Authors' reply

Dear Reviewers,

We thank you for a thorough review of the manuscript and constructive comments that we address point-by-point hereafter. We are confident that this resulted in an improved manuscript:

Reviewer #1: R. Fernandes

This paper discusses a straightforwards approach to quantitatively predict historical daily maps of binary snow cover status at 500m resolution over a watershed in Central Asia by exploiting statistical relationships between current colocated MODIS binary snow cover maps , in-situ snow depth measurements and digital elevation models. The approach is assessed with four binary snow cover maps derived from application of the MODIS algorithm to relatively cloud free Landsat TM imagery. Agreement rates range from 84% to 86.4% leading the authors to conclude that considering the ~ 92% accuracy of MODIS maps over the region the historical approach is reasonable and should be applicable.

The paper has some novelty in that it provides a repeatable approach for spatially extrapolating snow cover using station data.

The paper is mathemathically rigorous but suffers somewhat by introducing new notations for simple probability theory.

The paper does make generalizations and over simplifications in the use of satellite imagery and needs to provide more details on both the in-situ data and results.

A: We thank the reviewer for attesting the novelty to the presented research. See further below for answers related to in-situ data and results.

Much more care is required in presenting information related to the limits of this method. Especially accuracy and uncertainty during transition seasons and at edges of the snow field.

A: We acknowledge that some generalization was made in order to extrapolate the information on snow coverage to the period for which satellite imagery was not available. Discussion of limitations of the methodology will be extended in the revised version of the manuscript.

The conclusions are too broad - the study is for an ~ 120 x 120km region with specific land cover and climate conditions. I am not sure if the paper is applicable to this journal if it’s scope if limited but it is sufficiently novel that with some addition uncertainty information it should be of interest to readers.

A: The "limitations" and "conclusions" chapters will be reformulated and made more specific. In particular, we will explicitly point to the limited region of application in our case study, and to the fact that it might be difficult to exploit the statistical relationships between point measurements and aerial patterns in lowland areas. We believe, however, that the here presented methodology can be well applied also to other mountainous regions for which historical in-situ measurements are available.

Concerning the suitability of the presented manuscript to the journal scope: The presented methodology can be used to extend snow cover characteristics beyond the temporal coverage of the contemporary satellite missions, and the results can be used to analyze regional snow climatology of alpine basins, as well as for extracting information on the seasonal snow cover dynamics. Moreover, information on past snow cover patterns is very valuable for hydrological modeling, as it allows both to better constrain model parameters (Deuthmann, et al. 2014) and a better validation of model results. All these topics seem to us very pertinent to the scope of “The Cryosphere” and we thus believe to deliver a valuable methodological contribution.

My verdict: The paper should be revised and re-reviewed

I have five major review comments that should be addressed and then some minor ones.

1. The title of the paper suggests a generic methodology without caveats on the areas it is designed to be suitable (and areas it has been tested over). I strongly suggest the study be cast in terms of an approach applicable to mountainous areas that are relatively above the tree line , e.g. alpine regions (as I think applies to this study area). This is especially true considering a. the MR statistic used really only identified higher elevation areas, compared to a station, as being snow covered and lower elevation areas as being ’snow free’ - but in areas with different land cover conditions the presence/absence of snow may be related to factors other than elevation

   A: We agree with the reviewer that the suitability of here presented methodology in the regions other than mountainous area is not investigated. We will modify the title of the revised manuscript to:

   “Snow cover reconstruction methodology for mountain regions based on historic in situ observations and recent remote sensing data”

2. The study suggests that because there was a ~85% agreement rate with 4 landsat images the method and modis only provides 92% agreement the method is good. This is misleading. Snow cover mapping outside of melt periods tends to have large areas that are easy to classify and then transition zones that are problematic. For example, figure 9 shows that it is in the transition areas that the step 5 approach was being used for mapping and that this provided most of the uncertainty. If we interpret this conservatively from Table 4 about 50% of the area is mapped using step 5 but this contributes most of the uncertainty. In a sense the accuracy of mapping 50% of the area using step 1-4 is very good (>95%) and for the rest much worse (~70%). This should be noted - basically the approach is good in about 50% of the areas where step 5 is not used but is not competitive to mapping approaches in the other areas.

A: The reviewer raises a very important issue for validation of the presented methodology. Indeed we are aware of the fact that performance during melt/accumulation season is most challenging and selected the Landsat TM scenes accordingly. Figure 1 shows the snow cover area dynamics in 2004 plotted against air temperature at the Pendjikent climate station in the Zerafshan basin. As it is visible from the figure, the Landsat TM scenes for methodology validation were deliberately chosen in the transition period.

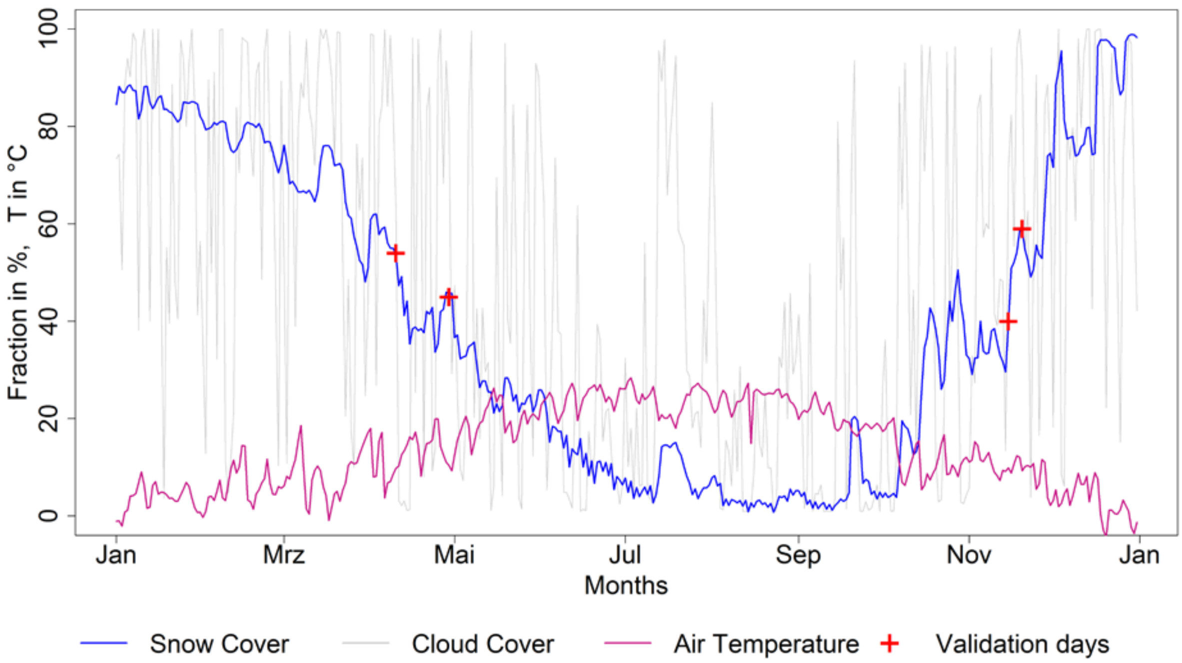

Figure 1. Snow cover area dynamics in the Zerafshan basin derived from MODIS daily snow cover time series for the year 2004. Cloud removal from original MODIS snow cover product

is done using the methodology by Gafurov and Bárdossy (2009). Air temperature is from Pendjikent meteorological station (see Figure 1 in the submitted manuscript).

Figure 2 shows the performance of snow reconstruction in the first 4 steps where accuracy is high. As it is visible from the figure, the fraction of the reconstructed snow cover area in these 4 steps is mostly above 50 %: The reconstructed fraction is reaching nearly 90 % for some summer months and nearly 70 % for winter months. Moreover, and as it was also mentioned earlier, the validation days were deliberately chosen from transition periods, where the reconstruction is expected to be particularly challenging. Although the reconstructed snow cover fraction within the first 4 steps is indeed lowest for the transition period (Figure 2), the total snow cover was reconstructed with an accuracy being as high as 84-86 %. For the time outside the snow melt or snow accumulation period, higher accuracies can be expected since a higher fraction can be reconstructed in the first 4 steps already.

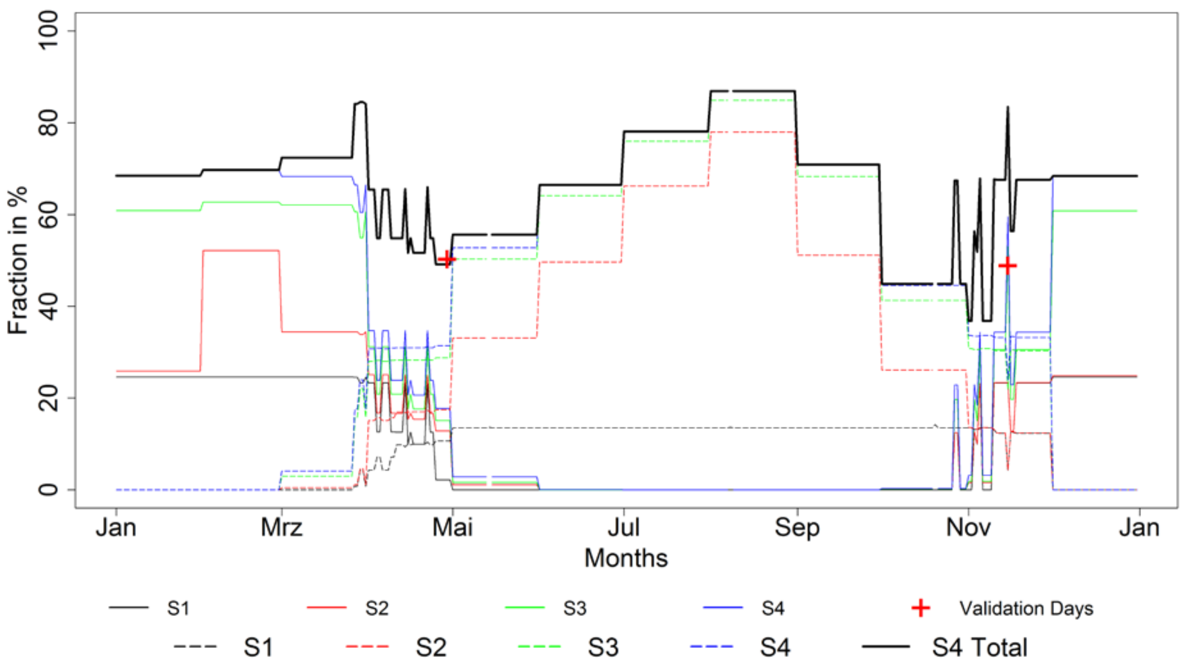

Figure 2. Fraction of the reconstructed snow (continuous lines) and land (dashed lines) cover area for the first 4 steps. “S4 Total” is the total snow and land fraction reconstruction after step 4. “Validation Days” indicate the days for which the snow reconstruction methodology was validated through Landsat cloud free maps. Note that the snow and land reconstruction for validation days only apply to the Landsat area outline in Figure 1b of the original manuscript.

3. There is historical snow cover data from AVHRR at >=1km available. AVHRR imagery can also be used to map snow cover. I understand you may not have it over your region at present but it exists and should be noted as an option.

A: We are aware of AVHRR data availability. In the page 2, lines 10-15 of the manuscript we mention the possibility of obtaining snow cover using AVHRR data. In particular, we also mention the study by Zhou et al. (2013) in which snow cover maps are generated from 1986 onwards. Unfortunately, no freely available AVHRR snow cover product was available for the area of interest at the time of writing. We will, however, stress the availability of AVHRR data for possible validation in the discussion part of the revised manuscript.

4. It is not clear you validated during either melt or onset season when snow cover may be harder to map from MODIS itself and hence your accuracy will be lower. Please find and add a test for each of these seasons.

A: As stated above, the validation days chosen in this study actually are from the melt (April 10 and April 29) and the onset (November 15 and November 20) seasons (see Figure 1 and 2). We will highlight this better in the revised manuscript. Similarly, we will stress that in the winter season, the basin is up to 100 % covered by snow, whilst in the summer season, the basin is covered by less than 5 % of snow (these 5% are from accumulation areas of glaciers within the basin), thus making a reconstruction practically meaningless. Currently, one can see this from Figure 1 where annual snow cover dynamics assessed from MODIS is plotted.

5. If snow cover is ephemeral I figure you will have a lower MR since the station and MODIS maps may be seeing different snow cover. Can you provide an image of the range of MR across the annual period (and more importantly when step 5 is being used to map a region for the worst/best case dates) since when MR<1 the method is far less accurate.

A: The methodology accounts such conditions as follows: Under ephemeral snow cover conditions, station or MODIS data may see different snow cover than the reality. In such cases (e.g. MODIS sees "land" although there is ephemeral snow whilst the station sees "snow" since it is a manual recording with a certain threshold), MR gets the value of < 1 and is not used in the first step of reconstruction. Only distinct snow cover records from both station and MODIS are used to identify snow covered areas in the first step. Reduced MR values that may partly be due to ephemeral snow cover are, however, used in step 5 in order to classify areas still undefined in the steps 1-4. The problem with the ephemeral snow cover partly contributes to the accuracy loss in step 5, and is addressed in the discussion of Figures 9 and 10. We will explicitly mention the problem with ephemeral snow cover in the discussion of step 5 in the revised manuscript.

Minor comments:

1. I would be more comfortable if much of the math was phrased in terms of conditional probabilities rather than new terms such as ’MR’. The ’MR’ is just the conditional probability a modis pixel is snow covered (or snow free) if a station is snow covered (or snow free).

A: Yes, we agree with the reviewer that MR is basically the conditional probability $P(S_p|S_s)$ of e.g. having snow in the pixel given (S_p) snow in the station record(S_s). We will change the notation accordingly.

2. The confidence in the MR will change with the number of cloud free MODIS estimates at a pixel. This should be factored in and modelled using binomial sampling theory.

A: Indeed, the number of cloud free MODIS data at a pixel plays a role in the confidence interval of estimated MR ($P(S_p|S_s)$ in revised manuscript) values. Figure 3 shows 95 % confidence intervals derived from a binomial distribution, exemplarily shown for MR=0.92 against different number of observations (samples). The figure shows that, as the number of observations used to compute MR between station and a pixel increases the variation in confidence interval decreases. See also the answer to the next minor comment where further consideration about confidence interval are presented.

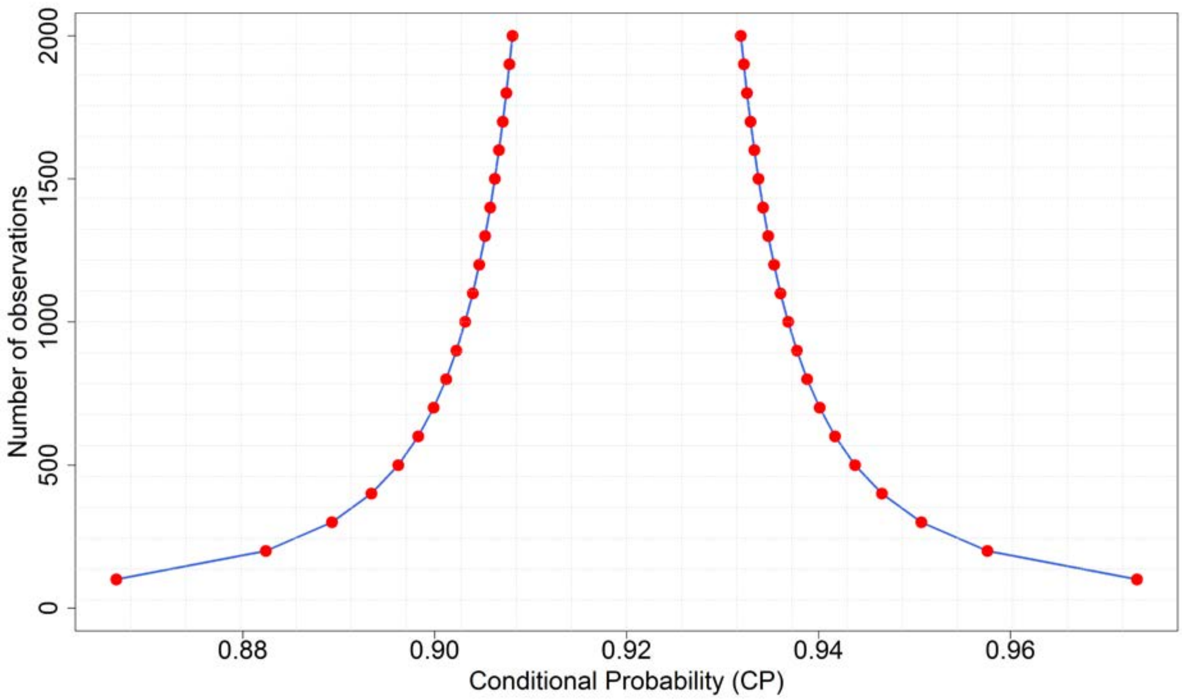

Figure 3. Confidence interval (95 %) derived from binomial distribution against number of observations, exemplarily plotted for $P(S_p|S_s)$=0.92.

3. Since the MR estimate is essentially a binomial probability you can model the confidence interval of MR when it is not equal to 1. For example, say the true probability a modis pixel says 'snow' if a snow station says 'snow' is 0.8. Say then we sample 10 cloud free years over the pixel and find 6 of 10 years say 'snow'. Then the probability the modis grid cell is snow covered based on the station being snow covered can be directly modelled using a binomial distribution.

A: Following the proposal of reviewer we tested snow cover reconstruction in the last step (step 5 when MR is not equal to 1) considering maximum lower bound of the confidence interval of MR value between each pixel and stations. The confidence bounds were estimated assuming the binomial distribution of the MR. They account for the number of observations used to compute MR at individual pixels (Figure 3). The confidence level of 95% was adopted. We compared the accuracy of the snow cover reconstruction between maximum estimate of MR as described in the submitted manuscript and the estimate of the maximum lower confidence bound in step 5 in relation to the Landsat images (Table 1). The results show no significant differences in the accuracy of the reconstruction when accounting for uncertainty related to the number of observations. This insignificant variation in accuracy may be explained by the number of observations used to construct MR values in this study. Figure 4 illustrates the frequency distribution of the number of observations used to compute MR values for all pixels based on the Oigaing station. As it is visible from the figure, the majority of the pixels had more than 500 and more than 1000 cloud free observations for constructing MR for snow and land, respectively. If we consider the frequency distribution of observations in Figure 4 and compare it to the confidence interval in Figure 3 for this case, it can be concluded that significant drop in accuracy is not expected as majority of MR values are computed based on more than 500 observations.
Nevertheless, the frequency distribution also shows that the number of observations used for constructing MR value for snow is smaller than number of observations used to construct MR for land. This is due to the fact that pixels in the Zerafshan basin are most of the time snow-free than snow-covered. Moreover, the probability of pixels being cloud free in summer months is higher than in winter months.
Since the value of number of observations can be important in regions with less observations, we will adjust our methodology considering confidence intervals for MR in the revised manuscript.

Table 1. Accuracy of snow cover reconstruction in (only) step 5 using two methods: “Maximum MR” as presented in the original manuscript and "Maximum lower bound” in which the maximum value of the lower bound of the confidence interval for the MR value was used to reconstruct snow cover. Four cases are distinguished: SS, LL, SL, and LS. The first (second) letter indicates the classification according to the presented algorithm (Landsat). “S” stands for “snow”, “L” for “land”. TRUE is the sum of SS and LL. FALSE is the sum of SL and LS.

Days	Method	SS	LL	SL	LS	TRUE	FALSE
Apr 10, 1998	Maximum MR	44.3	42.1	11.6	2.0	86.4	13.6
	Maximum lower bound	43.4	42.3	12.5	1.8	85.7	14.3
Nov 20, 1998	Maximum MR	18.2	66.6	2.1	13.1	84.8	15.2
	Maximum lower bound	18.1	67.0	2.2	12.7	85.1	14.9
Apr 29, 1999	Maximum MR	24.9	58.7	13.9	2.5	83.6	16.4
	Maximum lower bound	24.5	58.8	14.3	2.3	83.3	16.6
Nov 15, 1999	Maximum MR	42.3	41.7	10.6	5.4	84.0	16.0
	Maximum lower bound	41.7	42.0	11.2	5.1	83.7	16.3

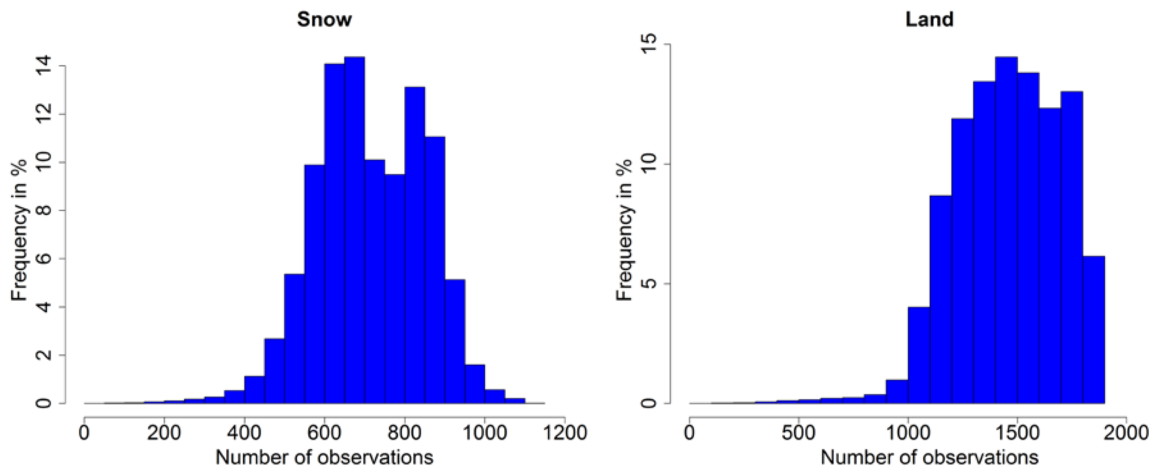

Figure 4. Frequency distribution of number of observations used to construct MR values for snow and land at each pixel and Oigaing station.

4. I am not opposed to using elevation zones and some sort of buffer to make monthly climatologies but I would rather the potential for a shift in snow covered area for a month be directly checked using modis time series.

A: There are two reasons for why we did not check potential variations in snow climatology using MODIS time series: 1) MODIS can currently offer only limited historical time series (~ 14 years to date) which we consider to be insufficient for quantifying snow line changes in the past in a robust manner, b) original MODIS snow cover data are disturbed by cloud cover which makes the analysis uncertain. We consider imposing an elevation buffer based on a long-term independent proxy as a more conservative approach.

5. I found the use of a neighbourhood filter to change snow cover status (test 4) too arbitrary to be useful and potentially dangerous – what about areas with lakes, open areas and forests, hollows where snow gathers etc. I suggest it may work in your region but would definitely mask these areas and use them only with some additional test to verify land cover and topography is relatively constant in the filter window.

A: We agree on reviewer's concern and, indeed, cases as mentioned above can occur. These limitations will be mentioned in the revised version of the manuscript, including additional notes in the methodology section. In the Zerafshan basin, however, big water surfaces for which applying this step could be problematic do not exist. Moreover, the basin is mainly treeless which would exclude also potential misclassification due to forested areas. Information about land cover distribution in the study area will be included in the revised version of the manuscript. Additional application and validation studies in other areas will shed more light on the relevance related to these aspects.

6. Figure 3 should show the study area outline

A: The study are will be outlined in Figure 3 of the revised manuscript.

7. Figure 1 should show the study area in the lower panel as an outline only - we need to see the higher resolution dem.

A: The outline of the Zerafshan basin in the lower panel of figure 1 will be set as transparent for better visibility of the DEM. As the snow cover reconstruction was conducted for the whole domain of the lower panel in figure 1, we think it is important to illustrate this domain as whole. It is correct, however, that the validation was performed only for the area covered by Landsat (outlined currently).

8. I like FIgure 1 it convinces the reader that in your region there is some robustness to your approach but please refer to my point 2 since I feel that the transition zone that covers 50% of the area seems to have most of the errors.

A: See the reply to major comment # 2.

9. The Landsat map is based on the MODIS algorithm as is the MODIS snow cover maps used to calibrate similarity indices - you should caution there could be some potential for bias in your validation due to this fact.

A: We fully agree. This will be mentioned in the revised version of the manuscript

10. Brown 1999 did not use remote sensing imagery to map snow cover.

A: We re-checked the Brown (1999) paper. In the methodology section (page 2344) the author states that high correlation was found between snow index (which is based on interpolated snow cover using station data) for November – April period and NOAA SCE (Snow Cover Extent) values (which are based on NOAA weekly satellite snow cover data as explained in the datasets section (b)). In our manuscript we referred to the NOAA SCE product as "AVHRR snow cover" to be consistent throughout the text where AVHRR data is mentioned few times and is a NOAA family sensor. However, in the revised version of manuscript we will mention NOAA SCE instead of AVHRR to be consistent with original publication as mentioned by Brown (1999).

11. You need to provide more details on the snow depth measurement especially a. if it also records no-snow conditions b. what threshold depth is used for snow c. what do you do if there are trace measurements?

A: More detailed information on observed snow depth data including threshold depth for snow measurement will be included in the revised version of the manuscript.

Reviewer #2:

This manuscript discusses a methodology to reconstruct historical snow coverage using recently available remote sensing data and long-term point observations of snow depth from existing meteorological stations for the Zerafshan River basin in Central Asia. Information on snow cover and snow depth is crucial for seasonal forecasts of water availability and for calibration and validation of hydrological models.
The outcome of this paper is important to better understand the possibility of reconstruct historical snow data. The paper is mathematically rigorous.
While the results of this paper should be interesting, the paper should be in my opinion published but it should be thoroughly revised and extended according to quality requirements imposed by the journal and referees.

General Comments:

1. This methodology is useful to reconstruct historical snow coverage certain region (Zerafshan River basin), but the title of the paper suggests a generic methodology without caveats on the area location.

A: See the reply to "major comment #1" of the first reviewer.

2. I think, much more care is required in presenting information related to the limits of this method for transition seasons and at edges of the snow field.

A: Limitations of the methodology will be extended in the revised version of the manuscript

3. And I agree with comment 5 by R. Fernandes :“If snow cover is ephemeral I figure you will have a lower MR since the station and MODIS maps may be seeing different snow cover. Can you provide an image of the range of MR across the annual period (and more importantly when step 5 is being used to map a region for the worst/best case dates) since when MR<1 the method is far less accurate”

A: See the reply to "minor comment #5" of the first reviewer.

References:

Gafurov, A. and Bárdossy, A.: Cloud removal methodology from MODIS snow cover products, Hydrol. Earth Syst. Sc., 13, 1-13, 2009

Brown, E.: Northern Hemisphere Snow Cover Variability and Change, 1915-97, J. Climate, 13, 2339-2355, 1999

Duethmann, D., Peters, J., Blume, T., Vorogushyn, S. and Güntner, A.: The value of satellite-derived snow cover images for calibrating a hydrological model in snow-dominated catchments in Central Asia, Water Resour. Res., 50, 2002-2021, doi: 10.1002/2013WR014382, 2014