Thank you for providing us valuable suggestions and they do help improve the paper. According to the reviewers' comments, we revised the paper carefully and tried to give satisfactory answers to the reviewers' questions. The corresponding modifications are highlighted with red font in the revised paper.

The summaries of the revision for this paper are as follows:

First, the AVHRR GAC snow dataset have been updated with the final released version, which is based on based on the algorithm SCAMOD (Metsämäki et al. 2015). Furthermore, we extend the description of AVHRR GAC snow retrieval in the revised manuscript.

Second, more in-depth analysis was made regarding the performance of AVHRR GAC snow over different elevation regions (0-200, 200-500, 500-1500, 1500-2500, 2500-3500, 3500-4500, 4500-5500, >5500). Furthermore, the effect of landcover type, slope, aspect, and topographical variability were analyzed for different elevation regions.

Third, the structure of the manuscript has been improved. The accuracy of MODIS based on in situ sites was discussed along with AVHRR in Section 4.1. Furthermore, the comparison between AVHRR GAC and MODIS snow regarding the accuracy and temporal stability is also presented in this section. The comparison between AVHRR GAC snow and MODS snow regarding their absolute values as well as the comparison between AVHRR GAC snow and Landsat snow were presented in Section 4.2 (Comparison based on medium to high resolution data).

Fourth, based on the new results and analyses, more comprehensive conclusions were presented.

Fifth, some details, including figures and grammar, were improved.

For the specific comments for each reviewer, we have made detailed reply as following.

Reviewer1:

General Comments:

Interesting concept to build a long-term snow cover data product from AVHHR GAC. There are many challenges to creating and validating an AVHRR GAC snow cover dataset over the Hindu Kush Himalaya (HKH) region for the period 1982-2013. Description of, and explanation of how the AVHRR GAC snow cover dataset was created are insufficient. There are significant problems regarding how the snow cover dataset was created, and with how the satellite validation datasets were processed and analyzed. A map of the snow cover dataset is not presented. A AVHHR GAC snow map must be shown in support of evaluation/analysis. The quality/accuracy of the snow cover dataset cannot be evaluated. The conclusions are not substantiated by the analysis presented. The manuscript is not suitable for publication.

Re: The manuscript has been revised from the following aspects:

First, the AVHRR GAC snow dataset have been updated with the final released version. Because the final AVHRR GAC snow data published and accessible for everyone is different from what we have previously employed in the paper. Our team have improved the retrieval algorithm, because there was a need to retrieve also snow on ground with an identical procedure as for viewable snow.

The final AVHRR GAC data dataset (openly accessible here https://catalogue.ceda.ac.uk/uuid/5484dc1392bc43c1ace73ba38a22ac56) in the whole time series was based on the algorithm SCAMOD (Metsämäki et al. 2015). Consequently, many results and conclusions have been reworked.

Second, we expand the description of the generation of AVHRR GAC snow cover product in **Section 2.2** and include the snow cover maps for illustration.

Third, we have made more in-depth analysis of the performance of AVHRR GAC snow. In particular, the study area has been divided into eight groups according to their elevations (0-200, 200-500, 500-1500, 1500-2500, 2500-3500, 3500-4500, 4500-5500, >5500) in order to take the topography into consideration. Furthermore, the effect of landcover type, slope, aspect, and topographical variability were analyzed for different elevation regions in **Section 4.2.3** (Pixel-based comparison and potential influential factors on accuracy).

Fourth, the structure of the manuscript has been improved. The accuracy of MODIS based on in situ sites was discussed along with AVHRR in **Section 4.1**. Furthermore, the comparison between AVHRR GAC and MODIS snow regarding the accuracy and temporal stability is also presented in this section. The comparison between AVHRR GAC snow and MODS snow regarding their absolute values as well as the comparison between AVHRR GAC snow and Landsat snow were presented in **Section 4.2** (Comparison based on medium to high resolution data).

Last but not the least, based on the new results and analyses, more comprehensive conclusions were presented in **Section 5**.

Specific Comments:

Line 44: The GCOS does recommend area covered by snow cover as an essential climate variable daily at 1 km or higher resolution. The GAC resolution is much lower than recommended. The lower resolution needs to be discussed relevant to GCOS observation requirements.

Re: On the one hand, GCOS definitely is in need of long, meaning climate-relevant time scales of >30 years, thus this is only covered by AVHRR. On the other hand, although the GAC resolution is lower than recommended, this snow dataset is unique over such a long time scale at daily resolution available.

In order to clarify this point, we have added the sentences as "*……the ESA Climate Change Initiative (CCI) has emphasized the necessity of generating consistent, high quality long-term datasets over the last 30 years as a timely contribution to the ECV databases. In this demand, a global time series of daily fractional snow cover product has been generated from AVHRR GAC data (Naegeli et al., 2021). This snow dataset is unique as it spans 4 decades and thus provides information about an ECV at climate-relevant time scales.*" in the **Introduction**.

Discussion is also lacking on explaining the relevance of coarse GAC resolution product in the HKH region, especially when higher resolution snow cover data sources are available.

Re: We would like to point out the AVHRR GAC snow is a global product for all land areas, excluding Antarctica and Greenland ice sheets. It provides daily products for the period 1982-2019.

This dataset is unique as it spans 4 decades and thus provides information about an ECV at climate-relevant time scales. In fact, this is the best spatial resolution over such a long-time scale at daily resolution available. It is important to note that the evaluation of AVHRR GAC snow over HKH is just a typical representation of its performance over mountainous area.

The HKH was selected as the study area partly because of its particular sensitiveness to climate change and thus reliable daily snow cover data across this area are in great demand, and partly because this area is featured by rich diversity of climates, hydrology, ecology, biology, and topography. Then it provides a favorable condition to explore the influential factors (e.g., elevations, landcover type, slope, aspect, and topographical variability) on the accuracy of snow mapping.

In order to clarify this point, we have emphasized the relevance of HKH and AVHRR GAC snow in **Introduction** as "*The Hindu Kush Himalaya (HKH), which is often called as the freshwater tower of Asia, comprises the highest concentration of snow outside the polar regions. The snow cover of this area plays a crucial role in the water supply of several major Asian rivers (Immerzeel et al, 2009). On the other hand, the HKH area is of special interest due to large area, rich diversity of climates, hydrology, ecology, and biology (Wester et al., 2019). Variations in snow cover affect the precipitation, near-ground air temperature, and summer monsoon in Eurasia and across the Northern Hemisphere (Hao et al., 2018). Given the fact that HKH is particularly sensitive to climate change and thus shows strong interannual variability, reliable daily snow cover data over a long time series across this area are in great demand.*".

Section 2.2: I do not understand how you created a AVHRR GAC snow extent. I read the FCDR product user guide. I downloaded and looked at the FCDR data products. There is no snow cover extent dataset in any of the products. There is no AVHRR channel reflectance data in those products that could be used to calculate NDSI.

Re: In the revised manuscript, we have extended the description of AVHRR GAC snow as "*The AVHRR GAC snow cover extent time series version 1 derived in the frame of the ESA CCI+ Snow project is the most recent long-term global snow cover product available (Naegeli et al., 2021). It covers the period 1982-2019 at a daily temporal and 0.05° spatial resolution. The product is based on the Fundamental Climate Data Record (FCDR) consisting of daily composites of AVHRR GAC data (https://doi.org/10.5676/DWD/ESA_Cloud_cci/AVHRR-PM/V003) produced in the ESA Cloud CCI project (Stengel et al., 2020). The data were pre-processed with an improved geocoding and an inter-channel and inter-sensor calibration using PyGAC (Devasthale et al., 2017). Snow cover extent retrieval method was developed and improved based on the ESA GlobSnow approach described by Metsämäki et al. (2015) and complemented with a pre-classification module. Alongside the daily reflectance and brightness temperature information, an excellent cloud mask including pixel-based uncertainty information is provided (Stengel et al., 2017, 2020). All cloud free pixels are then used for the snow extent mapping, using spectral bands centred at about 630 nm and 1.61 µm (channel 3a or the reflective part of channel 3b), and an emissive band centred at about 10.8 µm. The water bodies, permanent ice bodies and missing values are flagged. SCAmod* *retrieves both the snow cover on top of the canopy as well as on ground below the canopy by taking the canopy density into account. Here, we focus on the latter variable as this is most suitable for the comparison with in situ stations.*

To reduce the effect of cloud coverage, a temporal filter of ±3 days of each individual snow cover observation was applied after Foppa and Seiz (2012). The AVHRR GAC FCDR snow cover product comprises only one longer data gap of 92 days between November 1994 and January 1995 resulting in a 99 % data coverage over the entire study period of 38 years. In this study, we will focus on the evaluation of raw daily retrieval of AVHRR GAC snow extent (denoted by "AVHRR_Raw") since additional uncertainty will be introduced with the gap-filling process." in **Section 2.2.**

The fractional snow method of Salomonson and Appel (2006) is based on a using Landsat TM data at 30 m to estimate the fractional amount of snow in a MODIS 500 m pixels. That relationship is not directly applicable to GAC data at 1x5 km^2resolution. Please explain why you applied this regression to GAC data to estimate FSC.

Re: As we explained before, the final AVHRR GAC snow adopted by the revised manuscript is different from what we have previously employed. Our team have improved the retrieval algorithm, because there was a need to retrieve also snow on ground with an identical procedure as for viewable snow. The final algorithm was developed and improved based on the ESA GlobSnow approach described by Metsämäki et al. (2015) and complemented with a pre-classification module. The detailed description of this method can be seen in Section 2.2 (AVHRR GAC snow extent retrieval) in the revised manuscript.

NDSI cannot be calculated from AVHHR data until 1998 when with NOAA-15 Channel 3A at 1.6 um was added to AVHRR. Before then there was no shortwave infrared channel covering 1.6 um. It is not possible to calculate NDSI from AVHRR data prior to 1998. But your dataset record is 1982-2013. How can it possibly be consistent across major design changes in the AVHRR?

Re: The old retrieval algorithm of snow cover extent considers the high reflectance of snow in the visible spectra and the low reflectance in the short-wave infrared ($SWIR$). In order to construct the long-term AVHRR snow extent dataset, the $SWIR$ reflectance values stem either from **channel 3a** centred around 1.61 μm or from the **reflective part of channel 3b** centred around 3.75 μm (Roger et al., 1993). We derived the **reflective part of channel 3b (ref3b)** using the method proposed by Baum (1999):

$$ref3b = \frac{I_3(measured) - I_3(BT_4)}{\frac{1}{\pi}(F_3\mu_0) - I_3(BT_4)} \quad (2)$$

Where $I_3(measured)$ is the measured radiance (mW m^{-2} sr^{-1} cm) for channel 3b, BT_4 is the 11 μm brightness temperature, F_3 is the integrated solar spectral irradiance (mW m^{-2} cm) weighted by the spectral response function for channel 3b, and μ_0 is the cosine of the solar zenith angle. To account for changing solar zenith (sunz) angles the Top of Atmosphere (TOA) reflectance was corrected by a division of the cosine(sunz).

Reference:

Baum, B. A. (1999). A grouped threshold approach for scene identification in avhrr imagery. Journal of Atmospheric & Oceanic Technology, 16(6).

Roger, J. C. & Vermote, E. F. (1993). A method to retrieve the reflectivity signature at 3.75 µm from AVHRR data. Remote Sens. Environ. 64, 103–114.

The calculation of NDSI from AVHRR GAC is not explicitly given. Without explanation of how NDSI was calculated the entirety of the validation discuss and results is doubtful.

Re: The Normalised Difference Snow Index (NDSI) were determined by the high reflectance in the visible spectra and the low reflectance in the short-wave infrared:

$$NDSI = \frac{visRED - SWIR}{visRED + SWIR} \quad (1)$$

In the case of here used AVHRR GAC data, $visRED$ is represented by channel 1 centred around 0.63 µm. The $SWIR$ reflectance values stem either from **channel 3a** centred around 1.61 µm or **from the reflective part of channel 3b** centred around 3.75 µm (Roger et al., 1993). We derived the reflective part of **channel 3b (ref3b)** using the method proposed by Baum (1999):

$$ref3b = \frac{I_3(measured) - I_3(BT_4)}{\frac{1}{\pi}(F_3\mu_0) - I_3(BT_4)} \quad (2)$$

where $I_3(measured)$ is the measured radiance (mW m^{-2} sr^{-1} cm) for channel 3b, BT_4 is the 11 µm brightness temperature, F_3 is the integrated solar spectral irradiance (mW m^{-2} cm) weighted by the spectral response function for channel 3b, and μ_0 is the cosine of the solar zenith angle. To account for changing solar zenith (sunz) angles the Top of Atmosphere (TOA) reflectance was corrected by a division of the cosine(sunz).

Section 2.3.2: The estimate of FSC for TM/ETM data is flawed. Salomonson and Appel (2006) did not derive the FSC regression to estimate TM/ETM FSC, they derived a FSC estimate for MODIS data based on the higher spatial resolution TM/ETM data. That regression is not appropriate to estimate FSC in TM/ETM data.

Re: Although the method by Salomonson and Appel (2006) is originally designed for MODIS FSC products with a mean absolute error of less than 10% (Salomonson and Appel, 2004), we assumed that such an accuracy can be achieved with higher resolution data in this paper. This treatment follows the recommendations of (Metsamaki et al., 2015), which applied fractional snow method by Salomonson and Appel (2006) to Landsat data for the evaluation of coarse-pixel snow extent products.

In order to clarify this point, we have added the sentences as "*Following the recommendation of Metsamaki et al. (2015), the fractional snow method by Salomonson and Appel (2006) was employed to generate reference FSC from Landsat TM/ETM imagery. This method is originally designed for MODIS FSC products, with a mean absolute error of less than 10% (Salomonson and Appel, 2004). In this paper, we assumed that such an accuracy can be achieved with higher resolution data.*" in **Section 2.3.2** in the revised manuscript.

Reference:

Metsamaki, S., Pulliainen, J., Luojus, &K., et al.: Introduction to globsnow snow extent products with considerations for accuracy assessment, Remote Sens. Environ., 156, pp. 96-108, 2015.

No AVHRR GAC snow map is shown! No snow maps are shown! The research describes building an AVHRR GAC snow cover map, but none is shown. Visual evidence of the snow maps must be presented.

Re: We have included the snow cover maps for illustration in **Section 2.2** in the revised manuscript.

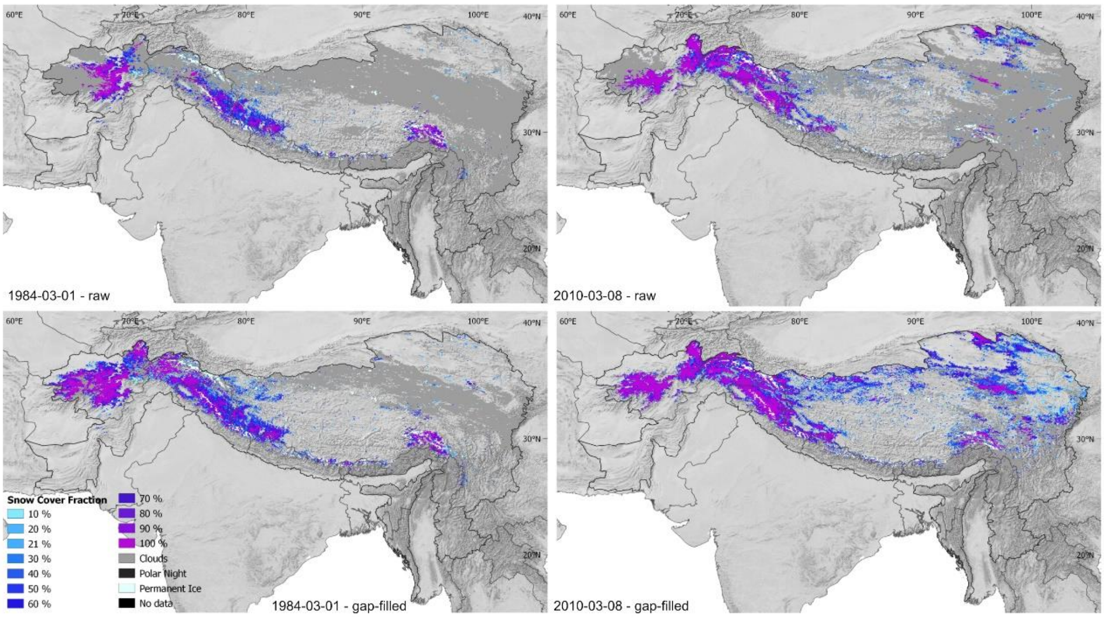

Figure 2. AVHRR GAC raw (top row) and gap-filled (bottom row) snow cover for entire HKH in March 1984 (left) and March 2010 (right).

Line 145: It is unclear how the 30 m FSC data was used to identify FSC in a GAC pixel. Were many 30 m FSC pixels mapped into a GAC pixel then was some spatial averaging or binning of observations done to estimate FSC in a GAC pixel? A clear explanation of sampling and estimate of FSC in GAC pixel is needed.

Re: We would like to thank the reviewer for providing such a valuable comment. The resampling method was area-weighted average of contributing pixels. This point has been clarified in **Section 2.3.2** as "*These high-resolution data were then projected to a geographic projection and aggregated to AVHRR GAC pixel scale using the area-weighted average of contributing pixels to 'simulate' the reference FSC estimates at the AVHRR GAC pixel scale.*" in the revised manuscript.

Line 248: How were MOD10A1 data resampled and reprojected to a GAC pixel? Many MOD10A1 pixels can be covered by a GAC pixel. What method was used to resample? Was there any averaging or compositing observations to GAC pixel done?

Re: The resampling method was area-weighted average of contributing pixels. In order to clarify this point, we have added one sentence as "*In order to avoid spatial scale mismatch between AVHRR and MODIS pixel, MOD10A1 was reprojected to a geographic projection and aggregated* *to AVHRR GAC pixel scale using the area-weighted average of contributing pixels.*" in **Section 2.3.3** in the revised manuscript.

Because of the great amount of uncertainty regarding how the AVHHR GAC snow cover dataset was created and uncertainty of processing the validation datasets, only a cursory review of the methods and results was done.

Re: In the revised manuscript, we have extended the description of AVHHR GAC snow cover retrieval in **Section 2.2**. Additionally, the processing the validation datasets have been enhanced in **Section 2.3.2** (Landsat TM/ETM data and processing) and **Section 2.3.3** (MODIS snow cover product).

The conclusion that this AVHRR GAC dataset has consistent performance across the whole suite of AVHRR sensors is not substantiated.

Re: Since the AVHRR GAC snow dataset have been updated with the final released version. many results and conclusions have been reworked. Furthermore, the performance of AVHRR GAC snow in different months were analyzed in detail.

Based on the new results, it was found that the sensor-to-sensor consistency was found to differ slightly and unsystematically in ACC and Bias throughout the time series. While the consistent slight increasing trend on HSS is noteworthy. Particularly, the performance of AVHRR GAC snow is the worst on January and the best on October regarding the magnitude of ACC and HSS. But when temporal stability of accuracy was considered, it performs the best on November and the worst on January and December regarding the ACC. While that of HSS is the worst on December. The results of Bias provide different perspectives for the performance of AVHRR GAC snow. It generally overestimates snow on February, March, October, and November. By contrast, unbiased estimation is likely to occur on December and January.