Dear Editor we thank you for your comment.

Editor: The first reviewer discusses that the Nash criterion is not a good indicator for strongly seasonal signals. In this context, it seems questionable whether SWAT has any predictive power for the validation catchment (without re-calibration). The authors state: "The efficiency of inflow predictions (NS) dropped to 0.49 and the R2 to 0.72, which are however satisfactory. The observed and predictive accumulative flow is presented in Fig.5(b)".

A Nash value of 0.49 for such a strongly seasonal signal might have no predictive power (Schaefli and Gupta, 2007). A simple experiment illustrates this: if you generate a sine curve that has the same seasonality as the observed discharge, similar amplitude and the same mean, and no negative values (e.g. by shifting the sine curve), then the Nash value of this signal (compared to the observed discharge) most likely has a Nash value of between 0.4 and 0.5. Attached to this comment is a Matlab example, including an illustration.

Given the above, I think that we need more evidence that the model actually has predictive power. A key question is hereby whether the model can predict winter low flows (i.e. it gets the baseflow right), general timing of snow melt, general timing if high flows, autumn recession etc.

In response to your comment, we investigated further whether the model can predict the spring snowmelt timing, timing of highest flow, autumn recession period and the centre of mass (COM). We used the 15-day average of the daily runoff and results are presented in Fig. 1 and 2 and the Table 1 given below. Figure 1 shows the observed and simulated spring snowmelt timing and Fig. 2 the highest flow timing for each year of the period 1997-2010. Table 1 shows the difference in days between the observed and simulated centre of mass and autumn recession period.

The model predicts satisfactorily the spring snowmelt timing and the autumn recession period. The difference between the COM of the observed and the simulated runoff, Table 1, is low and for some years close to zero, which is also satisfactory. Overall, we believe that these additional data prove the predictability of the model.

There are only some inconsistencies between measured and simulated data for the general timing of the highest flow, Fig. 2, and for certain years. These inconsistencies have two possible explanations. Firstly, the Damma glacier watershed is characterised by very steep slopes (even up to nearly 80 degrees) and runoff originates mainly from snowmelt, glacier melt and rainfall (Magnuson et al., 2012). Consequently the watershed is characterised by very fast response, which in terms of the model parameters resulted on the high value of ALPHA_BF and the low value of the GW_Delay. On the other hand, the Göscheneralpsee

feeding area is less steep on average and for the two out of the four of its watersheds, runoff is drained through tunnels into the reservoir. These two factors explain the difference in the response and the fact that the simulated runoff peaks are higher and narrower than the observed ones.

In addition, the model doesn't differentiate between snow and glacier dynamics and only one parameter for both snowmelt and a glaciermelt rate is applied. This becomes more important in our study, since there is a difference between the glacier coverage of the two catchments. The Damma glacier is 50% covered by the glacier while the greater catchment is 20%.

Table 1 Difference of the centre of mass (COM) and autumn recession period in days, calculated from the 15-day average.

Year	COM	Autumn recession period
1997	6.8	1
1998	4.2	1
1999	1.0	0
2000	3.0	16
2001	0.6	1
2002	7.8	19
2003	0.6	5
2004	2.4	4
2005	4.3	0
2006	4.1	1
2007	8.1	1
2008	3.1	0
2009	4.6	0
2010	6.0	0

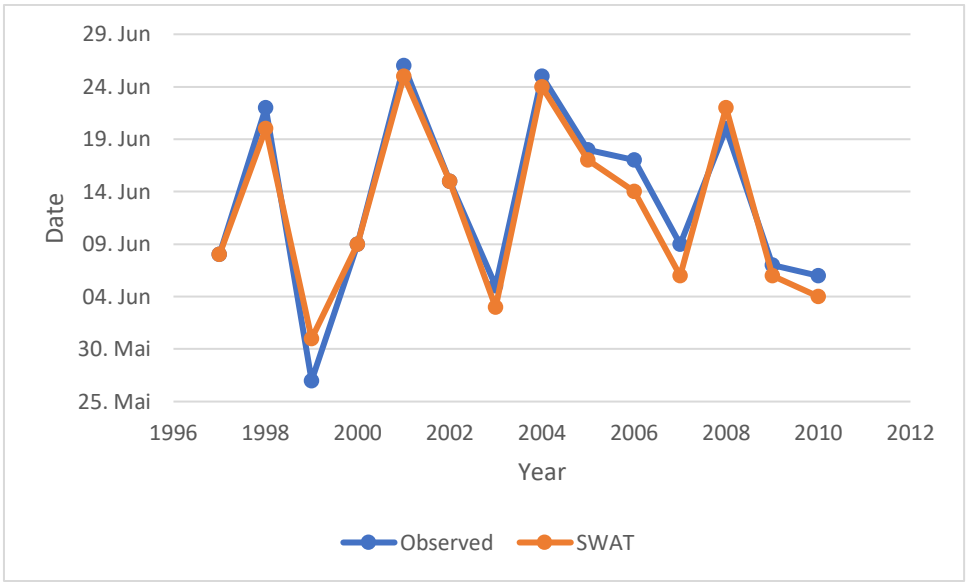

Figure 1 Comparison between the observed and simulated spring snowmelt timing. A 15-day average filter was applied on daily measurements.

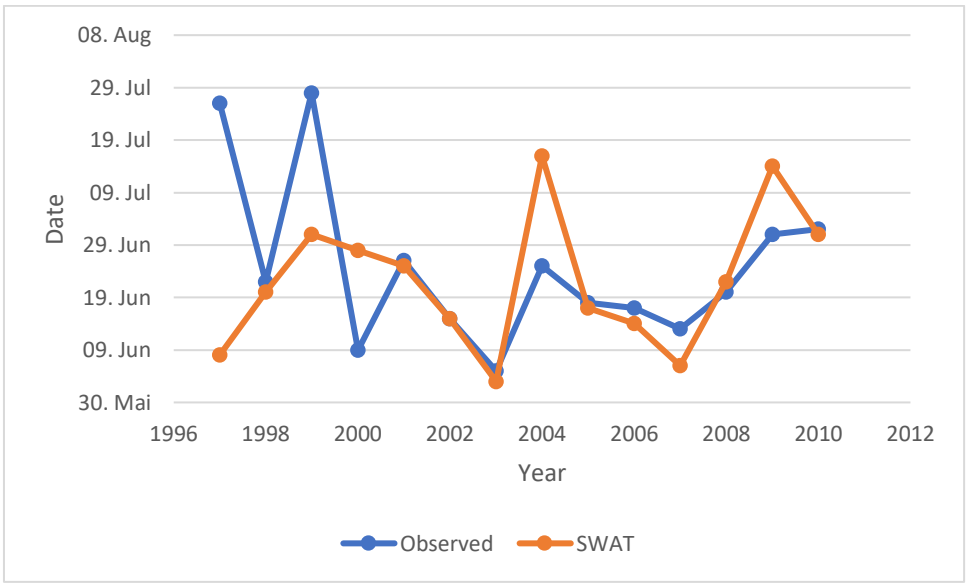

Figure 2 Comparison between the observed and simulated spring snowmelt timing. A 15-day average filter was applied on daily measurements.