that reflects the significantly lower annual precipitation on the Tibetan plateau compared to the western Himalayas (e.g. Bookhagen and Burbank, 2010; Ménégoz et al., 2013).
5 The RFE basin-wide annual mean of 701 mm/a lies well in the range of 675 ± 100 mm/a derived for the whole UIB by Reggiani and Rientjes (2015).

2.5 Model parameters

The most important parameter of a temperature-index model which is controlling daily snow and glacial melt is the degree-day factor [$\text{cm } ^\circ\text{C}^{-1} \text{ d}^{-1}$], which transforms the index variable degree-day [$^\circ\text{C d}$] into actual melt [cm d^{-1}].

10 In case of glaciers a constant degree-day factor of $0.70 [\text{cm } ^\circ\text{C}^{-1} \text{ d}^{-1}]$ as proposed by Schaper et al. (2000) was chosen, which also corresponds to degree-day factors reported from glaciers in the Himalayas at a comparable latitude (Hock, 2003), the approach for degree-day factors for snow is more elaborated. In a first step, optimal degree-day factors were obtained for each elevation zone and year by diagnostic calibration, i.e. by achieving the best possible fit between simulated and observed hydrographs for each year. From this calibration exercise it
15 appears, that degree-day factors are increasing by the time after melting has started in a particular elevation zone (Figure 4 and 5). As a generalised rule is needed in the forecasting procedure, zone-wise degree-day factor functions as suggested by Ismail et al. (2015) were developed by linear regression between the calibrated degree-day factors and time. The increase of the degree-day factors with the passage of time is because the snow absorbs energy due to its physical condition, in terms of increasing temperatures and solar radiations intensities. This process of energy
20 storage plays a pivotal role in the ripening of the snowpack, which melts rapidly as the snow melting season progresses. The extent to which degree-day factors increase is related to the calibration procedure because it was observed during the model calibration that in a certain elevation zone when the degree-day factors attain the value e.g. $(0.80 \text{ cm } ^\circ\text{C}^{-1} \text{ d}^{-1})$, the snow cover area in that very elevation zone has almost completely faded away so there is no need to further increase the values of degree-day factors. The limit to what extent the degree-day factors
25 increase at a certain spatio-temporal region depends upon various physiographic and climatic parameters and a research is on-going for evaluating the trend of degree-day factors in response to the aforementioned parameters. The start of snowmelt and correspondingly application of the developed degree-day factors generalised rule, is correlated with a certain threshold temperature for each elevation zone (see Table 1 and 2). The other model parameters required by SRM like temperature lapse-rate, recession coefficient, runoff coefficient for snow, lag-
30 time, etc., were applied basin-wide and kept constant for all years (see Table 3). The values of these parameters were determined according to the methods described by Martinec et al., (2011) and slightly adjusted to achieve a good fit over the whole calibration period. It has to be noted, that these parameter values will differ for other catchments.

2.6 Scenario approach for forecasting

35 In the forecasting period which starts from the 1st of April, the four model variables temperature, precipitation, snow covered area and glacier exposed area have to be predicted for the forthcoming 6 months of the Kharif cropping season (April – September). As the level of skill of seasonal climate forecasts for the Hindukush –

Commented [i4]: New text with physical reasoning of the degree-day factors increase. Parameters Table is also included.

Karakoram – Western Himalaya region for such a lead time is still not sufficient, a scenario approach already successfully applied in the Upper Jhelum catchment (Bogacki and Ismail, 2016) is used.

This scenario approach has a lot in common with traditional Ensemble Streamflow Prediction ESP developed at the U.S. National Weather Service as a method for generating long-term probabilistic streamflow outlooks (Day, 1985). Based on the assumption that past meteorology is representative of possible future events, ESP uses historical temperature and precipitation time series as forcings for the hydrological model to produce an ensemble of streamflow traces. A probabilistic forecast is created by statistical analysis of the multiple streamflow scenarios produced (Franz et al. 2008). Initial basin conditions are usually estimated by forcing the hydrological model with observed meteorology in a “warm-up” phase up to the time of forecast (Wood and Lettenmaier, 2008).

The seasonal scenario approach also uses historical temperature and precipitation as forcings for the SRM+G model. In contrast to ESP however, this approach is, like the other operational forecast models for UIB, primarily focussed on a deterministic forecast of total Kharif flow volume. Besides the “most likely” (median) flow, SRM+G forecasts only give an indication of the bandwidth of expected flows by the dry (20%) and wet (80%) quantiles as limits of the “likely” range.

Another notable difference are the initial basin conditions. SRM and SRM+G do not use any initial conditions, like soil moisture state of snow-water equivalent in other hydrological models. Instead however, the snow-cover area and the glacier exposed area are input variables to the model. For reasons of simplicity, the glacier exposed area is treated like the meteorological variables, i.e. the historical time-series are used. The depletion of the snow-covered area during the forecast period, which is the decisive factor for each forecast, is however predicted by so-called “modified depletion curves”. These modified depletion curves are derived from the conventional depletion curves of each elevation zone by replacing the time scale with the cumulative daily snow-melt depth (Martinec et al., 2011). The decline of the modified depletion curves depends on the initial accumulation of snow and represents the actual snow-water equivalent. When initial snow depth is low the modified depletion curve declines faster than in years when a lot of snow has accumulated. In the end of March, when the seasonal forecast is carried out, an elevation zone showing already some decline in snow covered area, and hence having also some cumulated degree-days, is chosen as “key zone”. Comparing the relation of decline in snow covered area versus cumulated degree-days with a statistical analysis of the modified depletion curves of previous years, the actual amount of snow is estimated and the future depletion anticipated accordingly, while assuming similar snow conditions for all elevation zones.

The major difference to other hydrological models as used in the ESP is the positive effect that usually the uncertainty in the initial conditions is progressively superseded by the actual meteorological conditions. In SRM+G however, if an erroneous depletion estimate is in effect then it will persist during the whole forecast period. As all ensemble traces are based on the chosen depletion curves, the initial estimate is crucially influencing each trace of the ensemble in the same direction.

Commented [i5]: Discussion about ESP has been included.

Commented [i6]: Discussion about ESP has been included.

2.7 Verification methods

Commented [17]: New chapter has been added about the verification methods.

Model verification comprises the simulation model as well as the forecasting model. The accuracy of the simulation model was evaluated by the two standard criteria used in SRM (Martinec et al., 2011), namely the relative volume difference equation (6)

$$D_v = \frac{V - V^*}{V} \times 100 \text{ [%]} \quad (6)$$

and the coefficient of determination R^2 equation (7)

$$R^2 = 1 - \frac{\sum_{i=1}^n (Q_i - Q_i^*)^2}{\sum_{i=1}^n (Q_i - \bar{Q})^2} \quad (7)$$

where V and V^* are the observed and the simulated annual flow volumes, Q_i and Q_i^* are the observed and the simulated daily discharge values, and \bar{Q} is the average observed daily discharge.

10 The skill of the forecasting model was assessed in comparison with IRSA's forecasts that are based on a statistical model and with forecasts from the UBC⁶ watershed model (Quick and Pipes, 1977) that is used by WAPDA's Glacier Monitoring Research Centre. The set of verification metrics was chosen taking into account that the existing operational forecasts for Kharif flows are traditionally issued in form of deterministic forecasts, thus only the 'most likely' values forecasted by these models are available.

15 The accuracy of a forecast is a measure of the error between predicted and observed values. The root mean squared error RMSE equation (8), mean percentage error MPE equation (9), and mean absolute percentage error MAPE equation (10)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f_i - o_i)^2} \quad (8)$$

$$MPE = \frac{1}{n} \sum_{i=1}^n \frac{(f_i - o_i)}{o_i} \quad (9)$$

$$20 \quad MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|f_i - o_i|}{o_i} \quad (10)$$

were used as deterministic metrics to assess the accuracy of the predicted mean Kharif flow volumes. In the above equations, f_i is the forecasted and o_i the observed flow volume and n the total number of considered forecasts. Both, RMSE and MAPE measure the average magnitude of the forecast errors, where RMSE penalises larger errors more than MAPE. The mean percentage error measures the deviation between average forecasted and average observed

25 flows, i.e. a positive MPE indicates over-forecasting and a negative under-forecasting respectively.

As a commonly used deterministic measure of association, the correlation coefficient, equation (11)

$$R = \frac{\sum_{i=1}^n (f_i - \bar{f})(o_i - \bar{o})}{\sqrt{\sum_{i=1}^n (f_i - \bar{f})^2} \sqrt{\sum_{i=1}^n (o_i - \bar{o})^2}} \quad (11)$$

⁶ University of British Columbia Watershed Model

was applied to assess the correspondence between forecasted and observed values. In addition, the uncentered anomaly correlation AC_u (Wilks, 2006) equation (12)

$$AC_u = \frac{\sum_{i=1}^n (f_i - \bar{c})(o_i - \bar{c})}{\sqrt{\sum_{i=1}^n (f_i - \bar{c})^2} \sqrt{\sum_{i=1}^n (o_i - \bar{c})^2}} \quad (12)$$

where \bar{c} is the climatological average value, was used as another measure of association. The anomaly correlation is designed to measure similarities in the patterns of anomalies from the climatological average between forecasted and observed values. An $AC \geq 0.6$ is usually regarded as an indication of some forecasting skill (Wilks, 2006). In the present context, the climatology average \bar{c} is equivalent to the average observed flows \bar{o} .

The ability of a non-probabilistic forecast to predict extreme conditions is usually assessed by defining discrete categories like below normal, normal, and above normal. The Heidke and Peirce skill scores for multi-categorical forecasts measure the fraction of correct forecasts in each category in relation to those forecasts which would be correct due purely to random chance. The Peirce skill score, equation (13)

$$PSS = \frac{\sum_{j=1}^m p(f_j, o_j) - \sum_{j=1}^m p(f_j) p(o_j)}{1 - \sum_{j=1}^m p(o_j)^2} \quad (13)$$

is unbiased in the sense that it assigns a marginal distribution to the reference random forecast which is equal to the (sample) climatology (Wilks, 2006). In the above equation, m is the number of categories, $p(f_j, o_j)$ the joint distribution of forecasts and observations, and $p(f_j)$ and $p(o_j)$ are the respective marginal distributions.

As the existing operational forecasts are primarily designed as point estimates of the mean flow volume and hence a comparison of probabilistic metrics between these models is not possible, only a basic probabilistic evaluation of the SRM+G scenario ensembles was carried out. The ranked probability score RPS, which is essentially an extension of the Brier score to the many-event situation (Wilks, 2006), was used as it reflects the overall performance of a multi-category probabilistic forecast (Franz et al., 2003). In order to calculate the RPS, first the quantiles of m categories have to be determined based on given non-exceedance probabilities of the observed values. Then, for each forecast the ensemble members as well as the observed flow are assigned to these categories and the respective cumulative distributions, equation (14)

$$F_i = \sum_{j=1}^m p_j(f_i) \quad O_i = \sum_{j=1}^m p_j(o_i) \quad (14)$$

are calculated, where F_i is the cumulative ensemble distribution of forecast i and $p_j(f_i)$ the relative frequency of an ensemble member falling into category j . For each forecast i there is only one observation o_i ; hence the category j the observation is falling in is given a relative frequency of $p_j(o_i) = 1$ while all others are set to 0. Finally, the RPS, equation (15) for n forecasts is the average of the sum of the squared differences of the cumulative distributions.

$$RPS = \sum_{i=1}^n \left\{ \sum_{k=1}^m \left[\sum_{j=1}^k p_j(f_i) - \sum_{j=1}^k p_j(o_i) \right]^2 \right\} \quad (15)$$

The RPS penalises forecasts more severely when their probabilities are further from the actual observations. The relative improvement or skill of a probability forecast over climatology as a reference forecast is assessed by the ranked probability skill score RPSS, equation (16)

$$RPSS = 1 - \frac{RPS}{RPS_{ref}} \quad (16)$$

5 where RPS_{ref} is the RPS calculated with a constant forecast, e.g. the average of the observed series.

3 Results and discussion

The development of the SRM+G forecasting model for the UIB has been an iterative process with the focus on creating an operational forecasting tool for Kharif flow volumes to the Tarbela reservoir. Thus not all improvements
10 have been tested individually while not changing the other components which would allow an independent assessment of the individual effects. Nevertheless are the results discussed below separately for each component.

3.1 Splitting of the UIB catchment

While the simulation results using a sole model for the whole UIB showed an acceptable agreement between simulated and observed flows in terms of R^2 and D_s , initial hindcast results proved to be not satisfactory, especially
15 for the Early Kharif (1st April – 10th June) season, which is the major snowmelt contribution period. The mean percentage error MPE between hindcasts and observations of -21.0% for the years 2003 – 2014 indicated a severe bias towards under-estimating the actual flows and the respective mean absolute percentage error MAPE of 25.9% was also unexpectedly large.

An analysis of MODIS snow cover data indicates that in the south-eastern part of the UIB namely the Tibetan
20 plateau already in March the snow cover is fading away rapidly while on the other hand, in the north-western part of the catchment the same elevation zone is still widely covered with snow (Figure 6). In Table 4 the snow cover area of the relevant elevation zones for the south-eastern (Upper) and north-western (Lower) part of the UIB is given on 1st March and 1st April as an example for the year 2003. While at an elevation of 4,000 m a.s.l. the snow cover area reduces from 82% to 71% in the Lower UIB, in the Upper UIB the snow cover area shrinks sharply from
25 79% to 50%. A similar behavior can be observed for most of the other years as well.

As in forecasting mode the depletion of the snow cover area during the whole forecasting period is predicted depending on the reduction in the “key zone” in March (see Chap. 2.6), the relatively larger depletion in the Upper
UIB leads to an under-estimation of the available snow-water equivalent for the whole catchment, which explains the subsequent under-estimation of Early Kharif flows by the initial hindcasts and in turn motivates the splitting of
30 the UIB into two sub-catchments and separate models respectively (see Chap. 2.3). As a result of this splitting, the MPE of the hindcasts for Early Kharif changed to a modest over-estimation of 4.2% while the MAPE could be reduced to 15.8%.

3.2 Glacier melt component

During a first diagnostic calibration of the degree-day factors it became obvious, that in late summer even with
35 extreme high degree-day factors it was usually not possible to reproduce the observed hydrograph. The analysis of

Commented [i8]: New text with reasoning of split the UIB into Upper and Lower sub-catchments.

Commented [i9]: New chapter included about the reasoning of incorporating the glacier component.

the snow cover depletion showed, that in most years snow has vanished from areas below 4,000 m a.s.l. already in June and elevation zones below 5,000 m a.s.l. usually become snow-free in July. Thus, the snow-melt contribution to the flow is rapidly diminishing in August and September (see Figure 7).

Although in July usually the monsoon season starts bringing the highest monthly precipitation depth to the UIB in July and August, the resulting flow component from rain is not sufficient to create the necessary discharge. Therefore, many studies postulate a substantial contribution from glacier melt to the annual flow in the UIB, e.g. Immerzeel et al. (2010) estimate a contribution of 40% from snow and 32% from glacier melt with the remaining 28% from rain, Charles et al. (2017) contributes 50% to snow- and 20% to glacial melt.

Figure 7 shows the monthly distribution of the three flow components as calculated by SRM+G before subjecting to the recession flow calculation according to eq. (1). SRM's simple recession flow approach is not mass conservative and does also not allow a direct attribution, at what day these flow components actually occur in the daily discharge Q_{n+1} . However, an overall water balance shows, that the difference between water going into the (virtual) storage and water taken out by the recession flow term Q_n is about 7% which seems acceptable in relation to the uncertainty associated with the input data. Having in mind the above limitation, the average (2003 – 2014) flow component distribution as simulated by SRM+G is 53%, 21%, and 26% for snow, glacier and rain respectively, which is well in the magnitude of the values found in other studies.

Figure 8 and 9 compare the hydrographs of simulation runs with and without the glacier component for Upper and Lower UIB. The effect is more visible in the Lower UIB as about 10.5% of the catchment is glaciated, while for the Upper UIB the glaciated area is merely 1.7%.

3.3 Simulation model verification

The simulation model was verified by comparing full year (1st January – 31th December) simulation runs using the actual temperature, precipitation and snow cover data versus observed daily flows. Examples of respective hydrographs for Upper and Lower UIB are given in Figure 10 and 11. The resulting coefficients of determination R^2 and relative volume differences D_v for each year and both, the Upper and Lower UIB model, are given in Table 5. Although the years 2003 – 2012 were used to calibrate certain model parameters (see Chap 2.5), they are regarded to validate the model in “forecasting mode”, as in particular the degree-day factor functions (Table 1 and 2) were applied as during a real forecasting procedure, i.e. the starting point was selected according to Tables 1 and 2 depending on the historic 10-days temperatures for each year. The years 2013 and 2014 on the other hand were not used at the time of model calibration; thus they represent a fully independent verification of the simulation model. The average R^2 of 0.86 and 0.88 for the Upper and Lower UIB respectively indicate that in general the two models simulate the variations of the observed hydrographs quite acceptable. Values for the years 2013 and 2014 are even above the average for Lower UIB, which is finally essential for Tarbela inflows, but extreme floods, like 2010 during monsoon season, are reproduced not that well. The average relative volume differences D_v of -1.89% and 0.03% for Upper and Lower UIB respectively show, that the simulation models, although not mass-conservative due to the recession approach (Chap. 2.2) and volume differences vary from year to year, are in terms of total flow volume not biased over the long run.

Commented [i10]: New chapter about simulation model verification

3.4 Evaluation of forecasting skills

Commented [i11]: New chapter about the evaluation of forecast skills

In order to evaluate the skills of the forecasting model, hindcasts were carried out for the years 2003 – 2014 using always all years, i.e. also the hindcasted year, as scenario members. For the years 2015 and 2016 real forecasts have been determined before 1st April of these years, thus without using the particular year as a scenario. In all cases, the expected depletion of snow cover area was predicted for each scenario member based on the respective situation in March of the specific year.

In Table 6 the ensemble medians of hind- / forecasts by SRM+G are compared with observed flows as well as with IRSA's forecasts that are based on a statistical model and with forecasts from the UBC⁷ watershed model (Quick and Pipes, 1977) that is used by WAPDA's Glacier Monitoring Research Centre. All values are total Kharif (1st April – 30th September) flow volumes in 10⁹ m³. Figure 12 presents all model results and observed flows and shows also the 20% and 80% quantiles of the SRM+G scenario ensemble. Table 7 summarises the metrics that are used to compare the forecast skills of the three models. The RMSE, MAE and MAPE show an improvement in accuracy by SRM+G and the MPE a reduction of bias where SRM+G tends to slightly under-estimate, while the two other forecasts moderately over-estimate the total Kharif flows. Both, the correlation coefficients R and the anomaly coefficients AC_u indicate however, that the association between forecasts and observations is weak for all three models. Here, the UBC forecasts show the best correlation followed by SRM+G and IRSA. The above aspects of model performance are synoptically visualised in the Taylor diagram (Taylor, 2001) in Figure 13 that was plotted using the R-package Plotrix (Lemon, 2006). All models are comparable far away from the point of observations given on the x-axis, with SRM+G having the smallest centered root mean square difference and UBC the best correlation coefficient.

In order to evaluate the model skills in forecasting extreme, i.e. dry or wet, conditions, the Peirce skill score was applied. The limits between the categories dry, normal, and wet conditions were defined as 20% and 80% non-exceedance of the observed historic Kharif flow series 2003 – 2016, which corresponds to quantiles of 56.8 km³ and 67.9 km³ respectively. Obviously IRSA and SRM+G forecasts have no skill in this respect, while UBC shows some, however limited skill compared to purely random chance.

As only point estimates of IRSA and UBC forecasts were available, the assessment of probabilistic skill only applies to the SRM+G scenario ensembles. Figures 14 and 15 show typical traces of ensemble members as well as the observed, mean, and historic trace. In most years the ensemble mean is closer to the observed value than the historic trace.

The ranked probability skill score RPSS was used to assess the overall performance of the probabilistic forecast. Same category limits as for the Peirce skill score were chosen, i.e. the 20% and 80% percentiles of the observed flow series. As no other probabilistic forecasts were available as reference forecast, the ranked probability score RPS of the scenario forecast ensembles was compared to the climatology, i.e. the average Kharif flow of the observed series of 62.7 km³. The RPS of scenario forecast ensembles and climatology were 0.348 and 0.462 respectively. The resulting RPSS of 0.25 indicates that the scenario ensemble shows some, however limited skill and improvement over a constant forecast of the historic average.

⁷ University of British Columbia Watershed Model

4 Conclusions

The snowmelt runoff model, which in combination with a scenario approach is successfully applied to predict seasonal flows e.g. in the Upper Jhelum catchment, was applied to the Upper Indus Basin in order to forecast the total Kharif inflow to the Tarbela reservoir. Several improvements had to be introduced to SRM in order to meet the specific requirements of the UIB.

Not surprisingly, a separate component had to be added to SRM in order to consider the flow component originating from glacier melt. Without this component, especially in the late summer months there is a lack of water to meet the observed hydrograph. All year simulation runs with SRM+G for the period 2003 – 2014 result in an average flow component distribution of 53%, 21%, and 26% for snow, glacier and rain respectively, which fits well to the values found in a number of other studies.

It is well known, that the Tibetan plateau receives significantly lesser precipitation than the western parts of the UIB. In addition, MODIS data shows that the snow cover is fading away in early spring much faster than in the other parts. In the present study, SRM's modified depletion curve approach for predicting the snow cover depletion during the forecast period has proven to be very sensitive to errors in the estimation of the actual snow-water equivalent. In such cases it is inevitable to split the catchment into more homogeneous units. Therefore, the superposition of flows from sub-catchments by using a time-lag was introduced to SRM+G, which leads to a significant improvement in the forecasts of snow-melt dominated Early Kharif flows in particular.

The scenario approach is a step towards probabilistic forecasting of seasonal flows in the UIB. As the accuracy of existing forecasts with a mean volume error of 10.9% and 11.4% is already quite high, the improvement by SRM+G having a MAPE of 9.5% is only limited. The bias however could be reduced to -2.0%. Association between forecasts and observations is rather weak for all three models, just as none of the models has significant skill in predicting extreme dry or wet conditions.

Regarding the scenario approach it is obvious, that as far as the variables precipitation and temperature are concerned, these tend toward the climatology, i.e. the long term averages. A variance in the forecasts is only introduced by the different estimates of the snow-cover depletion curves for each forecast. Thus a promising way to improve the association and sharpness of the scenario approach would be a selection of a subset of ensemble members according to forecasted seasonal anomalies in temperature and precipitation. A quick test using only the five lowest, middle, or highest ensemble members selected according to the (known) relative flow frequency of the forecasted year gives promising results, e.g. not only a MAPE of 4.9%, but also an AC_u of 0.78 and a PSS of 0.41.

The challenge of course is to forecast the seasonal anomaly in temperature and precipitation. In this respect further research is needed on how today's global forecast systems may allow a more specific selection of ensemble members particularly in the UIB, where the correlation to common teleconnections like the ENSO⁸ status is known to be weak.

Commented [i12]: Updated text with respect to the new analysis being carried out in the revised manuscript.

⁸ El Niño Southern Oscillation

Acknowledgements

The authors thank National Engineering Services Pakistan (Pvt.) Ltd. (NESPAK), Lahore and AHT Group AG, Essen, Germany that they could be part of the project team. They are highly grateful to Indus River System Authority (IRSA) and WAPDA's Glacier Monitoring Research Centre (GMRC) and Surface Water Hydrology Project (SWHP) of for sharing their forecast results as well as the daily discharge data. The authors are also thankful to the two anonymous reviewers for their valuable and constructive comments that substantially helped to improve the quality of the manuscript.

References

- Akram, A. A.: Indus Basin water resources, *Tiempo*, Issue 70,,
10 <http://www.environmentportal.in/files/Indus%20Basin.pdf>, 2009.
- Archer, D. R. and Fowler, H. J.: Spatial and temporal variations in precipitation in the Upper Indus Basin, global teleconnections and hydrological implications, *Hydrol. Earth Syst. Sci.*, 8, 47–61, doi:10.5194/hess-8-47-2004, 2004.
- Bogacki, W. and Hashmi, D.: Impact of Climate Change on the Flow Regime of the Mangla Basin. GWSP
15 Conference 'Water in the Anthropocene: Challenges for Science and Governance. Indicators, Thresholds and Uncertainties of the Global Water System, Bonn, http://conference2013.gwsp.org/uploads/media/Bogacki_Impact_of_Climate_Change_on_the_Flow_Regime_of_the_Mangla_Basin.pdf, 2013.
- Bogacki, W. and Ismail, M. F.: Seasonal forecast of Kharif flows from Upper Jhelum catchment. In Proc. IAHS
20 374, pp. 137–142. DOI: 10.5194/piahs-374-137-2016.
- Bookhagen, B. and Burbank, D.W.: Toward a complete Himalayan hydrological budget: Spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge, *J. Geophys. Res.*, 115, F03019, doi:10.1029/2009JF001426, 2010.
- Charles, S. P., Wang, Q. J., Ahmad, M.-U.-D., Hashmi, D., Schepen, A., and Podger, G.: Assessment of methods
25 for seasonal streamflow forecasting in the Upper Indus Basin of Pakistan, *Hydrol. Earth Syst. Sci. Discuss.*, <https://doi.org/10.5194/hess-2017-428>, in review, 2017.
- Clarke, M.: Climate Change Considerations for Hydropower Projects in the Indus River Basin, Pakistan. Impact Assessment in the Digital Era, 35th Annual Conference of the International Association for Impact Assessment, Italy, 2015.
- 30 Day, G. N.: Extended Streamflow Forecasting Using NWSRFS. In *Journal of Water Resources Planning and Management* 111 (2), pp. 157–170. DOI: 10.1061/(ASCE)0733-9496(1985) 111:2(157), 1985.
- Forsythe, N., Fowler, H. J., Kilsby, C. G. and Archer, D. R.: Opportunities from remote sensing for supporting water resources management in village/valley scale catchments in the Upper Indus Basin. *Water Resour. Manage.* 26 (4), 845–871, 2011.
- 35 Franz, K J.; Hogue, T.S.; Sorooshian, S: Snow Model Verification Using Ensemble Prediction and Operational Benchmarks. In *J. Hydrometeorol* 9 (6), pp. 1402–1415, 2008. DOI: 10.1175/2008JHM995.1.
- Franz, K. J., Hartmann, H. C., Sorooshian, S., and Bales, R.: Verification of National Weather Service Ensemble Streamflow Predictions for water supply forecasting in the Colorado River basin, *J. Hydrometeorol.*, 4, 1105–1118, 2003.
- 40 Hock, R.: Temperature index melt modelling in mountain areas. *Journal of Hydrology* 282, 104–115, 2003.
- Immerzeel, W.W., van Beek, L.P.H. and Bierkens, M.F.P.: Climate change will affect the Asian water towers. *Science*, 328(5984): 1382-1385, 2010.
- Immerzeel, W.W.; Droogers, P.; de Jong, S.M.; Bierkens, M.F.P.: Satellite Derived Snow and Runoff Dynamics in the Upper Indus River Basin." *Grazer Schriften der Geographie und Raumforschung* 45:303-312, 2010.

Commented [i13]: New reference.

Commented [i14]: New reference.

Commented [i15]: New reference.

Commented [i16]: New reference.

- Immerzeel, W. W., Droogers, P., de Jong, S. M., and Bierkens, M. F. P.: Large-scale monitoring of snow cover and runoff simulation in Himalayan river basins using remote sensing, *Remote Sensing of Environment*, 113, 40–49, doi:10.1016/j.rse.2008.08.010, 2009.
- 5 Ismail, M. F., Rehman, H., Bogacki, W., and Noor, M.: Degree Day Factor Models for Forecasting the Snowmelt Runoff for Naran Watershed, *Sci. Int. Lahore*, 27(3), 1961-1969, 2015.
- Kirpich, Z.P.: Time of concentration in small agricultural watersheds. *Civil Engineering*, 10 (6): 362-, 1940
- Lemon, J: Plotrix: a package in the red light district of R. *R-News*, 6(4): 8-12, 2006.
- Martinec, J., Rango, A., and Roberts, R.: *Snowmelt Runoff Model User's Manual, WinSRM Version 1.14.* Agricultural Experiment Station Special Report 100, New Mexico State University, Las Cruces, NM 88003, USA, 2011.
- 10 Martinec, J.: Snowmelt-Runoff Model for Stream Flow Forecasts, *Nord. Hydrol.*, 6, 145–154, 1975.
- Ménégoz, M., Gallée, H. and Jacobi, H. W.: Precipitation and snow cover in the Himalaya: from reanalysis to regional climate simulations. *Hydrol. Earth Syst. Sci.* 17, 3921–3936, 2013.
- Palazzi, E., Hardenberg, J. and Provenzale, A.: Precipitation in the Hindu-Kush Karakoram Himalaya: Observations and future scenarios, *J. Geophys. Res. Atmos.*, 118, 85–100, doi: 10.1029/2012JD018697, 2013.
- 15 Pellicciotti, F., Buerger, C., Immerzeel, W. W., Konz, M., and Shrestha, A. B.: Challenges and Uncertainties in Hydrological Modeling of Remote Hindu Kush–Karakoram–Himalayan (HKH) Basins. Suggestions for Calibration Strategies. In *Mountain Research and Development* 32 (1), pp. 39–50. DOI: 10.1659/MRD-JOURNAL-D-11-00092.1, 2012.
- Quick, M. C. and Pipes, A.: UBC watershed model, *Hydrol. Sci. Bull.*, 221, 153–161, 1977.
- Raup, B.H., Racoviteanu, A., Khalsa, S.J.S., Helm, C., Armstrong, R. and Arnaud Y.: "The GLIMS Geospatial Glacier Database: a New Tool for Studying Glacier Change". *Global and Planetary Change* 56:101–110, 2007.
- Reggiani, P. and Rientjes, T. H. M.: A reflection on the long-term water balance of the Upper Indus Basin. *Hydrol. Res.*, 46(3), 446–462, DOI:10.2166/nh.2014.060, 2015.
- 25 Schaper J, Seidel K.: Modelling daily runoff from snow and glacier melt using remote sensing data. In: *Proceedings of the 2nd EARSel Workshop—Special Interest Group Land Ice and Snow* (pp. 308-316), 2000.
- Schaper, J., Martinec, J., And Seidel, K.: Distributed Mapping of Snow and Glaciers for Improved Runoff Modelling. *Hydrological Processes*, 13(12-13): 2023–2031, 1999.
- Schaper, J., Seidel, K., and Martinec, J.: Precision snow cover and glacier mapping for runoff modelling in a high alpine basin. *IAHS Publ. no. 267*, 2001, 105–111, 2000.
- 30 Tahir, A. A.; Chevallier, P.; Arnaud, Y.; Ahmad, B.: Snow cover dynamics and hydrological regime of the Hunza River basin, Karakoram Range, Northern Pakistan. *Hydrol. Earth Syst. Sci.* 15 (7), pp. 2275–2290. DOI: 10.5194/hess-15-2275-2011.
- Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram. In *J. Geophys. Res.* 106 (D7), 7183–7192. DOI: 10.1029/2000JD900719, 2001
- 35 USDA: National Engineering Handbook. Section 4: Part 630 Hydrology Chapters. U.S. Department of Agriculture, Natural Resources Conservation Service. Available online at <https://www.wcc.nrcs.usda.gov/ftpref/wntsc/H&H/NEHhydrology/ch15.pdf>, updated on 2010
- Viviroli, D., Dürr, H. H., Messerli, B., Meybeck, M., and Weingartner, R.: Mountains of the world, water towers for humanity. Typology, mapping, and global significance. *Water Resour. Res.* 43 (7). DOI: 10.1029/2006WR005653, 2007.
- 40 Ward, E., Buytaert, W., Peaver, H., and Wheeler, L.: Evaluation of precipitation products over complex mountainous terrain: A water resources perspective. *Adv. Water Resour.* 34 (10), 1222–1231, 2011.
- Wood, A. W.; Lettenmaier, D. P.: An ensemble approach for attribution of hydrologic prediction uncertainty. In *Geophys. Res. Lett.* 35 (14), p. 642, 2008. DOI: 10.1029/2008GL034648.
- 45 Wilks, D.S. *Statistical Methods in Atmospheric Sciences*, 2nd ed. ed.; Elsevier Science & Technology Books: Amsterdam, The Netherlands, 2006.
- Xie, P., Yarosh, Y., Love, T., Jonowiak, J.E., and Arkin, P.A.: A real-time daily precipitation analysis over south asia. Preprint of the 16th Conference on Hydrology, Orlando, Florida. American Meteorological Society, Washington DC, USA. http://www.cpc.ncep.noaa.gov/products/fews/sasia_rfe.pdf, 2002.
- 50

Commented [i17]: New reference.

Commented [i18]: New reference.

Commented [i19]: New reference.

Commented [i20]: New reference.

Commented [i21]: New reference.

Commented [i22]: New reference.

Commented [i23]: New reference.

Table 1: Zone-wise degree-day factors and 10-daily temperature threshold depending on 10-daily periods after melting start for Lower UIB

10-daily Period	Elevation Zone (m a.s.l)							
	0-2500	3000	3500	4000	4500	5000	5500	>5500
T_{10d}^9	9.0	7.0	5.0	4.0	2.0	1.0	1.0	1.0
1	0.20	0.21	0.22	0.22	0.19	0.18	0.18	0.20
2	0.30	0.32	0.32	0.32	0.30	0.31	0.31	0.33
3	0.39	0.43	0.41	0.43	0.41	0.43	0.44	0.46
4	0.48	0.53	0.51	0.54	0.52	0.56	0.57	0.59
5	0.57	0.64	0.61	0.65	0.63	0.68	0.70	0.72
6	0.67	0.75	0.70	0.80	0.74	0.80	0.80	0.80
7	0.80	0.80	0.80	0.80	0.80			

5 Table 2: Zone-wise degree-day factors and 10-daily temperature threshold depending on 10-daily periods after melting start for Upper UIB

10-daily Period	Elevation Zone (m a.s.l)						
	3000	3500	4000	4500	5000	5500	>5500
T_{10d}	2.0	2.0	2.0	2.0	0.5	0.5	0.5
1	0.37	0.35	0.35	0.52	0.56	0.48	0.60
2	0.43	0.40	0.40	0.59	0.64	0.54	0.70
3	0.49	0.45	0.46	0.66	0.73	0.80	0.80
4	0.54	0.51	0.51	0.73	0.80		
5	0.60	0.56	0.56	0.80			
6	0.66	0.61	0.62				
7	0.71	0.66	0.67				

10

15

⁹ 10-Daily average temperature in °C in each elevation zone.

Table 3: SRM+G Model Parameters for both Upper and Lower UIB

Commented [i24]: Table about parameters has been incorporated.

Parameters	Symbol	Value	Units	Remarks
Temperature Lapse-Rate	γ	6.0	$^{\circ}\text{C km}^{-1}$	
Recession Coefficient	k_x	1.193	-	October-February
		1.060		March – September
	k_y	0.029		October-February
		0.020		March – September
Critical Precipitation	P_{crit}	1	cm	constant
Lag Time	L	54	h	2.5 days delay between melt and runoff at Tarbela
Critical Temperature	T_{crit}	0.5 – 3.0	$^{\circ}\text{C}$	variable
Rainfall Contributing Area	RCA	0	-	November – March
		1		April – October
Runoff Coefficient - Snow	c_S	0.80	-	constant
Runoff Coefficient - Glacier	c_G	0.70	-	constant
Runoff Coefficient - Rain	c_R	0.25-0.75	-	
Degree-Day Factor - Snow	α	0.15-0.80	$\text{cm } ^{\circ}\text{C}^{-1}\text{d}^{-1}$	
Degree-Day Factor - Glacier	α_G	0.70	$\text{cm } ^{\circ}\text{C}^{-1}\text{d}^{-1}$	constant

Table 4: Depletion of snow cover area for Upper and Lower UIB during March 2003

Elevation (m asl)	3500	4000	4500	5000	5500	>5500
1 st March						
Lower UIB	66%	82%	88%	87%	83%	94%
Upper UIB	58%	79%	58%	51%	58%	71%
1 st April						
Lower UIB	42%	71%	84%	84%	78%	92%
Upper UIB	24%	50%	48%	43%	51%	73%

5

10

Table 5: Coefficient of determination R² and relative volume difference D_v for Upper and Lower UIB

Year	Upper UIB		Lower UIB	
	R ²	D _v [%]	R ²	D _v [%]
2003	0.86	-17.3	0.92	4.6
2004	0.84	2.2	0.90	0.1
2005	0.89	15.3	0.83	-17.4
2006	0.85	8.4	0.91	-3.5
2007	0.80	4.0	0.88	-4.1
2008	0.94	-6.7	0.92	-1.4
2009	0.79	14.2	0.86	16.4
2010	0.90	-1.9	0.77	-16.3
2011	0.88	-9.1	0.88	4.5
2012	0.87	-16.0	0.89	11.9
2013	0.77	-9.3	0.93	-2.5
2014	0.88	-6.5	0.92	8.0
Average	0.86	-1.89	0.88	0.03

Commented [i25]: Table about Model calibration and validation results has been incorporated.

Table 6: Comparison of Kharif flow volumes [km³] 2003 - 2016

Year	Observed	IRSA	UBC	SRM+G
2003	67.8	64.0	63.5	63.1
2004	51.8	60.5	63.6	60.8
2005	68.9	69.0	73.3	60.9
2006	67.8	68.4	73.3	61.6
2007	60.5	74.9	70.1	61.0
2008	57.7	68.5	59.2	53.9
2009	57.6	63.7	67.2	62.4
2010	76.6	63.3	68.4	61.4
2011	60.0	67.2	70.8	59.9
2012	55.4	61.3	61.7	60.4
2013	65.6	64.9	58.8	59.8
2014	52.9	64.6	64.2	61.4
2015	67.2	63.3	61.3	58.9
2016	66.4	62.4	66.5	63.1

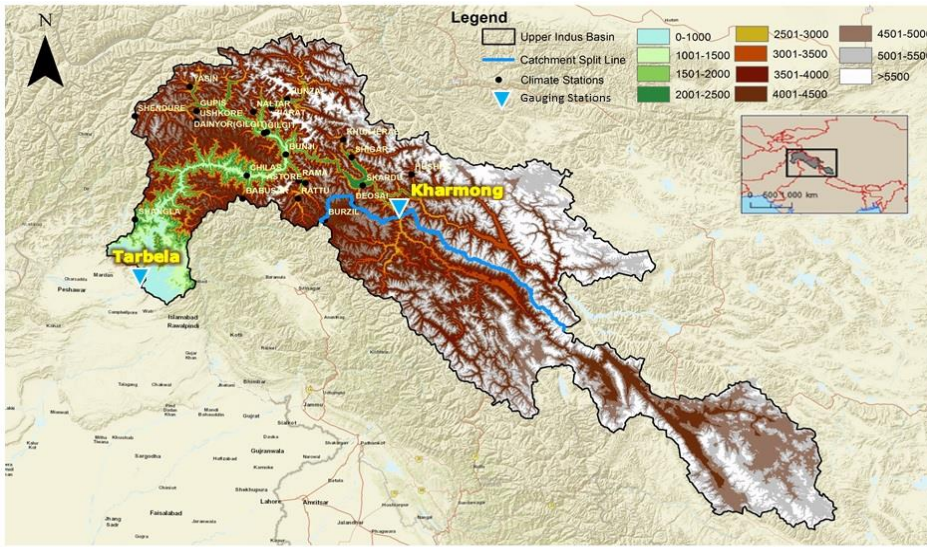
Commented [i26]: Comparison of Kharif flow volumes has been incorporated.

5

Table 7: Comparison of forecast skills between IRSA, UBC, and SRM+G

Model	MAE km ³	RMSE km ³	MPE %	MAPE %	R	AC _u	PSS
IRSA	6.5	8.0	5.8	10.9	0.107	0.085	-0.070
UBC	6.9	7.7	6.3	11.4	0.318	0.260	0.096
SRM+G	6.0	7.0	-2.0	9.5	0.223	0.168	-0.079

Commented [i27]: Table about comparison of forecast skills using different available methods has been incorporated.



Background Map Source: ESRI Open Street Map

Figure 1: Map of the Upper Indus Basin showing different elevations and splitting of UIB at the Kharmonj gauging station into Upper and Lower UIB

Commented [i28]: Figure 1 has been updated. Now the catchment as well as splitting is shown in a single figure.

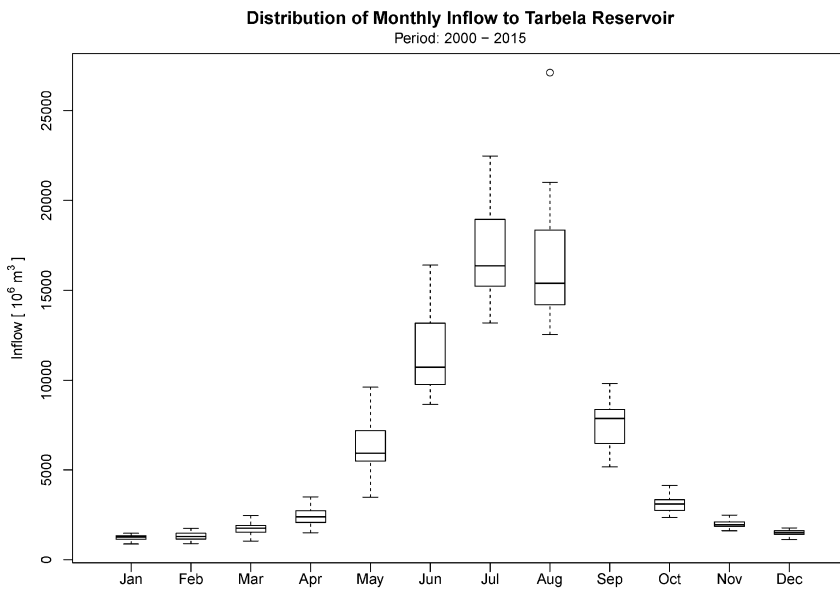


Figure 2: Monthly distribution of inflows to the Tarbela Reservoir from 2000 - 2015

5

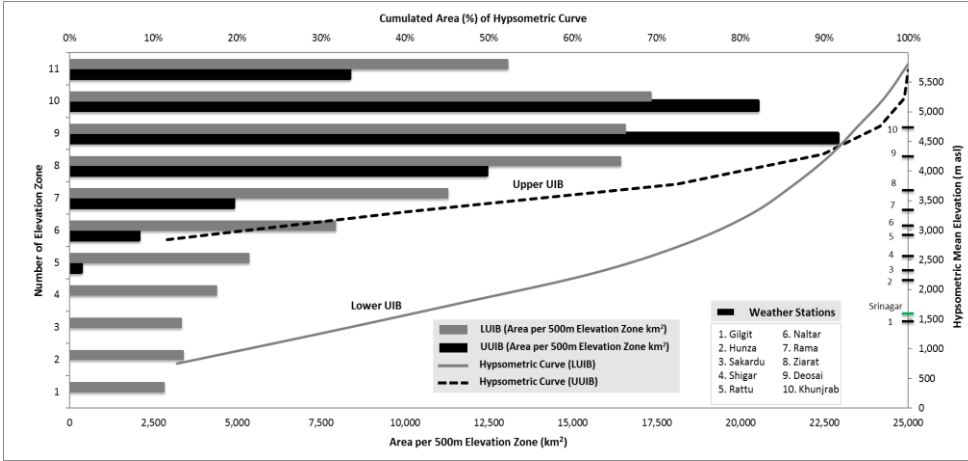


Figure 3: Hypsometric curves and the distribution of area under 500-m elevation bands for the Upper and Lower UIB. Eleven as well as seven elevation zones were made for Upper and Lower UIB and the elevation of the weather stations in western portion of the UIB are presented on the right hand side y-axis.

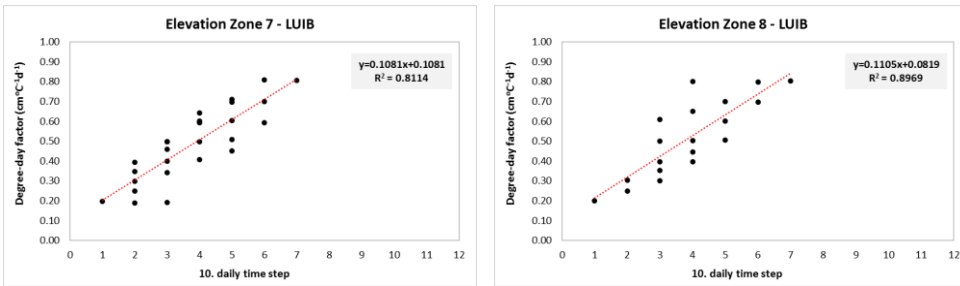


Figure 4: Increase of degree-day factors with time (10-days periods) after melting start for elevation zones 7 and 8 for Lower UIB. Degree-day factors are obtained by diagnostic calibration.

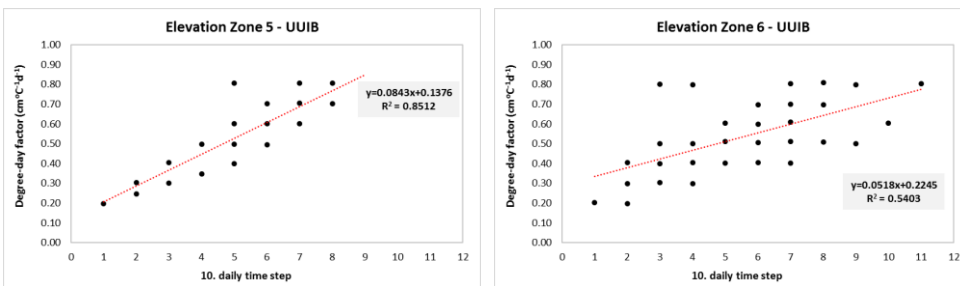


Figure 5: Increase of degree-day factors with time (10-days periods) after melting start for elevation zones 5 and 6 for Upper UIB. Degree-day factors are obtained by diagnostic calibration.

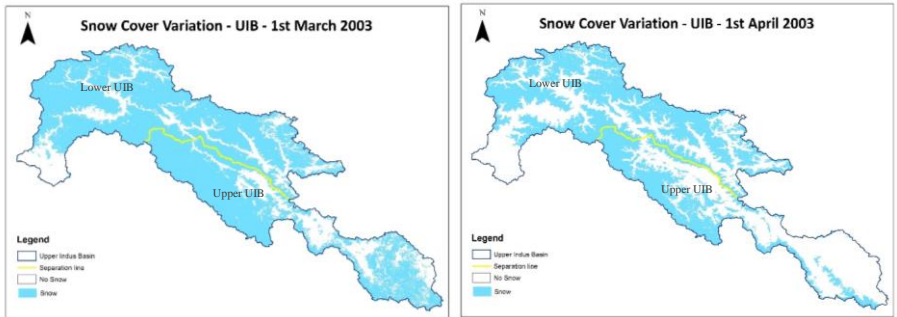


Figure 6: Snow cover variation in the month of March and April 2003 in UIB

Commented [i29]: New figure showing MODIS snow cover information in the lower and upper UIB on 1st March and 1st April 2003.

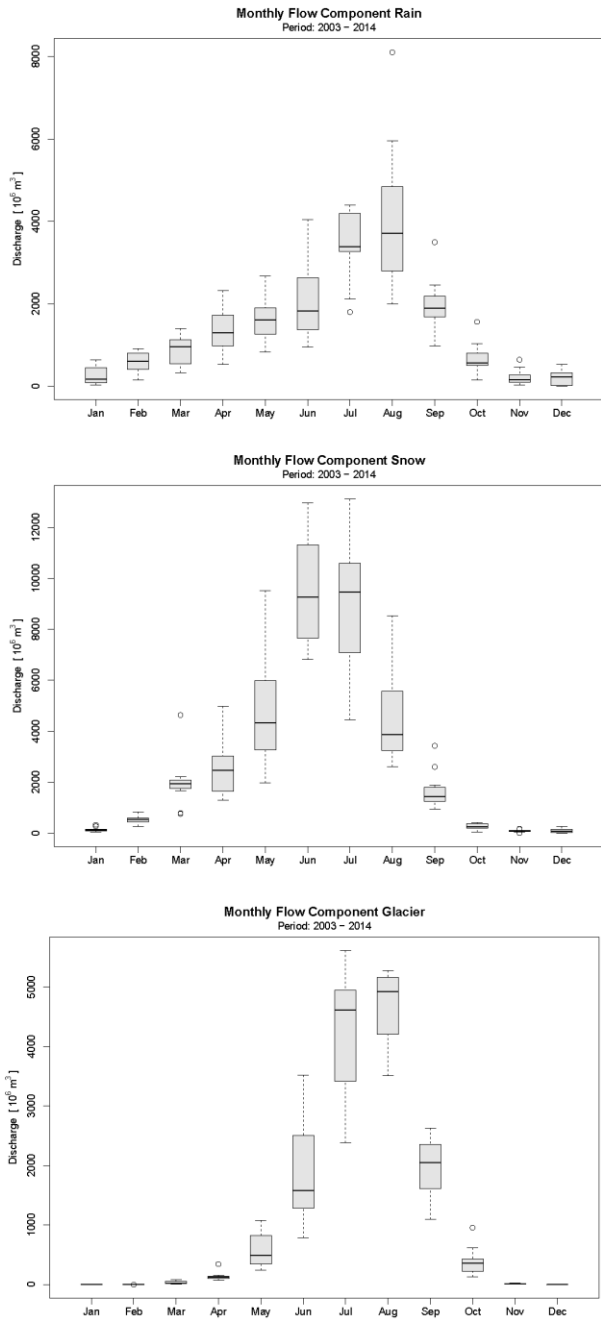


Figure 7: Monthly distribution of flow components in UIB

Commented [i30]: New figure included about the monthly distribution of flow component in UIB.

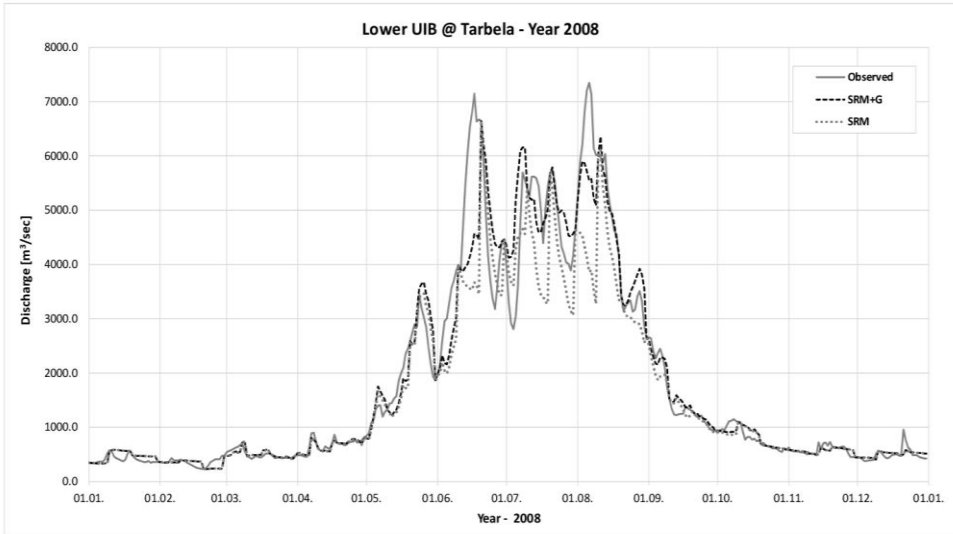


Figure 8: Comparison of SRM+G (with glaciers) and SRM (without glaciers) for Lower UIB – 2008

Commented [i31]: New figure added for the comparison of SRM+G (with glaciers) and SRM (without glaciers).

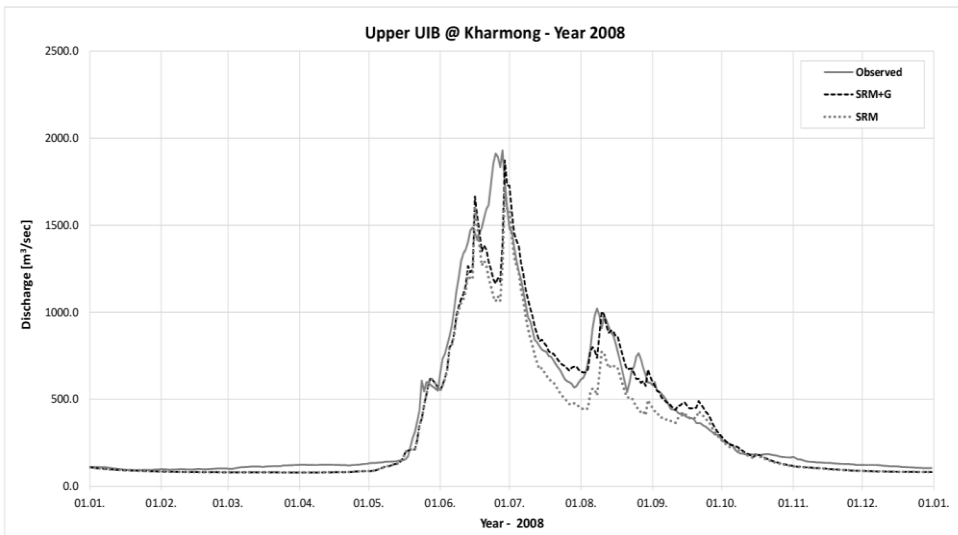


Figure 9: Comparison of SRM+G (with glaciers) and SRM (without glaciers) for Upper UIB - 2008

Commented [i32]: New figure added for the comparison of SRM+G (with glaciers) and SRM (without glaciers).

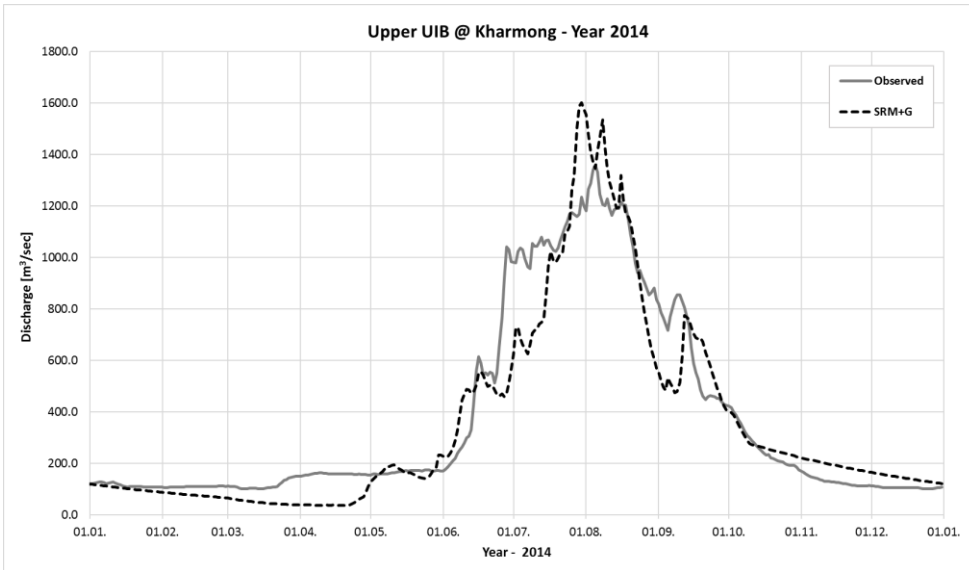


Figure 10: Results of validation of final Upper UIB flow forecast model (dashed line) compared to observed flows at Kharmong (solid line) for the year 2014.

Commented [i33]: updated

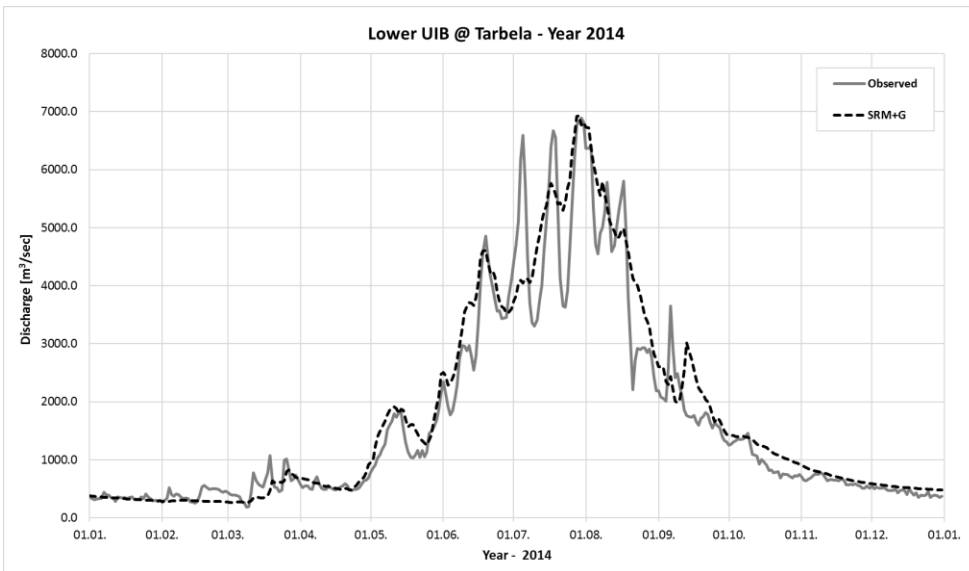


Figure 11: Results of validation of final Lower UIB flow forecast model (dashed line) compared to observed inflows at Tarbela (solid line) for the year 2014.

Commented [i34]: updated

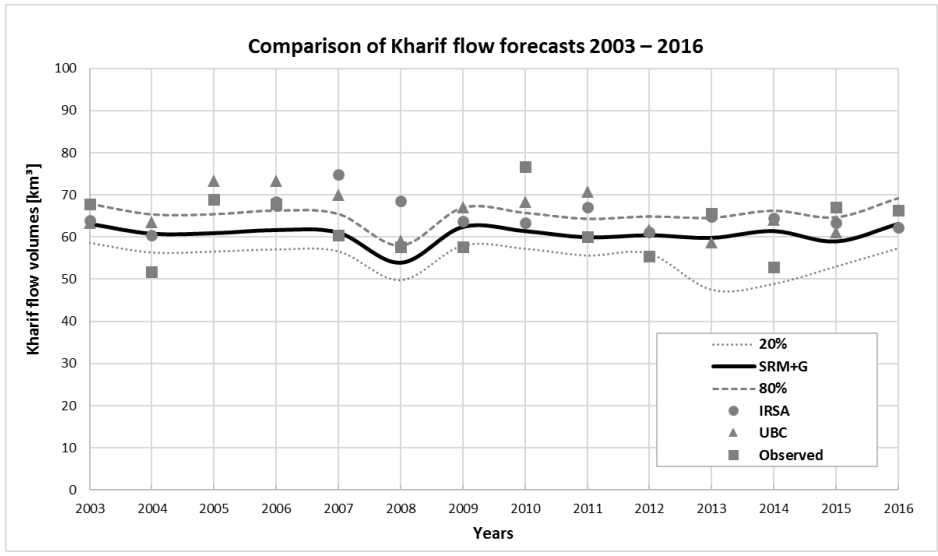


Figure 12: Comparison of Kharif flow forecast with 20% and 80% quantiles of SRM+G scenario ensembles years 2003 - 2016

Commented [i35]: New figure added about the dry and wet quantiles of SRM+G.

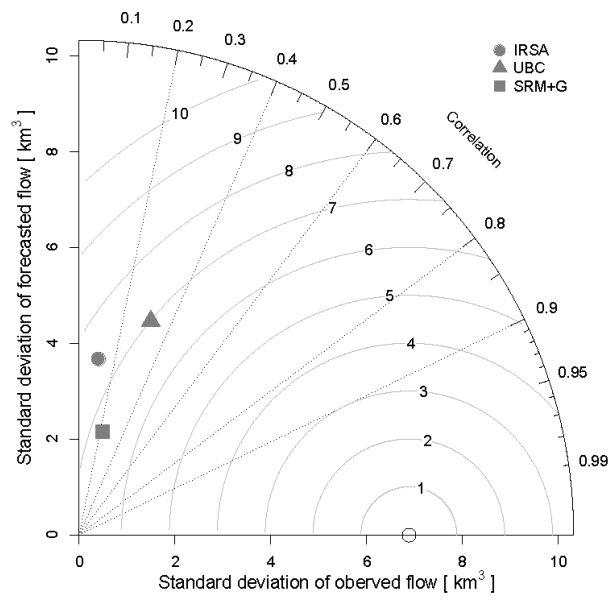


Figure 13: Taylor diagram of IRSA, UBC, and SRM+G model performance

Commented [i36]: New figure has been added for analysing the models performance.

5

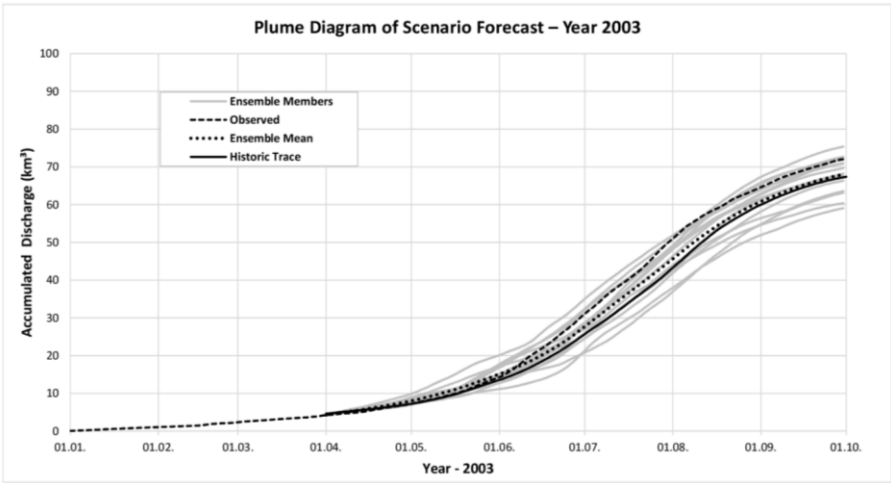


Figure 14: Plume diagram of ensemble member traces in 2003

Commented [i37]: New figure has been added about the ensemble member trace 2003

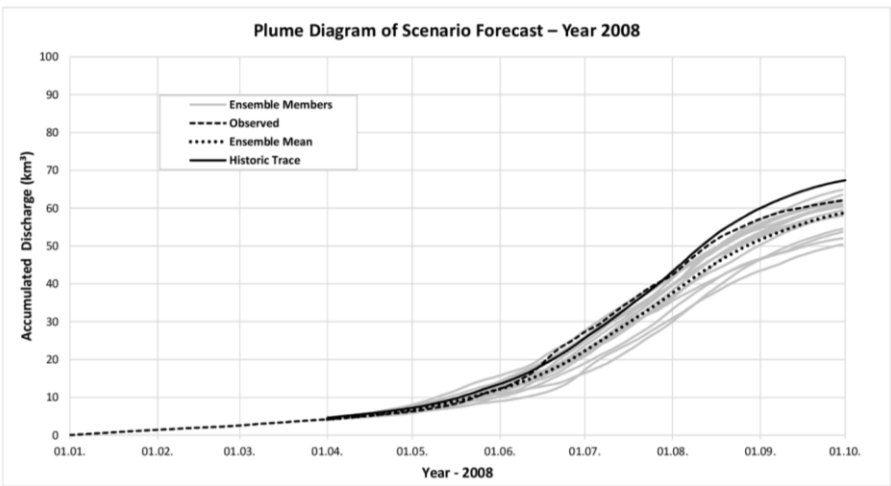


Figure 15: Plume diagram of ensemble member traces in 2008

Commented [i38]: New figure has been added about the ensemble member trace 2008