Reply to reviewer#2

Response: We thank David Shean for providing a thorough and constructive review of our manuscript. Please find below our preliminary answers to the main comments. The response to the specific/minor comments will be detailed in our final response letter if the editor considers that our manuscript is appropriate for The Cryosphere.

Summary (including already major comments)

One of my main questions is whether TC is really the best journal for this paper. This work is relevant to the glaciology and high-mountain cryosphere/hazard communities. But much of the focus is on the data products, including methods and detailed cal/val. There are limited "cryosphere science" results in the paper, which is not a complaint, as I value data/methods papers, but it makes me wonder whether another journal (ESSD?) might be more appropriate?

Response: Our rationale to target The Cryosphere is that our dataset is specifically designed for and of interest to the glaciological community. We note that the REMA paper by Howat et al. (https://doi.org/10.5194/tc-13-665-2019) had a similar scope (i.e. "data products, including methods and detailed cal/val"), did not include glaciological findings and was published also in The Cryosphere. It is actually one of the most cited articles in The Cryosphere. Our strong preference is to keep our manuscript in The Cryosphere should the editor decide it is publishable and within scope.

For an operational observation/monitoring program like this, it is essential to use version-controlled processing and data products for provenance and understanding of potential issues and evolution of the methods (and associated products). Ideally, low-level image products (Level-1B) products would be available, low-level DEM (uncorrected) products, in addition to high-level, co-registered and corrected DEM and difference products. I also wonder about the evolution of the proprietary processing routines used by Airbus to prepare the delivered products over the years (there should at least be version numbers in the metadata that can be tracked) - presumably the current Level-0 to Level-1B processor is better than what was used 10 years ago.

Response: As described in our data availability statement, we are not allowed to redistribute low-level Pléiades imagery. This should change by the end of 2024 (or rather sometimes in 2025 to be more realistic) with a program named the Pléiades World heritage (PWH, described very briefly here: https://dinamis.data-terra.org/acces-aux-donnees/) thanks to which the full Pléiades archive will be opened to the academic sector (details of licensing to be confirmed). This will satisfy the request of expert users, who might wish to process low-level data. We stress that the PGO was designed to satisfy the need of fellow glaciologists directly interested in the coregistered DEMs. We note that the coregistration of the PGO DEMs to GLO-30 is a simple translation so a user could, if needed, easily obtain the uncorrected DEMs.

Unfortunately, I was unable to access the "open" data products during my review! I followed the instructions prepared by the authors, which state "this is when you have to wait a delay (up to 7 days!) for your account to be active." After 3-4 weeks, I still did not have them appear in my account. The different platforms are a bit confusing to navigate (though not suggesting current NASA distribution is necessarily easier). But this "custom approval" approach is really not conducive to supporting new users or easy "on-demand" data access.

Response: It was unfortunate that the reviewer could not access these data in an easy manner. Due to licensing constraints, we did not have a lot of freedom to choose our

distribution platform. We apologize to the referee and would have gladly shown the reviewer how to solve this data access issue. This lag of several days between the initial request and the final access should be solved by the end of year 2024. Two "sample" products (in the Alps and Tibetan Plateau) are available until 7 July 2024 here: https://filesender.renater.fr/?s=download&token=50321ce9-ab69-4ddd-a5e2-56dbb4e0de36

One suggestion - I understand the budget constraints and option to define custom polygons around glaciers of interest for delivery of clipped Level-1B image products, instead of the full Level-1B images (full sensor swath). The "extra" pixels for full-width images, and resulting DEM coverage, would provide a much larger sample of "stable" terrain, which would improve co-registration and correction beyond the relatively narrow buffer of the clipped products around priority glaciers. Perhaps an arrangement can be made with Airbus to deliver the full Level-1B images for internal processing and correction, but then PGO would only release the clipped DEMs.

Response: Budget constraints were unavoidable, and we preferred to cover more glaciers than order entire images with a lot of off-glacier terrain. In the long term, it will be one of the added values of the PGO to have "forced" the Pléiades satellites to acquire imagery over glaciers. These full level-1B images are now stored in the image catalogue and will be available, at some point (see Pléiades World Heritage program). This was important because Airbus does not aim at building an archive: they acquire images on demand from the customers (and commercial requests are rarely on glaciers).

Major comments (including also some specific comments)

limited value of tri-stereo

Response: We are unaware of studies that show tri-stereo yield clearly superior results on natural terrain (not urban), but we would be happy to include those and modify the statement if the referee knows of studies that demonstrate this.

"before" and "after" co-registration metrics

Response: These metrics are distributed with the data product (that the reviewer could unfortunately not access). See sample products now available on a FTP server.

2m vs 20 m coregistration

Response: We will clarify the text (and possible change Figure 6) to make sure there is no confusion

I recommend more careful documentation of which products were used when co-registering a PGO DEM (2 m) to GLO-30 (30 m) vs. comparing two PGO DEMs (both 2 m), and how this might impact results.

Response: This general comment was not clear to us. Maybe one of the specific comments will help us address this? We will make sure we clarify the specification of the GLO-30 DEM and the workflow for DEM coregistration.

5th-order polynomial is needed for the cross-track direction.

Response: It is true that the across track biases are generally smaller than along track biases. In our full response (if the editor agrees our article is suitable to TC), we will show several examples where the spatially-structured biases should be corrected. A fifth-order polynomial performed suitably well to remove bias. We note that the corrections are all graphically (.jpg or .png) illustrated in the data products downloaded by the users. In case

those generic corrections are not satisfying (and this is indeed sometimes the case), the user can easily come back to the DSM.

mosaic/composite created for the 2+ DEMs over a single site for a given campaign/year?

Response: We prefer not to create mosaics at the cost of distributing many products. The Pléiades DSM are sometimes acquired several months apart during a period of strong melt (and hence elevation changes) so we prefer to leave it to the users to make the mosaic and maybe apply a seasonal correction.

Are the observed horizontal CE90 values within the vendor spec of 8.5 m and the stated GLO30 horizontal geolocation accuracy (~4 m if I recall correctly)? I believe the 8.5m CE90 is for off-nadir angles up to 30°.

Response: We did not compute CE90 but will check this during revision.

The systematic +4.5 m northward geolocation error is puzzling AND systematic 2.4 m vertical error

Response: We agree. We will check with CNES and Airbus DS to see if they made similar observations while following the geometric performance of the Pleiades satellites. We hope they will be able to share their knowledge for a satellite mission which is ageing and hence not their priority. Our plan is to continue to explore the source of these biases during the revision of our manuscript.

A 30 m threshold for "unreliable" sounds high for image geolocation spec of 8.5 m CE90 combined with GLO-30 geolocation spec of ~4 m

Response: We will share examples of the dem_align.py outputs where shifts of 15-20 m were found and did not result from lack of stable terrain or unwrapping errors in GLO-30. The observation of such shifts justifies the use of a 30-m threshold. Maybe the geolocation performances of Pléiades or GLO-30 cited by the reviewer do not hold everywhere on Earth? And a CE90 value of 8.5 m also leaves room for a few cases for which the horizontal offset will reach two to three times this amount.

Ideally, one would mask the GLO-30 products before co-registraiton using the various AUX products containing error estimates and flags for artifacts, though I'm not sure about AUX layer availability for latest GLO-30 products

Response: There are indeed some quality layers for GLO-30 products, in particular a Heigh Error Mask. However, as stated in the documentation "these are random errors and do not include any kind of systematic errors, such as elevation offsets related to erroneous orbital parameters." So the main artefacts in the GLO-30 DEM (from phase unwrapping errors) are not flagged in the auxiliary data. We think that the DEM coregistration (based on a large number of pixels) is sufficiently robust to outliers so that the results should not be affected by these pixels with lower confidence. We also note that coregistering again the entire database of DEMs is a major endeavour and justified only if significant improvements are achieved.

Earlier, the authors suggest users should use their judgment when evaluating the corrections (including co-registration and systematic bias removal), but here it sounds like there are some automated checks to apply or not apply the co-registration? Or were these checks performed manually?

Response: There were manual checks. This is also why we provide some plots (including the output of dem_align.py) so that the user knows the corrections that were applied and how they impacted the residuals off glacier. We propose to include a few such plots in the appendix of our article.

I was pretty confused when I started reading about PGO mass balance from 2000-2009, since the PGO DEMs don't start until 2016-2017. I suggest more clearly stating that you are not comparing mb measurements from PGO DEMs, but simply re-aggregating Hugonnet et al. results for a subset of glaciers with PGO coverage.

Response: We acknowledge that this section was confusing and will be reworked.

An area coverage threshold of 50% was used to identify PGO glaciers. How was the mean mass balance computed for these glaciers? Hypsometric interpolation? Or just a simple mean for areas with valid coverage, which could be biased if 50% is only over ablation area or only over accumulation area. What happens if you only consider glaciers with 95% more coverage? Do you get the same result?

Response: We only use the 50% coverage to decide if a glacier is included or not in our subsequent analysis. Then no interpolation has to be performed by us because the mass balance data comes from Hugonnet et al. and is not recalculated from the elevation change maps (neither from PGO nor from Hugonnet et al. dh maps). Our revisions will ensure this point is clear to the reader.

The question of "Are PGO sites representative of the Earth's glaciers?" is important, but a bit subjective. They could be, but you can also get "the right answer for the wrong reasons" using an independent dataset. I think the more pertinent question for this effort might be "Are PGO mass balance measurements accurate?" I realize that many other papers have demonstrated Pleiades DEMs offer accurate glacier mass balance measurements, but it would have been nice to see some mass balance computed using the available repeat PGO DEMs compared with Hugonnet mass balance from the same period.

Response: The question "Are PGO mass balance measurements accurate?" is partly answered by our analysis of the uncertainty of the elevation change maps using Lidar or "stable" terrain. Of course, one would have to properly propagate these elevation change uncertainties to take into account other sources of errors from the density conversion factor and the outlines. We are reluctant to use the Hugonnet et al. elevation change maps to evaluate the PGO ones because (i) the time periods differ and (ii) the precision of the PGO dh maps are in most cases superior. See for example the comparison of the elevation change maps on Meighen Ice Cap (Figure 4 in https://doi.org/10.1088/1361-6633/acaf8e) and evaluation of the glacier-wide mean elevation change for several glaciers in Norway (Fig. 15 in https://doi.org/10.1017/aog.2023.70)

What percentage of the PGO glaciers already have 5-year repeat coverage?

Response: As this percentage is evolving every semester, we do not think this is really useful for the article itself.

The Figure A3 of Berthier et al., in press was cited but not available during review.

Response: The figure A3 is available: https://doi.org/10.1017/jog.2023.100