natureresearch

Peer Review File

Manuscript Title: Accelerated global glacier mass loss in the early twenty-first century

Editorial Notes: *none*

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referee #1 (Remarks to the Author):

A] The manuscript details a globally comprehensive description of the state of mountain glaciers and ice caps that leverages use of the ASTER archive, REMA, ArcticDEM, TanDEM-X and other sources.

Its significant results are:

1) it is the first global high resolution survey with consistent methodology to provide estimates of mountain glacier and ice cap change for all glaciers on earth.

2) it shows that glacier mass wastage has accelerated unequally on regional scales

3) it provides up to date estimates of sea level contributions from glacier regions

4) it confirms the presence of the influence of several climate oscillations on regional glacier mass balance

B] The significance of the work is high and is eagerly awaited by much of the glaciological and climate community. Most of the methods are relatively standard but have been scaled to an extraordinary degree, a mammoth task. The uncertainty calculations are complicated but seem to cover all the bases. I recommend a statistician to look them over. The paper somewhat surprisingly lacks citations or acknowledgments of previous studies and methods that it has built upon to set the stage for the work and conclusions - this can be easily remedied.

C] The data is vast and hopefully will be available to the community soon. The approach is valid and builds upon previous studies. The diagrams could do with a little work (see below).

D] I got a little lost with the plethora of steps for the uncertainty calculations, and don't feel I can fully comment on whether they are valid or not. I am surprised that there isn't a little more work on ocean/ice processes when considering marine terminating glaciers. Air temperatures and precipitation are treated well, but the paper would be improved if the effect of ocean and lake melt were commented upon.

E] The conclusions are robust, but I suggest should be more expansive. The last paragraph of the main text is weak, especially when several national space agencies are contemplating and proposing instrumentation with better capabilities than ASTER. The paper should focus more upon its findings and implications.

F] Please be careful to make sure you are clarify and make your text a little clearer - this is purely "word smithing," for example, in the abstract you present that mass loss has increased by 97 Gt per year (~0.3mm sea level) over the last 20 years - it would be better to state that IN TOTAL mass loss has increased by....

Extended tables 1 and 2 are goldmines, but are very hard to read when printed out. Maybe include them as csv files or the like?

Please cite work by Bevis et al on the previously identified North Atlantic Anomaly and the well known "pause" in mass loss for Greenland around 2013.

The last part of the abstract is messy and the first paragraph of the paper (the rationale) could be simplified and made clearer.

Claims that ASTER is under utilized should be backed up. As should claims that most large scale data sets have been used.

I suggest that the variation of the density chosen for the ice and firn be varied slightly to get an idea of the uncertainty in the density that could be a further bound on uncertainties.

Line 106 - you quote a total but provide a rate leaving the reader to search for the equivalent amount for glaciers vs Greenland vs Antarctica. Indeed the total amount seems to be only mentioned in the diagram (5325 Gt total)

I encourage you to think about the conclusions section of the paper more clearly - there are several means of collecting stereo imagery with national agencies proposing new missions. You show agreement with GRACE and ICESat in many situations so a more compelling argument for your findings and the granularity you observe is warranted in the conclusions.

G] Your references are robust, but I think there are some more places where citations (such as the previously mentioned paper on the north Atlantic anomaly) could be improved upon. There is work on the Russian Arctic, Dave Sheans work on HIMA, available at NSIDC, etc that has not been cited and provides context for your numbers.

H] I think the main work for improving clarity on the paper should focus on the introduction and conclusions. FIG1 should have the total mass loss per circle adjacent to each circle. This would aid in things like sea level fingerprinting studies.

In figure 2 large surge events in Svalbard and the Russian Arctic do not seem to be reflected in the diagram, which is surprising.

Overall, I suggest this paper is appropriate for a nature journal and only requires some wordsmithing and the opportunity to punch up its conclusions for higher impact.

Referee #2 (Remarks to the Author):

Review of ‘Accelerated global glacier mass loss in the early twenty-first century’

In this paper, the authors use satellite stereo imaging to estimate glacier mass loss over 2000-2019. This new approach complements other estimates from models, satellite gravimetry, altimetry, and in-situ measurements and the estimated mass loss falls largely in line with these other estimates. This study confirms that glacier melt forms a substantial contribution to sea-level rise since 2000. As I am more of a sea-level than a glacier remote sensing expert, I’ll limit my review to the sea-level related parts of this manuscript. The new method is interesting and the high spatial resolution could be useful to determine the regional sea-level response to glacier mass loss. The method also allows for a more accurate distinction between mass loss on the ice sheets and the peripheral glaciers. However, I think the manuscript can benefit from some extra analyses, and I urge the authors to carefully check all comparisons with ice-sheet mass balances.

The study could profit from a more thorough comparison to other mass loss estimates, for example from GRACE observations and models. Currently, Figure 3 shows a comparison with some estimates, but recent studies like Ciraci et al. (2020) and Bamber et al. (2018) are not included. Does this study point at a significant departure of global-mean mass loss from these previous estimates? From a quick scan, it seems that the global mass-loss estimates from these studies are all in line with the results discussed here. I also wonder how the observed mass loss compares to the changes on longer time scales, for example from Parkes and Marzeion (2018) and Zemp et al. (2019): are the accelerations and trends you see unique and unprecedented over the observational record, or are they within the range of variability that has occurred before?

Can the authors provide data sets that are usable for computing the impact of glacier mass loss on global and regional sea-level? Global studies like WCRP Global Sea Level Budget Group (2018) or regional studies like Rietbroek et al. (2016) require time series of mass loss aggregated by RGI glacier region, or even better, on a grid. The latter can be directly used to compute the resulting spatial sea-level rise patterns due to glacier mass loss (see for example Adhikari et al. 2019).

I also urge the authors to thoroughly check all comparisons between glacier and ice mass loss. Trends and accelerations are very sensitive to the chosen start and end points of the time series. Furthermore, there's likely temporal autocorrelation present in the time series, which is not included in the trend and acceleration uncertainty calculations as far as I can see. Therefore, I am reluctant to accept the statements where trends and accelerations between glaciers and ice sheets are compared. These comparisons beyond the level of 'similar in magnitude' should either be removed, or substantiated by comparing all time series over the exact same period with a full analysis of the error budget. See the line-by-line comments below for some examples.

Line-by-line comments

L23-24: "Yet, due to the scarcity of homogeneous mass loss observations, their recent evolution is only partially known as a geographic and temporal patchwork."
I think this statement is a bit too gross. How about GRACE and GRACE-FO observations? GRACE/FO provides direct mass loss observations since 2002, and there are many approaches available to determine global and regional glacier mass from these observations, for example Reager et al. 2015, Bamber et al. 2018, Ciraci et al. 2020, and Wouters et al. 2019.

L29-30: "equivalent to 24 ± 5% of observed sea-level rise."
Because of the relatively short time series and the presence of serious serial correlation in sea-level time series, estimates trends and accelerations in sea level are highly sensitive to the exact beginning and end period, and the uncertainty in the trend in altimetry-derived sea level is still not negligible (Ablain et al. 2019). Since the sea-level time series from the quoted source only covers 1993-2017, I wonder whether the trends have been computed over the same interval, and how this uncertainty range has been determined. As an alternative, global-mean sea level time series up to present can be found from traditional data centers, such as AVISO, CSIRO, or NASA.

L32-33: "Collectively, glaciers presently lose more mass, and at more accelerated rates, than the Greenland or Antarctic ice sheets."
I agree with this statement for Antarctica, but for Greenland I don't. Over 2002-present, GRACE observations suggest Greenland and glaciers to be of similar magnitude (Ciraci et al. 2020, Velicogna et al. 2020, Frederikse et al. 2019). Unless the authors can somehow prove a significant difference between glaciers and Greenland, I think this statement should be reframed as 'larger than Antractica, similar to Greenland', or something along this line.

L44 ‘tide lines’ replace with ‘high-tide level’.

L48: “largest estimated contributor to current sea-level rise after thermal expansion”
This statement is again very sensitive to the time period it refers to. Do the authors refer to 1993-2016 here?

L119-L121: “We thereby infer”
To which period do the authors refer here? Sea level has been accelerating since the 1960s, but the underlying causes vary substantially with time (e.g. Dangendorf et al. 2019).

L223-225: Where does this factor 10 come from? From IPCC SROCC table 4.1? That table suggests an uncertainty on the order of 0.1-0.2 mm/yr, translating into 30-70 gt/yr, which is definitely not a factor 10 larger than the numbers presented here.

L225-L227: “We distinctly constrain the trend of glaciers towards larger mass losses and their contribution to sea-level rise and its acceleration…”:
Larger mass loss than what? Previous estimates? The numbers seem to be in line with most other estimates (Figure 3b, Ciraci et al. 2020, Zemp et al. 2018)

L227-L231: “no relaxation is in sight for the globally accelerated mass loss rates of Earth’s glaciers”
I’m not a glacier expert, but how can you draw this conclusion from such short time series? How have you separated internal variability from forced melt? For example, model estimates show mass loss rates in the 1930s to be much larger than what we observe today (Parkes and Marzeion, 2018), which came down a lot in the 1950s. In the next paragraph, this problem is directly discussed, which makes me wonder how to interpret this statement.

References

Ablain, M., Meyssignac, B., Zawadzki, L., Jugier, R., Ribes, A., Spada, G., Benveniste, J., Cazenave, A., & Picot, N. (2019). Uncertainty in satellite estimates of global mean sea-level changes, trend and acceleration. Earth System Science Data, 11(3), 1189–1202. https://doi.org/10.5194/essd-11-1189-2019

Adhikari, S., Ivins, E. R., Frederikse, T., Landerer, F. W., & Caron, L. (2019). Sea-level fingerprints emergent from GRACE mission data. Earth System Science Data, 11(2), 629–646. https://doi.org/10.5194/essd-11-629-2019

Bamber, J. L., Westaway, R. M., Marzeion, B., & Wouters, B. (2018). The land ice contribution to sea level during the satellite era. Environmental Research Letters, 13(6), 063008. https://doi.org/10.1088/1748-9326/aac2f0

Ciracì, E., Velicogna, I., & Swenson, S. (2020). Continuity of the Mass Loss of the World’s Glaciers and Ice Caps From the GRACE and GRACE Follow-On Missions. Geophysical Research Letters, 47(9). https://doi.org/10.1029/2019GL086926

Dangendorf, S., Hay, C., Calafat, F. M., Marcos, M., Piecuch, C. G., Berk, K., & Jensen, J. (2019). Persistent acceleration in global sea-level rise since the 1960s. Nature Climate Change. https://doi.org/10.1038/s41558-019-0531-8

Frederikse, T., Landerer, F. W., & Caron, L. (2019). The imprints of contemporary mass redistribution on local sea level and vertical land motion observations. Solid Earth, 10(6), 1971–1987. https://doi.org/10.5194/se-10-1971-2019
Parkes, D., & Marzeion, B. (2018). Twentieth-century contribution to sea-level rise from uncharted glaciers. Nature, 563(7732), 551–554. https://doi.org/10.1038/s41586-018-0687-9

Parkes, D., & Marzeion, B. (2018). Twentieth-century contribution to sea-level rise from uncharted glaciers. Nature, 563(7732), 551–554. https://doi.org/10.1038/s41586-018-0687-9

Reager, J. T., Gardner, A. S., Famiglietti, J. S., Wiese, D. N., Eicker, A., & Lo, M.-H. (2016). A decade of sea level rise slowed by climate-driven hydrology. Science, 351(6274), 699–703. https://doi.org/10.1126/science.aad8386

Rietbroek, R., Brunnabend, S.-E., Kusche, J., Schröter, J., & Dahle, C. (2016). Revisiting the contemporary sea-level budget on global and regional scales. Proceedings of the National Academy of Sciences, 113(6), 1504–1509. https://doi.org/10.1073/pnas.1519132113

Velicogna, I., Mohajerani, Y., A, G., Landerer, F., Mouginot, J., Noel, B., Rignot, E., Sutterley, T., Broeke, M., Wessem, M., & Wiese, D. (2020). Continuity of Ice Sheet Mass Loss in Greenland and Antarctica From the GRACE and GRACE Follow-On Missions. Geophysical Research Letters, 47(8). https://doi.org/10.1029/2020GL087291

WCRP Global Sea Level Budget Group. (2018). Global sea-level budget 1993–present. Earth System Science Data, 10(3), 1551–1590. https://doi.org/10.5194/essd-10-1551-2018

Wouters, B., Gardner, A. S., & Moholdt, G. (2019). Global Glacier Mass Loss During the GRACE Satellite Mission (2002-2016). Frontiers in Earth Science, 7, 96. https://doi.org/10.3389/feart.2019.00096

Zemp, M., Huss, M., Thibert, E., Eckert, N., McNabb, R., Huber, J., Barandun, M., Machguth, H., Nussbaumer, S. U., Gärtner-Roer, I., Thomson, L., Paul, F., Maussion, F., Kutuzov, S., & Cogley, J. G. (2019). Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016. Nature, 568(7752), 382–386. https://doi.org/10.1038/s41586-019-1071-0

Referee #3 (Remarks to the Author):

The manuscript presents a new, detailed reconstruction of global glacier mass changes by combining a massive set of DEMs, "covering 20 times the land area of the Earth", derived from stereo satellite imagery and coregistered to the TanDEM-X global DEM. The resulting three-dimensional array of elevation time series has 100 by 100-meter resolution. Using a Gaussian Process (GP) regression, a continuous function of elevation change is estimated with a combination of kernels accounting for interannual and seasonal variations. The new reconstruction derives mass changes for the 19 regions of the Randolph Glacier Inventory (RGI) with smaller errors than previous results. The high-temporal resolution of the elevation change reconstruction allows the calculation of mass changes for shorter periods within the study's 19-year span. The results are presented in several ways, including plots and tables of total mass loss divided into mass loss estimates from land-terminating and marine glaciers, as well as five-year mass loss, and mean elevation-change rate time series for each of the 19 regions (Figure 1 and Extended Table 1). The comprehensive results are expected to serve as an essential observational baseline for improving model results and driving policy decisions.

I acknowledge the innovative nature of the new reconstruction and the critically important conclusions derived from the results about the accelerating global glacier mass loss and its spatiotemporal distribution. However, the manuscript also has several weaknesses that are summarized below.

The manuscript lacks a clear focus, and it is hard to follow. For example, although it discusses mass loss and acceleration of mass loss from all glaciers not directly connected to the Greenland

and Antarctic ice sheets relative to the mass loss from the ice sheets (e.g., lines 32-33 in abstract, lines 106-107 & 115-121), it does not present any explicit comparison of global glacier mass loss and polar ice sheet mass loss. While the acceleration of glacier mass loss is emphasized (e.g., lines 112-115), very little is shown about the spatiotemporal distribution of this acceleration, or, in general, about spatiotemporal evolution of the mass loss pattern. For example, extended data Fig. 3 demonstrates that glacier thinning at a single location exhibits a complicated temporal pattern in the 19 years of the study. Therefore, short-term snapshots of mean elevation change rates, such the shown for the Karakoram anomaly in extended data Fig. 8, could provide valuable insight into the temporal variation of mass loss on a regional scale, which cannot be adequately represented by average values. Presenting and analyzing the regional scale interannual variation of mass loss would be an essential new contribution made possible by the new DEM-based reconstruction. It could replace the section of the reconciliation of regional estimates, which could be shown in the extended data section.

In my knowledge, the study is the first one approximating glacier elevation change by combining all (or almost all) high spatial resolution DEMs derived from stereo satellite imagery. The methodology is based on previous research, such as the spatial DEM co-registration of Nuth and Kaab (2011) and the application of Gaussian processes for estimating the temporal covariance of glacier elevation change. However, the manuscript does not provide sufficient details on how these approaches were applied. For example, how did they derive the spatiotemporal function of the co-registration error over a changing surface (eq. 1 in Methods)? What algorithms, spatiotemporal domains, and approaches were used to condition the kernel functions to model the glacier elevation changes (eq. 2 in Methods)? Moreover, no errors are derived or presented for the spatial distribution of elevation changes aggregated into tiles, for example the 1 by 1 degree tiles shown in Figure 2 or in extended Figure 8. Finally, the use of IceBridge photogrammetric DEMs, instead of IceBridge airborne lidar data needs to be justified. Unlike the IceBridge airborne lidar data (ATM, LVIS) that have subdecimeter accuracy, the accuracy of IceBridge DEMs is not well characterized and thus could results in less reliable validation.

Detailed comments:
Line 30: this statement is confusing. The mass loss did not increase by 97 ±20 Gt per year each year but at the end of the two decades a 97 ±20 Gt/yr is added to the initial value.

Lines 32-33: I suggest mentioning mass loss values for the ice sheets

Line 35: is the “North Atlantic anomaly” a new term introduced in this manuscript?

Line 64: what is surface elevation imaging? Does it include stereo DEMs and lidar or only stereo DEMs derived from images? If yes, both optical and SAR?

Line 97-99: figures in this manuscript does not show the entire Antarctic continent. Did the study consider all Antarctic glaciers in RGI?

Line 106: the manuscript would benefit from a clear definition and/or inclusion of specific values for the followings:
(1) Periphery of the ice sheets, how are they defined, what is the area excluded from RGI for calculating specific results? (e.g., line 110: the glaciers beyond the periphery of the ice sheets);
(2) Marine terminating vs. land terminating glaciers (see, for example, lines 127-128). What is the area and contribution of marine terminating glaciers to mass loss? I suggest showing their contributions separately in Extended data table 1.
(3) A clear and consistent naming of the regions in RGI. For example, the region that is referred as Svalbard in line 153 is called 07, Svalbard and Jan Mayen (SJM) in extended data table 1.

Line 114-116: how much did the glacier surface area decrease and what is the related change in glacier mass loss?

Line 118: explain the meaning of "acceleration of GIS mass loss nearly extended to the period 2000-2019 is not statistically significant …..."? It is an interesting idea to use the results of this study to confirm the conclusions of Velicogna et al., 2020 – but a stronger argument is needed.

Line 121-122: what is a "climate change-driven sea level trend"?

Line 141: are changes shown in Fig. 2 statistically significant? Including a figure showing the errors would be useful.

Line 151: selecting one study that agrees well seems to be a biased evaluation. What is the meaning of "most confidently resolved"? A study with the smallest uncertainty?

Line 154: please explain what a "minor difference in Svalbard could originate in the delayed attribution of mass transfer from massive surge events, accounted for prematurely by our geodetic method" means?

Line 161-165: I suggest avoiding the expression of "geodetic study." Geodetic methods, i.e., repeat measurements of absolute surface height using remote sensing methods provide regional estimates compared to in situ measurements of mass changes, for example, using stakes. However, with the advent of remote sensing applications, including gravimetry, lidar, repeat photogrammetry, using this terminology has its limitations. Do you include gravimetry-based estimates, such as GRACE, for example?

Lines 172-173: an additional map with a different map projection (polar stereographic?) wouold be needed to show the Antarctic mass loss pattern.

Lines 184-197: this section, describing the temporal evolution/acceleration of mass loss could be a highlight of the manuscript. Instead, it is very vaguely written and illustrated. In particular, the meaning of including two percentages for the different regions, one for mass loss rate increase and another one for regional increase in thinning rate, is not clear and needs further clarification.

Line 198-208: this paragraph is very hard to follow and the statements need better explanations and illustrations.

Line 211: I expect that changes are related to the warming of the atmosphere, not the troposphere.

Line 221: please elaborate on the meaning of complete. Does it mean that no additional surface elevation change is available and can be fused with the record presented here?

Author Rebuttals to Initial Comments (please note that the authors quoted the reviewers in black text and responded in blue text):

Referee #1 (Remarks to the Author):

A] The manuscript details a globally comprehensive description of the state of mountain glaciers and ice caps that leverages use of the ASTER archive, REMA, ArcticDEM, TanDEM-X and other sources.

Its significant results are:

1) it is the first global high resolution survey with consistent methodology to provide estimates of mountain glacier and ice cap change for all glaciers on earth.

2) it shows that glacier mass wastage has accelerated unequally on regional scales

3) it provides up to date estimates of sea level contributions from glacier regions

4) it confirms the presence of the influence of several climate oscillations on regional glacier mass balance

B] The significance of the work is high and is eagerly awaited by much of the glaciological and climate community. Most of the methods are relatively standard but have been scaled to an extraordinary degree, a mammoth task. The uncertainty calculations are complicated but seem to cover all the bases. I recommend a statistician to look them over. The paper somewhat surprisingly lacks citations or acknowledgments of previous studies and methods that it has built upon to set the stage for the work and conclusions - this can be easily remedied.

We thank referee #1 for the thoughtful comments on our paper and for reminding us to acknowledge previous work. We have added several early references that exploited ASTER DEMs, such as Nuimura et al. (2012) and Willis et al. (2012), which were only partially listed in the SI in our first submission. Studies that developed or employed methods similar to ours are referenced throughout the Methods section of our paper. Many of these were omitted from the introduction of the main text given the strict space limitations of *Nature* (up to 30 references).

C] The data is vast and hopefully will be available to the community soon. The approach is valid and builds upon previous studies. The diagrams could do with a little work (see below).

We confirm that all the data will be available through the Data availability statement and can be manipulated using code that we provide (https://github.com/rhugonnet/ww_tvol_study).

D] I got a little lost with the plethora of steps for the uncertainty calculations, and don't feel I can fully comment on whether they are valid or not. I am surprised that there isn't a little more work on ocean/ice processes when considering marine terminating glaciers. Air temperatures and precipitation are treated well, but the paper would be improved if the effect of ocean and lake melt were commented upon.

We have extended our analysis of land- and marine-terminating glacier changes through the addition of Extended Data Table 3. We have modified the related paragraph:

"Marine-terminating glaciers collectively represent 40% of Earth's total glacierized area, yet only contribute 26% to the observed mass loss (Fig. 1). This smaller contribution to sea-level rise is uniform for all maritime regions, except where losses of marine-terminating glaciers are dominated by recent large surge events (e.g. Svalbard and Jan Mayen, Extended Data Fig. 7). The delayed and asynchronous response of tidewater glaciers to changes in climate[26] may partly explain why most marine-terminating glaciers show reduced mass loss. Despite differing mass loss rates, relative acceleration of land- and marine-terminating glaciers within each maritime region are similar (Extended Data Table 3). Notable exceptions exist for glaciers in the Antarctic and Subantarctic, where few land-terminating glaciers are present, and in regions of strong surge-driven mass losses.".

E] The conclusions are robust, but I suggest should be more expansive. The last paragraph of the main text is weak, especially when several national space agencies are contemplating and proposing instrumentation with better capabilities than ASTER. The paper should focus more upon its findings and implications.

We have removed the last paragraph of the conclusions and, instead, used this space to reiterate the major findings and implications of our study. The new conclusions build more clearly on the spatiotemporal resolution of our dataset, following remarks from the editor and referee #3:

"From the spatiotemporally resolved nature of our assessment, novel possibilities arise to harness observations of the satellite era. Not only instrumental for glaciers, such resolved estimates also hold the potential to constrain recent ice sheet mass balance, in particular for the outlet glaciers that are prone to the highest long-term sea-level rise. The improved ability to deconvolve glacier signals from gravimetric observations might foster the detection of nearly two decades of changes in terrestrial water storage. In time, we expect our observational baseline to help drive the development of the next generation of global glaciological and hydrological models, and to ultimately result in more reliable projections at all scales[14]. In light of the rapid, ongoing change of the cryosphere, the increasingly reliable projections made possible by accurate, global-scale observations are critical for the design of adaptation strategies, with impacts ranging from further sea-level rise[4,11] to changes in water management for some of the most vulnerable regions on Earth[12,15].".

F] Please be careful to make sure you are clarify and make your text a little clearer - this is purely "word smithing," for example, in the abstract you present that mass loss has increased by 97 Gt per year (~0.3mm sea level) over the last 20 years - it would be better to state that IN TOTAL mass loss has increased by....

A comment also made by referee #3. We now provide a statement with the value of acceleration per decade for clarity: "*We identify a mass loss acceleration of 48 ± 16 Gt yr^{-1} per decade, explaining 6-19% of the observed acceleration of sea-level rise.*".

Extended tables 1 and 2 are goldmines, but are very hard to read when printed out. Maybe include them as csv files or the like?

Extended Data Tables are now provided through the Data availability statement.

Please cite work by Bevis et al on the previously identified North Atlantic Anomaly and the well known "pause" in mass loss for Greenland around 2013.

Thanks for pointing to this relevant reference. Bevis et al. (2019) is now cited.

The last part of the abstract is messy and the first paragraph of the paper (the rationale) could be simplified and made clearer.

We have clarified the last two sentences of the abstract about perspectives:

"*We anticipate our individually-resolved estimates to foster the understanding of drivers that govern the distribution of glacier mass change, and to extend our capabilities of predicting these changes at all scales. Predictions robustly benchmarked against observations are critically needed to design adaptive policies for the management of local water resources and cryospheric risks as well as for regional-to-global sea-level rise*".

We have streamlined the first paragraph on the main text by removing some non-essential details and repetitions:

"*About 200 million people live on land predicted to fall below the high-tide lines of rising sea levels by the end of the century*[11]*, while more than one billion could face water shortage and food insecurity within the next three decades*[4]*. Glaciers distinct from the ice sheets play a prominent role in these repercussions as the largest estimated contributor to twenty-first century sea-level rise after thermal expansion*[2]*, and as one of the most climate-sensitive constituents of the world's natural water towers*[12,13]*. Current glacier retreat temporarily mitigates water stress on populations reliant on ice* *reserves by increasing river runoff[1], but this short-lived effect will eventually decline[14]. Understanding present-day and future glacier mass change is thus crucial to avoid water scarcity-induced socio-political instability[15], to predict the alteration of coastal areas due to sea-level rise[4], and to assess the impacts on ecosystems[16] as well as on cryosphere-related hazards[3].".*

Claims that ASTER is under utilized should be backed up. As should claims that most large scale data sets have been used.

We have replaced "*most large-scale*" with simply "*large-scale*" and removed "*under-exploited*" to simplify the narrative. Those claims were based on the fact that we here use 100% of ASTER and ArcticDEM and REMA data covering glaciers, while less than 5% of these archives were used at once in previous studies. The claim on using most large-scale datasets was based on the fact that we have used all optical datasets covering at least two first-order RGI regions, and exclude radar datasets (due to penetration of signals into ice).

I suggest that the variation of the density chosen for the ice and firn be varied slightly to get an idea of the uncertainty in the density that could be a further bound on uncertainties.

We agree with referee #1 in that the use of a single estimate for density would be problematic. The approach suggested by referee #1 is accounted for through our uncertainty propagation. Making the density vary slightly corresponds to propagating the available information on density uncertainty (±60 kg m^{-3} following Huss (2013); ±120 at our 95% confidence level) through our calculations. At the scale of RGI regions, we apply the uncertainty as if completely correlated in space. This leads to more conservative density-based mass change uncertainties than in previous regional studies based on DEM differencing (this is because the studies often considered RGI subregions as independent; e.g., Brun et al. (2017), Braun et al., (2019)). Additionally, the current formulation of the density uncertainty proposed by Huss (2013) linearly scales with the most negative specific mass changes. This is also conservative, since it is known that the value of 850 kg m^{-3} is best applicable with the most negative rates. This effect thus provides uncertainties that are likely too large for regions with strong mass loss. Note that the uncertainty of the density conversion factor is currently limiting all existing geodetic studies. Challenges arise from both the poorly constrained dependencies between density conversions and specific mass change rates, and the poorly known correlation of density conversions in time and space. Progress in the understanding of these conversion factors might be possible by coupling our new observational baseline with glaciological models, a possibility that is not feasible in the frame of this study.

We have added a short paragraph in the Supplementary Discussion to clarify these aspects.

Line 106 - you quote a total but provide a rate leaving the reader to search for the equivalent amount for glaciers vs Greenland vs Antarctica. Indeed the total amount seems to be only mentioned in the diagram (5325 Gt total)

We have added values for the Greenland and Antarctic ice sheets (GIS, AIS) in Table 1, referenced at this point in the text. We also added a short paragraph in the Methods to clarify the comparison to GIS and AIS, following remarks from referee #2.

We have modified values of Fig. 1 into mass change rates instead of cumulative mass change to match those of the rest of the paper.

I encourage you to think about the conclusions section of the paper more clearly - there are several means of collecting stereo imagery with national agencies proposing new missions. You show agreement with GRACE and ICESat in many situations so a more compelling argument for your findings and the granularity you observe is warranted in the conclusions.

We have removed the section about future satellite missions. Modifications of the conclusions is detailed in a previous answer, based on similar remarks by referee #3 and the editor.

G] Your references are robust, but I think there are some more places where citations (such as the previously mentioned paper on the north Atlantic anomaly) could be improved upon. There is work on the Russian Arctic, Dave Sheans work on HIMA, available at NSIDC, etc that has not been cited and provides context for your numbers.

We agree with the referee and indeed struggled when deciding which key references to add in the main text given the stringent space limitations of *Nature*. The regional study suggested by referee #1 for HMA (Shean et al. (2020)) was already referenced in the main text. Other regional studies were compared in Supplementary Table 1 (previously Table S4, which corresponds to the IPCC SROCC table for glacier mass change). To our knowledge, we have compared our results to all recent global and regional studies that cover full RGI first-order regions. For reasons of space, we cannot integrate in the paper a comparison to all sub-regional (e.g., RGI second-order regions) or local studies in the but leave the means to perform this comparison by providing the tools to aggregate our dataset over any period and region (see Code & Data availability statements). RGI second-order regions will also be directly provided through the Data availability statement.

H] I think the main work for improving clarity on the paper should focus on the introduction and conclusions. FIG1 should have the total mass loss per circle adjacent to each circle. This would aid in things like sea level fingerprinting studies.

Modifications of the introduction and conclusions are detailed in previous answers.

Estimates of total mass loss are now displayed on Fig. 1 for regions with mass loss rates larger than 4 Gt yr^{-1}. Other values can be found in Extended Data Table 1.

In figure 2 large surge events in Svalbard and the Russian Arctic do not seem to be reflected in the diagram, which is surprising.

On Fig. 2, surges cannot be seen due to the large amount of non-surging glaciers whose mass changes are aggregated within the same 2°x2° tiles, dampening the strong mass loss of these individual glaciers. However, these rapid changes can be observed on Extended Data Fig. 6 that is aggregated by 1°x1° tiles, and at the pixel scale on Extended Data Fig. 7.

Overall, I suggest this paper is appropriate for a nature journal and only requires some wordsmithing and the opportunity to punch up its conclusions for higher impact.

We again thank the referee for these valuable comments which, we feel, have strengthened our paper.

Referee #2 (Remarks to the Author):

Review of ‘Accelerated global glacier mass loss in the early twenty-first century’

In this paper, the authors use satellite stereo imaging to estimate glacier mass loss over 2000-2019. This new approach complements other estimates from models, satellite gravimetry, altimetry, and in-situ measurements and the estimated mass loss falls largely in line with these other estimates. This study confirms that glacier melt forms a substantial contribution to sea-level rise since 2000. As I am more of a sea-level than a glacier remote sensing expert, I’ll limit my review to the sea-level related parts of this manuscript. The new method is interesting and the high spatial resolution could be useful to determine the regional sea-level response to glacier mass loss. The method also allows for a more accurate distinction between mass loss on the ice sheets and the peripheral glaciers.

However, I think the manuscript can benefit from some extra analyses, and I urge the authors to carefully check all comparisons with ice-sheet mass balances.

The study could profit from a more thorough comparison to other mass loss estimates, for example from GRACE observations and models. Currently, Figure 3 shows a comparison with some estimates, but recent studies like Ciraci et al. (2020) and Bamber et al. (2018) are not included. Does this study point at a significant departure of global-mean mass loss from these previous estimates? From a quick scan, it seems that the global mass-loss estimates from these studies are all in line with the results discussed here. I also wonder how the observed mass loss compares to the changes on longer time scales, for example from Parkes and Marzeion (2018) and Zemp et al. (2019): are the accelerations and trends you see unique and unprecedented over the observational record, or are they within the range of variability that has occurred before?

We thank the referee for his input and provide answers to the questions raised above.

On the comparison to additional observational studies.

We have added the comparison to Ciracì et al. (2020) which was published shortly before we submitted. The study of Bamber et al. (2018) is a compilation of several existing glacier estimates, for the most part already compared either in Fig. 3 or in Supplementary Table 1. We thus do not add a separate comparison to Bamber et al. (2018) as we feel that it would not provide additional insight into the origin of the differences in the estimates. To our knowledge, Zemp et al. (2019), Gardner et al. (2013), Wouters et al. (2019) and the added Ciracì et al. (2020) are the only observational global or near-global recent studies on glacier mass change. We found a good overall agreement with Gardner et al. (2013) and Wouters et al. (2019) but our glacier mass loss rates are smaller compared to Zemp et al. (2019) and Ciracì et al. (2020). The comparison is provided on Fig. 3.

On the acceleration.

While studies such as Zemp et al. (2019) did not isolate a statistically significant acceleration due to large uncertainties, the recent acceleration derived from Ciracì et al. (2020) is significant. Note, however, that their value has very large uncertainties, and that they cannot isolate glaciers in the Greenland Periphery or the Antarctic and Subantarctic from the ice sheets.

We have now added several sentences to more clearly position our estimates of acceleration against those mentioned in existing studies:

"Observational studies were yet unable to discern significant accelerated glacier mass loss[19,21]*, with the exception of a recent gravimetric study*[20] *that estimated an acceleration of 50 ± 40 Gt yr-1 per decade excluding peripheral glaciers. Despite its large uncertainties, this estimate is in agreement with our results."*.

On the comparison to modelling studies.

Past modelling studies are calibrated on a limited, post-2000 observational dataset. It is therefore difficult to compare the overlapping trends as the modelled ones are dominated by calibration uncertainties of previous observational assessments. Due to those limitations, we chose not to compare with any modelling estimate.

Can the authors provide data sets that are usable for computing the impact of glacier mass loss on global and regional sea-level? Global studies like WCRP Global Sea Level Budget Group (2018) or regional studies like Rietbroek et al. (2016) require time series of mass loss aggregated by RGI glacier region, or even better, on a grid. The latter can be directly used to compute the resulting spatial sea-level rise patterns due to glacier mass loss (see for example Adhikari et al. 2019).

We now provide datasets in the form of aggregated mass change grids for direct use by the sea-level community. This decision was triggered by the comment of referee #2 and discussions with our colleague B. Meyssignac, an expert in the domain of sea-level change assessments.

I also urge the authors to thoroughly check all comparisons between glacier and ice mass loss. Trends and accelerations are very sensitive to the chosen start and end points of the time series. Furthermore, there's likely temporal autocorrelation present in the time series, which is not included in the trend and acceleration uncertainty calculations as far as I can see. Therefore, I am reluctant to accept the statements where trends and accelerations between glaciers and ice sheets are compared. These comparisons beyond the level of 'similar in magnitude' should either be removed, or substantiated by comparing all time series over the exact same period with a full analysis of the error budget. See the line-by-line comments below for some examples.

These are valid points raised by the referee which we address below.

On the comparison of glaciers with the GIS and AIS.

We have now added a dedicated section in Methods, and an explicit comparison to the AIS and GIS in Table 1.

A major difference for trend comparison that was not explicit in the main text is that a subtraction is sometimes needed to compare ice sheet and glacier mass loss. In some AIS and GIS studies, mass losses are reported including glaciers in the periphery and the periphery component needs to be removed (e.g., IMBIE (2018), (2020); Velicogna et al. (2020)), while in other studies, such as Smith et al. (2020), the estimate is provided for the ice sheet only. Once this subtraction is performed, the estimates of ice sheet mass loss agree with our statements: mass loss from glaciers is larger than mass loss from the GIS or AIS. We underline that our resolved estimates of mass change for peripheral glaciers, and possibly also outlet glaciers, will help to improve the comparability of different ice sheet estimates (e.g., future IMBIE exercise).

On the temporal correlation for deriving acceleration.

We have now added a short dedicated section in the Methods to justify and discuss temporally correlated uncertainties.

Accelerations are derived through weighted-least squares using the 5-year estimates and propagating their uncertainties which are assumed to be uncorrelated in time. Contrary to sea-level altimetry, there is no temporal correlation present in ASTER or WorldView surface elevation observations. However, our Gaussian Process (GP) regression does create some autocorrelation through temporal interpolation at the local scale. Given the short periods characterizing the local kernels of the GP and the average of ~39 observations per pixel in 20 years of observations, this autocorrelation remains limited in time. Once aggregated amongst the half billion pixels to derive global volume change, correlated errors compensate again and leave only little space for a residual temporal autocorrelation. This is backed up by our comparison to ICESat and IceBridge that shows small elevation change biases at the scale of RGI regions (Table S3 in the SI). At the global scale, we thus consider that volume changes have negligible temporal correlation for periods longer than 5 years. The remaining factor leading to temporal correlation is the conversion from volume to mass change. Based on Huss (2013), we can assume 5-year periods to be largely uncorrelated at the scale of RGI regions, but not totally. However, once aggregated at the global-scale, these correlated errors compensate again and we can assume the density-based correlation over 5-year periods to be negligible. It is also worth noting that we applied density uncertainties as if those were fully spatially correlated at the regional scale. This leads to conservative mass change uncertainties, which are then propagated to the acceleration uncertainties.

Line-by-line comments

L23-24: “Yet, due to the scarcity of homogeneous mass loss observations, their recent evolution is only partially known as a geographic and temporal patchwork.”

I think this statement is a bit too gross. How about GRACE and GRACE-FO observations? GRACE/FO provides direct mass loss observations since 2002, and there are many approaches available to determine global and regional glacier mass from these observations, for example Reager et al. 2015, Bamber et al. 2018, Ciraci et al. 2020, and Wouters et al. 2019.

We modified the statement into :”*Yet, due to the scarcity of constrained mass loss observations, glacier evolution during the satellite era is only known as a geographic and temporal patchwork*[4,5].”.

Regarding our choice for the statement in the abstract, it is essentially a reword of the conclusions of the recent IPCC SROCC report. The latter is based on both the current state of glacier mass change estimates and recent communications made by the report's coordinating lead author (R. Hock). The same point is also nicely illustrated by Gardner et al. (2013) - a study including several specialists from the GRACE community - who had to decide region by region which method(s) was(were) more reliable, essentially producing a "patchwork" to find the most reliable estimate. We now reference the IPCC and Gardner et al. (2013) at this point.

While GRACE and GRACE-FO provide almost direct mass loss observations for ice sheets when isostatic rebound is subtracted, glaciers remain difficult to resolve due to the coarse spatial resolution of the gravimetric signal captured by GRACE, the subtraction of the solid Earth response, and the need for deconvoluting hydrological signals. These combined issues have resulted in an absence of estimates for glaciers located near ice sheet peripheries (that cannot be differentiated with the ice sheet themselves), poorly constrained estimates for lightly glacierized regions (such as New Zealand and Central Europe) and estimates that differ significantly between GRACE-based studies even in regions where uncertainties are supposed to be small. For example Iceland mass change rates for Ciracì et al. (2020) are estimated at -16 ± 4 Gt yr^{-1} and for Wouters et al. (2019) at -10 ± 2 Gt yr^{-1} with only a couple year difference in the study periods; and for Russian Arctic supposedly best constrained by gravimetric methods: −10.6 ± 1.7 for Wouters and -20.2 ± 12 Gt yr^{-1} for Ciracì. Fig. 3a now allows for a full comparison

We have also clarified our related statement in the introduction: "*Notwithstanding recent progress in glacier monitoring from space*[18]*, global-scale remote sensing-based studies have been so far limited to (i) the coarse spatial resolution of satellite gravimetry, unable to reliably disentangle glacier mass change signals from those of the ice sheets, solid Earth and hydrology in many regions*[5,19,20]*.*".

L29-30: "equivalent to 24 ± 5% of observed sea-level rise."

Because of the relatively short time series and the presence of serious serial correlation in sea-level time series, estimates trends and accelerations in sea level are highly sensitive to the exact beginning and end period, and the uncertainty in the trend in altimetry-derived sea level is still not negligible (Ablain et al. 2019). Since the sea-level time series from the quoted source only covers 1993-2017, I wonder whether the trends have been computed over the same interval, and how this uncertainty range has been determined. As an alternative, global-mean sea level time series up to present can be found from traditional data centers, such as AVISO, CSIRO, or NASA.

We agree with referee #2 and have added a short Methods section on the sea-level time series.

We have contacted M. Ablain who has provided us with an updated time series (AVISO-based) and uncertainties for the exact period of 2000-2019 (2000-01-01 to 31/12/2019). We have updated all sea-level values based on this estimate in the text.

L32-33: “Collectively, glaciers presently lose more mass, and at more accelerated rates, than the Greenland or Antarctic ice sheets.”

I agree with this statement for Antarctica, but for Greenland I don’t. Over 2002-present, GRACE observations suggest Greenland and glaciers to be of similar magnitude (Ciraci et al. 2020, Velicogna et al. 2020, Frederikse et al. 2019). Unless the authors can somehow prove a significant difference between glaciers and Greenland, I think this statement should be reframed as ‘larger than Antractica, similar to Greenland’, or something along this line.

We have clarified the differentiation between the GIS, AIS and their peripheral glaciers in Methods. With this clear differentiation between glaciers, GIS and AIS, the difference in trend between glaciers and the GIS is significant (Table 1).

We have clarified our related statement in the abstract: “*Glaciers presently lose more mass, and at similar or larger accelerated rates, than the Greenland or Antarctic ice sheets taken separately*$^{7–9}$.”.

And also in the main text: “*From 2000 to 2019, global glacier mass loss totalled 266 ± 16 Gt yr-1, a mass loss 47% larger than that of the GIS, and more than twice that of the AIS*$^{7–9}$ *(Table 1).*”.

L44 ‘tide lines’ replace with ‘high-tide level’.

Replaced by “*high-tide lines*”, as used in Kulp et al. (2019), and to avoid repetition of “*level*” in the sentence.

L48: “largest estimated contributor to current sea-level rise after thermal expansion”

This statement is again very sensitive to the time period it refers to. Do the authors refer to 1993-2016 here?

We have added “*twenty-first century*” in the sentence. The statement indeed refers to the

World Climate Research Programme (WCRP) assessment for the period 1993-2016. This statement remains valid along the rest of the early twenty-first century, as demonstrated by our study.

L119-L121: “We thereby infer”

To which period do the authors refer here? Sea level has been accelerating since the 1960s, but the underlying causes vary substantially with time (e.g. Dangendorf et al. 2019).

We have added “*since 2000*” in the sentence for clarity.

L223-225: Where does this factor 10 come from? From IPCC SROCC table 4.1? That table suggests an uncertainty on the order of 0.1-0.2 mm/yr, translating into 30-70 gt/yr, which is definitely not a factor 10 larger than the numbers presented here.

We have now modified this sentence for clarity:

“Benefiting from the nearly complete spatial coverage afforded by ASTER stereo-imagery, our global estimate of recent glacier mass change (-274 ± 18 Gt yr^{-1} for the 2006-2015 IPCC reference period) shows strongly reduced uncertainties compared to the latest IPCC report[4] (-278 ± 226 Gt yr^{-1}) and a recent global study[21] (-335 ± 144 Gt yr^{-1}).”.

Here the latter estimate cited is from Zemp et al. (2019).

L225-L227: “We distinctly constrain the trend of glaciers towards larger mass losses and their contribution to sea-level rise and its acceleration…”:

Larger mass loss than what? Previous estimates? The numbers seem to be in line with most other estimates (Figure 3b, Ciraci et al. 2020, Zemp et al. 2018)

We here meant accelerated trends by the term “*trend towards larger mass losses*”.

We have now modified this sentence for clarity:

“We resolve the time-varying nature of this mass change signal for nearly all of Earth’s glaciers which, globally, reveals a significant accelerated mass loss.”.

L227-L231: “no relaxation is in sight for the globally accelerated mass loss rates of Earth’s glaciers”

I’m not a glacier expert, but how can you draw this conclusion from such short time series? How have you separated internal variability from forced melt? For example, model estimates show mass loss rates in the 1930s to be much larger than what we observe today (Parkes and Marzeion, 2018), which came down a lot in the 1950s. In the next paragraph, this problem is directly discussed, which makes me wonder how to interpret this statement.

We have now removed this statement, which was meant to relate our observational mass changes and climate sensitives to model predictions.

We again thank referee #2 for the review that greatly improved the clarity of our comparison to sea-level and ice sheet change estimates. In particular, we feel that the important statements on the differentiation between the GIS, AIS and glaciers at their peripheries, now summarized in Table 1, will benefit the glaciological and sea-level budget community.

References

Ablain, M., Meyssignac, B., Zawadzki, L., Jugier, R., Ribes, A., Spada, G., Benveniste, J., Cazenave, A., & Picot, N. (2019). Uncertainty in satellite estimates of global mean sea-level changes, trend and acceleration. Earth System Science Data, 11(3), 1189–1202. https://doi.org/10.5194/essd-11-1189-2019

Adhikari, S., Ivins, E. R., Frederikse, T., Landerer, F. W., & Caron, L. (2019). Sea-level fingerprints emergent from GRACE mission data. Earth System Science Data, 11(2), 629–646. https://doi.org/10.5194/essd-11-629-2019

Bamber, J. L., Westaway, R. M., Marzeion, B., & Wouters, B. (2018). The land ice contribution to sea level during the satellite era. Environmental Research Letters, 13(6), 063008. https://doi.org/10.1088/1748-9326/aac2f0

Ciracì, E., Velicogna, I., & Swenson, S. (2020). Continuity of the Mass Loss of the World's Glaciers and Ice Caps From the GRACE and GRACE Follow-On Missions. Geophysical Research Letters, 47(9). https://doi.org/10.1029/2019GL086926

Dangendorf, S., Hay, C., Calafat, F. M., Marcos, M., Piecuch, C. G., Berk, K., & Jensen, J. (2019). Persistent acceleration in global sea-level rise since the 1960s. Nature Climate Change. https://doi.org/10.1038/s41558-019-0531-8

Frederikse, T., Landerer, F. W., & Caron, L. (2019). The imprints of contemporary mass redistribution on local sea level and vertical land motion observations. Solid Earth, 10(6), 1971–1987. https://doi.org/10.5194/se-10-1971-2019

Parkes, D., & Marzeion, B. (2018). Twentieth-century contribution to sea-level rise from uncharted glaciers. Nature, 563(7732), 551–554. https://doi.org/10.1038/s41586-018-0687-9

Parkes, D., & Marzeion, B. (2018). Twentieth-century contribution to sea-level rise from uncharted glaciers. Nature, 563(7732), 551–554. https://doi.org/10.1038/s41586-018-0687-9

Reager, J. T., Gardner, A. S., Famiglietti, J. S., Wiese, D. N., Eicker, A., & Lo, M.-H. (2016). A decade of sea level rise slowed by climate-driven hydrology. Science, 351(6274), 699–703. https://doi.org/10.1126/science.aad8386

Rietbroek, R., Brunnabend, S.-E., Kusche, J., Schröter, J., & Dahle, C. (2016). Revisiting the contemporary sea-level budget on global and regional scales. Proceedings of the National Academy of Sciences, 113(6), 1504–1509. https://doi.org/10.1073/pnas.1519132113

Velicogna, I., Mohajerani, Y., A, G., Landerer, F., Mouginot, J., Noel, B., Rignot, E., Sutterley, T., Broeke, M., Wessem, M., & Wiese, D. (2020). Continuity of Ice Sheet Mass Loss in Greenland and Antarctica From the GRACE and GRACE Follow-On Missions. Geophysical Research Letters, 47(8). https://doi.org/10.1029/2020GL087291

WCRP Global Sea Level Budget Group. (2018). Global sea-level budget 1993–present. Earth System Science Data, 10(3), 1551–1590. https://doi.org/10.5194/essd-10-1551-2018

Wouters, B., Gardner, A. S., & Moholdt, G. (2019). Global Glacier Mass Loss During the GRACE Satellite Mission (2002-2016). Frontiers in Earth Science, 7, 96. https://doi.org/10.3389/feart.2019.00096

Zemp, M., Huss, M., Thibert, E., Eckert, N., McNabb, R., Huber, J., Barandun, M., Machguth, H., Nussbaumer, S. U., Gärtner-Roer, I., Thomson, L., Paul, F., Maussion, F., Kutuzov, S., & Cogley, J. G. (2019). Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016. Nature, 568(7752), 382–386. https://doi.org/10.1038/s41586-019-1071-0

Referee #3 (Remarks to the Author):

The manuscript presents a new, detailed reconstruction of global glacier mass changes by combining a massive set of DEMs, "covering 20 times the land area of the Earth", derived from stereo satellite imagery and coregistered to the TanDEM-X global DEM. The resulting three-dimensional array of elevation time series has 100 by 100-meter resolution. Using a Gaussian Process (GP) regression, a continuous function of elevation change is estimated with a combination of kernels accounting for interannual and seasonal variations. The new reconstruction derives mass changes for the 19 regions of the Randolph Glacier Inventory (RGI) with smaller errors than previous results. The high-temporal resolution of the elevation change reconstruction allows the calculation of mass changes for shorter periods within the study's 19-year span. The results are presented in several ways, including plots and tables of total mass loss divided into mass loss estimates from land-terminating and marine glaciers, as well as five-year mass loss, and mean elevation-change rate time series for each of the 19 regions (Figure 1 and Extended Table 1). The comprehensive results are expected to serve as an essential observational baseline for improving model results and driving policy decisions.

I acknowledge the innovative nature of the new reconstruction and the critically important conclusions derived from the results about the accelerating global glacier mass loss and its spatiotemporal distribution. However, the manuscript also has several weaknesses that are summarized below.

The manuscript lacks a clear focus, and it is hard to follow. For example, although it discusses mass loss and acceleration of mass loss from all glaciers not directly connected to the Greenland and Antarctic ice sheets relative to the mass loss from the ice sheets (e.g., lines 32-33 in abstract, lines 106-107 & 115-121), it does not present any explicit comparison of global glacier mass loss and polar ice sheet mass loss. While the acceleration of glacier mass loss is emphasized (e.g., lines 112-115), very little is shown about the spatiotemporal distribution of this acceleration, or, in general, about spatiotemporal evolution of the mass loss pattern. For example, extended data Fig. 3 demonstrates that glacier thinning at a single location exhibits a complicated temporal pattern in the 19 years of the study. Therefore, short-term snapshots of mean elevation change rates, such the shown for the Karakoram anomaly in extended data Fig. 8, could provide valuable insight into the temporal variation of mass loss on a regional scale, which cannot be adequately represented by average values. Presenting and analyzing the regional scale interannual variation of mass loss would be an essential new contribution made possible by the new DEM-based reconstruction. It could replace the section of the reconciliation of regional estimates, which could be shown in the extended data section.

We thank referee #3 for the comments and provide point by point answers below.

On the comparison to AIS and GIS.

We have added Table 1 comparing glacier mass loss to estimates for the AIS and GIS. Providing values for this comparison was also mentioned by referee #1. Following referee #2's comments, we have also added a short section in the Methods to clarify the distinction between the AIS, GIS, and glaciers in the peripheries.

On the spatiotemporal distribution of the mass loss pattern.

We have now added Extended Data Fig. 7 representing the 1x1°, 5-year elevation change rate pattern at the global scale in 4 panels for 2000-2004, 2005-2009, 2010-2014 and 2015-2019. The tile size is inversely scaled with the error. Previous Extended Data Fig. 8 on the Karakoram anomaly is now included in this new Figure. We also added multiple regional panels to Extended Data Fig. 6 (previously Extended Data Fig. 7) to show the high spatial resolution in several regions of interest.

As stated in the article, we note that the current uncertainty analysis shows that mean elevation change rates for periods shorter than 5 years are not always statistically significant, depending on observational coverage. We thus do not detail our analysis for periods shorter than 5 years.

We point out that Fig. 4A summarizes the patterns of acceleration and deceleration between the 2 decades of our study at a fine spatial scale and with scaling to errors. Tests showed that plotting the difference in elevation change rates (Fig. 4A) rather than the direct mass change (as now presented in Extended Data Fig. 7) made it easier to visualize temporal changes.

We understand that referee #3 wants to see even more of the spatiotemporal resolution analyzed. However, we do not think that the paper can accommodate a more detailed analysis than the global, regional, and tile scales. The same is true for periods of 5 to 20-year periods, and on top of the methods described to derive the entire dataset. The current analysis already leads to multiple findings: the constrained global glacier mass change estimates, the identification of significant global acceleration, the contrasted regional patterns of loss, the subregional temporal anomalies, and their link to recent changes in climate.

In my knowledge, the study is the first one approximating glacier elevation change by combining all (or almost all) high spatial resolution DEMs derived from stereo satellite imagery. The methodology is based on previous research, such as the spatial DEM co-registration of Nuth and Kaab (2011) and the application of Gaussian processes for estimating the temporal covariance of glacier elevation change. However, the manuscript does not provide sufficient details on how these approaches were applied. For example, how did they derive the spatiotemporal function of the co-registration error over a changing surface (eq. 1 in Methods)? What algorithms, spatiotemporal domains, and approaches were used to condition the kernel functions to model the glacier elevation changes (eq. 2 in Methods)? Moreover, no errors are derived or presented for the spatial distribution of elevation changes aggregated into tiles, for example the 1 by 1 degree tiles shown in Figure 2 or in extended Figure 8.

We have strived to clarify the points highlighted by referee #3 in the new Methods section. Those aspects were originally described in the SI. We have now removed all duplication, merged most sections into the main Methods, and left only some specific aspects to the SI to facilitate navigation.

On the co-registration error.

In the main Methods, we have clarified that the errors related to slope and quality of stereo-correlation were found to be consistent at the global scale and applied similarly for all elevation observations:

"We found that the empirical variances for the slope ${\sigma_\alpha}^2$ *and the quality of stereo-correlation* ${\sigma_q}^2$ *were consistent between regions, and used them to condition a model at the global scale to account for the measurement error independently for each elevation observation* $h(t, x, y)$*"*.

In the SI, we have clarified that the co-registration error is calculated independently for each DEM as the Root Mean Square Error on ice-free terrain, after corrections and co-registration to TanDEM-X. We mention that this approach is conservative as it might double-count effects from steep slopes and low quality of stereo-correlation:

"The co-registration error σ_c *would ideally have to be estimated on pixels with low slopes and good qualities of stereo-correlation to avoid double-counting the effect of other errors. However, as this is not possible for some DEMs because of the limited amount of flat terrain available, we conservatively used the RMSE of elevation differences over all available stable terrain to derive* σ_c*."*.

Together, those three errors (i.e. co-registration, steep slopes, quality of stereo-correlation) are used to derive the elevation measurement error of Equation (1).

On the Gaussian Process kernels.

We have clarified in the main Methods that the kernels are conditioned by empirical variograms and applied similarly at the global-scale:

"We thus conditioned the parameters of the ESS, RBF and RQ kernels at the global scale based on our empirical variograms, while the PL kernel was determined directly from the observations of each pixel (x,y) (Extended Data Fig. 3b).".

Our sampling of empirical variograms demonstrated that temporal covariance was similar between regions. This aspect has been left in the SI:

"The empirical temporal variances varied little between regions, and we found no significant variability with external factors (such as slope), as we did for the elevation measurement error.".

As discussed in the SI, sensitivity tests of Gaussian Process regression showed that kernel parameters of the same magnitude had little influence on the final estimates. Gaussian Process regression for our application is essentially an interpolation method, for which the resulting time series is primarily driven by the observations. The temporal covariance only serves to mitigate non-linear and seasonal biases:

"The variances described here do not directly condition the mean of the GP elevation time series, which is interpolated from available observations, but only leave the opportunity to find periodicity and local variations in those observations within an order of magnitude.".

We found by sensitivity tests that Gaussian Process parameters mostly impacted empirical confidence intervals derived by the method. This is also due to the fact that ASTER data has a rough vertical precision (~5 m) and thus the Gaussian Process regression cannot always deconvolve local or periodic signals from the measurement error. These specific aspects are described in the SI:

"The limited influence of GP parameters within the same order of magnitude is due to the relatively large measurement error of ASTER elevations (of about 5 m) which generally prevents complete deconvolution of local and periodic signals. This effect is later accounted for by our uncertainty propagation of interpolation biases (Section 4.3) and, when aggregated at different spatial scales, is essentially what defines the temporal resolution of our dataset (Supplementary Discussion). We found that the confidence interval of the regression was the most impacted by parameter changes and was thus validated in a later analysis (Section 3.4).".

We validated those confidence intervals when comparing with ICESat and IceBridge, which were found to be conservative by a factor of two. We have moved this statement to the main Methods:

"Standardized comparison additionally demonstrated that our elevation time series uncertainties are conservative by a factor of over two. We reach the same conclusions at the scale of RGI regions, and also perform these verifications with several variables of importance.". These variables of importance are shown on Extended Data Fig. 4.

Additionally, the biases that still might be created by the Gaussian Process interpolation are quantified when assessing the long-range spatial correlations (20-500 km) with observational time lag (Extended Data Fig. 5). Those are then propagated to uncertainties, which is now described succinctly in the main Methods and in details in the SI:

"The uncertainty in the mean elevation change dh is highly subject to spatial correlations due to instrument resolution (0-150 m), uncorrected ASTER instrument noise46 (0-20 km), and the interpolated nature of our elevation time series (0-500 km). For the later, it is explained by the fact that neighboring pixels of a given region generally share similar temporal data gaps, and are hence likely to have similar interpolation biases which correspond to long range correlations.".

On the elevation change errors.

We have added errors of mean elevation change per 5-year period on Extended Data Fig. 8. Errors between 10-year periods were also represented on Fig. 4A.

As Figures (in particular maps) cannot always accommodate a visual representation of errors, we chose not represent errors on Fig. 1 and Fig. 2. Note that the 2000-2019 changes shown in these Figures are rather well constrained (longer period, smaller uncertainties) and that the related uncertainties are available in Extended Data Table 1 or other Figures. At all spatial and temporal scales, errors were computed and are available through the provided data (see Data Availability statement). We also provide the means to aggregate our results at any scale while accounting for spatial correlations (see Code Availability).

Finally, the use of IceBridge photogrammetric DEMs, instead of IceBridge airborne lidar data needs to be justified. Unlike the IceBridge airborne lidar data (ATM, LVIS) that have subdecimeter accuracy, the accuracy of IceBridge DEMs is not well characterized and thus could results in less reliable validation.

We agree with the referee that ATM or LVIS data has higher vertical precision. It however has a smaller spatial coverage. Our primary purpose when validating our 100 m x 100 m time series is not only to have a sufficient vertical precision, but mostly to compare to a wide spatial sampling. Because the vertical precision of a single pixel of our time series is typically around 5 m (described on Extended Data Fig. 3), our validation analysis does not require sub-decimeter accuracy. The IceBridge photogrammetric DEMs, which supposedly do not contain systematic biases as they are vertically calibrated on the ATM and LVIS data (https://nsidc.org/data/IODEM3/versions/1) thus provide a sufficient vertical accuracy (at least < 0.5 m, corresponding to the horizontal resolution of the optical imagery used to derive these DEMs through photogrammetry). This is for example pointed out in the IODMS3 documentation (https://nsidc.org/data/iodms3): "*The DEMs contain substantially more elevation detail (resolution) than the LIDAR data.*".

Detailed comments:

Line 30: this statement is confusing. The mass loss did not increase by 97 ±20 Gt per year each year but at the end of the two decades a 97 ±20 Gt/yr is added to the initial value.

This comment was also made by referee #1. We now provide a statement with the value of acceleration per decade for clarity: "*We identify a mass loss acceleration of 48 ± 16 Gt yr^{-1} per decade, explaining 6-19% of the observed acceleration of sea-level rise.*".

Lines 32-33: I suggest mentioning mass loss values for the ice sheets

Following previous comments, we have added Table 1. These values cannot be accommodated directly in the abstract for reasons of space.

Line 35: is the “North Atlantic anomaly” a new term introduced in this manuscript?

Yes. We clarified this with the wording “*newly-identified*”.

Line 64: what is surface elevation imaging? Does it include stereo DEMs and lidar or only stereo DEMs derived from images? If yes, both optical and SAR?

Changed to “*optical and radar surface elevation imaging*” for clarity.

Line 97-99: figures in this manuscript does not show the entire Antarctic continent. Did the study consider all Antarctic glaciers in RGI?

Yes, all glaciers in region 19 (Antarctic and Subantarctic) of the RGI are accounted for in our estimates. Figure space is the only reason that some small ice caps in East Antarctica are not shown. The related information has been moved from the SI to Methods.

Line 106: the manuscript would benefit from a clear definition and/or inclusion of specific values for the followings:

(1) Periphery of the ice sheets, how are they defined, what is the area excluded from RGI for calculating specific results? (e.g., line 110: the glaciers beyond the periphery of the ice sheets);

We have added a statement in the main text, and Table 1 that now provides separate mass change rates for glaciers, the AIS, and the GIS. We have moved the Inventories section from the SI to the Methods where we describe the area excluded from RGI (connectivity level 2 glaciers):

“*In the Greenland Periphery (region 5), we did not analyze the 955 glaciers highly connected to the ice sheet (RGI connectivity level 2) with an area of 40,354 km², as these are generally included by studies on the GIS*[7,9].”.

We also added a section in the Methods to describe the dissociation between glaciers and ice sheets, as well as a statement in the main text:

“*Our analysis includes 200,000 km² of glaciers located in the Greenland Periphery and in the Antarctic and Subantarctic that are distinct from the Greenland Ice Sheet (GIS) and the Antarctic Ice Sheet (AIS). Our estimates of these ice masses, referred to as peripheral glaciers, are instrumental in our comparison with recent ice sheet studies (see Methods).*”.

(2) Marine terminating vs. land terminating glaciers (see, for example, lines 127-128). What is the area and contribution of marine terminating glaciers to mass loss? I suggest showing their contributions separately in Extended data table 1.

We have added Extended Data Table 3, showing area and contributions of marine-terminating glaciers per maritime region. This table supports additional statements in the discussion:

"*Despite differing mass loss rates, relative acceleration of land- and marine-terminating glaciers within each maritime region are similar (Extended Data Table 3). Notable exceptions exist for glaciers in the Antarctic and Subantarctic, where few land-terminating glaciers are present, and in regions of strong surge-driven mass losses.*".

(3) A clear and consistent naming of the regions in RGI. For example, the region that is referred as Svalbard in line 153 is called 07, Svalbard and Jan Mayen (SJM) in extended data table 1.

We corrected such occurrences for "*Svalbard and Jan Mayen*", "*Antarctic and Subantarctic*" and "*Greenland Periphery*". We now refer to regions consistently by their full RGI naming. We have removed all acronyms, not widely used, from Figures and Tables to simplify access to a broader readership.

Line 114-116: how much did the glacier surface area decrease and what is the related change in glacier mass loss?

Global glacier surface areas have decreased by about 10% in 20 years (Zemp et al. (2019)), a factor that only affects the computation of specific mass change (or mean elevation change) rates but not the computation of mass change (in Gt).

We have moved this information from the SI to a section of the Methods:

"*Over the 20-year study period, these time-evolving areas correspond to a nearly 10% decrease of glacier areas around the globe, a non-negligible change when assessing mean elevation change rates.*".

Line 118: explain the meaning of "acceleration of GIS mass loss nearly extended to the period 2000-2019 is not statistically significant …..."? It is an interesting idea to use the results of this study to confirm the conclusions of Velicogna et al., 2020 – but a stronger argument is needed.

We have removed this statement to streamline the comparison of acceleration between glaciers, GIS and AIS. This statement was taken from Velicogna et al. (2020) to provide context to our estimates of acceleration.

Line 121-122: what is a "climate change-driven sea level trend"?

This statement was removed for simplification as sea-level rise trends are now compared to Ablain et al. (2019) (AVISO-based).

The original statement referred to Nerem et al. (2018) that used the term "*climate change-driven sea-level*" to designate sea-level trends after the removal of the terrestrial water storage (TWS) components of El Nino Southern Oscillation or Pinatubo. A comparison to these sea-level trends seemed more relevant as glaciers do not contribute much to the TWS variation that directly impacts the estimation of sea-level rise acceleration. We contacted S. Nerem and, unfortunately, extending his assessment to the end of 2019 was not possible for him.

Line 141: are changes shown in Fig. 2 statistically significant? Including a figure showing the errors would be useful.

Changes displayed on Fig. 2 are almost all significant as the spatial coverage is nearly complete and those represent the full 20 years of study period. Our Figure displayed non-significant changes in grey. We now better detail the related criteria in the caption:

"*Disks scale with the glacierized area of each tile and are colored according to the mean elevation change rate (grey if less than 50% surface covered or if 95% confidence interval larger than 1 m yr^{-1}, only applies to 0.4% of glacierized area).*".

We chose to display glacier areas on Fig. 2, in order to provide both a visualization of glacier areas (Fig. 2) and errors (Fig. 4A, Extended Data Fig. 7) throughout the article. Errors are also available through the separately provided data (see Data Availability statement).

Line 151: selecting one study that agrees well seems to be a biased evaluation. What is the meaning of "most confidently resolved"? A study with the smallest uncertainty?

We modified the sentence for clarity and we have added a statement to relate the reliability of gravimetric estimates: "*Our regional mass change estimates closely match those of a recent* *gravimetric study[19] in remote polar regions (Arctic Canada, Svalbard and Jan Mayen, and the Russian Arctic) where gravimetric uncertainties are considered small due to weak competing signals (Fig. 3). We note, however, large discrepancies between this gravimetric study[19] and a more recent one[20] in both Iceland and the Russian Arctic.".*

Line 154: please explain what a "minor difference in Svalbard could originate in the delayed attribution of mass transfer from massive surge events, accounted for prematurely by our geodetic method" means?

We have removed this statement to shorten the related paragraph.

The original statement meant that our GP interpolation method does not always capture very rapid changes in elevation (see Fig. S3 in the SI), resulting in a redistribution of mass change during adjacent years for the largest surges.

Line 161-165: I suggest avoiding the expression of "geodetic study." Geodetic methods, i.e., repeat measurements of absolute surface height using remote sensing methods provide regional estimates compared to in situ measurements of mass changes, for example, using stakes. However, with the advent of remote sensing applications, including gravimetry, lidar, repeat photogrammetry, using this terminology has its limitations. Do you include gravimetry-based estimates, such as GRACE, for example?

We agree with referee #3 that "*geodetic*", although widely used by the DEM glacier community, can lead to confusion. For clarity, we have now replaced "*geodetic*" by "*DEM-based*" at all instances.

Lines 172-173: an additional map with a different map projection (polar stereographic?) wouould be needed to show the Antarctic mass loss pattern.

We agree with referee #3 in that a Polar Stereographic projection for Antarctica would look more familiar to some readers. We have strived to show more than 90% of glacier surfaces of region 19 (Antarctic and Subantarctic) on Fig. 2 and Fig. 4A with a legend that dissociates West and Peninsula from East Antarctica. Given the stringent space limitation of the journal, we think that this display is sufficient to illustrate the point highlighted by our statement at l.172-173.

Lines 184-197: this section, describing the temporal evolution/acceleration of mass loss could be a highlight of the manuscript. Instead, it is very vaguely written and illustrated. In particular, the meaning of including two percentages for the different regions, one for mass loss rate increase and another one for regional increase in thinning rate, is not clear and needs further clarification.

We have modified this section following referee #3 comments. We now summarize key values for specific regions:

"Elsewhere on Earth, glacier thinning accelerated. The combined mass loss of accelerating regions increased from 148 ± 20 Gt yr-1 in 2000-2004 to 247 ± 20 Gt yr-1 in 2015-2019. Two-thirds of this increased loss derives from three regions: Alaska (38%), High Mountain Asia (19%), and Western Canada and US (9%). Glaciers in the latter region experienced a fourfold increase in thinning rates. Most notably, glaciers in Northwestern America are responsible for nearly 50% of the accelerated mass loss. The widespread and strong increase of thinning of High Mountain Asian glaciers brought a large sub-region of sustained thickening in central-western Asia down to a generalized thinning in the late 2010s (Extended Data Fig. 7), suggesting the end of the so-called Karakoram anomaly[10]. Smaller glacierized regions also underwent strong, sometimes drastic acceleration of thinning. New Zealand, for example, shows a record 1.52 ± 0.50 m yr-1 thinning rate in 2015-2019 , which is a nearly sevenfold increase compared to 2000-2004.".

Line 198-208: this paragraph is very hard to follow and the statements need better explanations and illustrations.

We have modified the paragraph now presenting more illustrations:

"Analysis of climate data reveals that much of the regional patterns of mass change uncovered by our resolved estimates are consistent with large-scale, decadal changes in annual precipitation and temperature (Fig. 4, Extended Data Fig. 7). Strong dipoles that reflect concordant spatial patterns between precipitation change and mass change are observed notably in Northwestern America, southern Greenland Periphery and the Southern Andes. The southern Andean dipole is consistent with the mega-drought[31] of the 2010s that drove increased glacier mass loss in the Central Andes. In the Coast Mountains of Western Canada and in southeast Alaska, glaciers were severely deprived of precipitation that instead benefited neighbouring regions of central Alaska and continental US, correspondingly showing either stable or reduced mass loss. The North Atlantic anomaly coincided with cool, wet conditions of the last decade. Weaker dipoles can also be observed within the European Alps or Scandinavia. In both regions, glacier thinning slightly accelerated in the northeast and decelerated in the southwest.".

Line 211: I expect that changes are related to the warming of the atmosphere, not the troposphere.

We have changed to "*atmosphere*".

Line 221: please elaborate on the meaning of complete. Does it mean that no additional surface elevation change is available and can be fused with the record presented here?

We have replaced the original sentence to clarify the statement altogether:

"*Benefiting from the nearly complete spatial coverage afforded by ASTER stereo-imagery, our global estimate of recent glacier mass change (-274 ± 18 Gt yr^{-1} for the 2006-2015 IPCC reference period) shows strongly reduced uncertainties compared to the latest IPCC report*[4] *(-278 ± 226 Gt yr^{-1}) and a recent global study*[21] *(-335 ± 144 Gt yr^{-1}).*".

Here the later estimate cited is from Zemp et al. (2019).

We thank referee #3 for his input, which has improved the clarity of our statements, the traceability of our methods, and increased the exhaustivity and visualization of the mass change record presented throughout the article.

References for the answers to editor and referees (alphabetical)

Ablain, M. et al. Uncertainty in satellite estimates of global mean sea-level changes, trend and acceleration. Earth System Science Data 11, 1189–1202 (2019)

Bamber, J. L., Westaway, R. M., Marzeion, B. & Wouters, B. The land ice contribution to sea level during the satellite era. Environ. Res. Lett. 13, (2018)

Bevis, M. et al. Accelerating changes in ice mass within Greenland, and the ice sheet's sensitivity to atmospheric forcing. Proc. Natl. Acad. Sci. U. S. A. 116, 1934–1939 (2019)

Braun, M. H. et al. Constraining glacier elevation and mass changes in South America. Nat. Clim. Chang. (2019) doi:10.1038/s41558-018-0375-7

Brun, F., Berthier, E., Wagnon, P., Kääb, A. & Treichler, D. A spatially resolved estimate of High Mountain Asia glacier mass balances, 2000-2016. Nat. Geosci. 10, 668–673 (2017)

Ciracì, E., Velicogna, I. & Swenson, S. Continuity of the Mass Loss of the World's Glaciers and Ice Caps From the GRACE and GRACE Follow-On Missions. Geophys. Res. Lett. 47, 226 (2020)

Gardner, A. et al. A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009. Science 340, 852–857 (2013)

Huss, M. Density assumptions for converting geodetic glacier volume change to mass change. The Cryosphere 7, 877–887 (2013)

IMBIE Team. Mass balance of the Greenland Ice Sheet from 1992 to 2018. Nature 579, 233–239 (2020)

Kulp, S. A. & Strauss, B. H. New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding. Nat. Commun. 10, 4844 (2019)

Nerem, R. S. et al. Climate-change-driven accelerated sea-level rise detected in the altimeter era. Proc. Natl. Acad. Sci. U. S. A. 115, 2022–2025 (2018)

Nuimura, T., Fujita, K., Yamaguchi, S. & Sharma, R. R. Elevation changes of glaciers revealed by multitemporal digital elevation models calibrated by GPS survey in the Khumbu region, Nepal Himalaya, 1992-2008. J. Glaciol. 58, 648–656 (2012)

Pörtner, H. O. et al. IPCC Special Report on the Ocean and Cryosphere in a Changing Climate. IPCC Intergovernmental Panel on Climate Change: Geneva, Switzerland (2019)

Shean, D. E. et al. A Systematic, Regional Assessment of High Mountain Asia Glacier Mass Balance. Front. Earth Sci. 7, 435 (2020)

Velicogna, I. et al. Continuity of Ice Sheet Mass Loss in Greenland and Antarctica From the GRACE and GRACE Follow-On Missions. Geophys. Res. Lett. 47, L11501 (2020)

Willis, M. J., Melkonian, A. K., Pritchard, M. E. & Rivera, A. Ice loss from the Southern Patagonian Ice Field, South America, between 2000 and 2012. Geophys. Res. Lett. 39, (2012)

Wouters, B., Gardner, A. S. & Moholdt, G. Global Glacier Mass Loss During the GRACE Satellite Mission (2002-2016). Front. Earth Sci. 7, 535 (2019)

Zemp, M. et al. Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016. Nature 568, 382–386 (2019)

Reviewer Reports on the First Revision:

Referee #1 (Remarks to the Author):

I apologize that this review is a little late.
The flow of the paper is now much more understandable.

I see no large revisions needed for publication of this manuscript.

Please explain the rationale for splitting RGI regions in the diagrams but not the tables? Please be explicit in the caption of figure 1 about what the numbers in the circles are.

Referee #2 (Remarks to the Author):

Review of "Accelerated global glacier mass loss in the early twenty-first century" by Hugonnet et al.

This is the second round of review, and I'd like to thank the authors for the extensive work they've put into revising this paper. The authors have satisfactorily addressed most of my comments, and have put together a very extensive online dataset. I only have some minor points left, which are listed below.

Page 3: 'Intergovernmental Panel on Climate Change (IPCC)': I'd refer to the specific report (SROCC) in this case, instead of the IPCC as a whole.

Page 4: While many of the time series have been extended, I still don't fully agree with the comparisons with the Greenland and Antarctic Ice Sheet contributions. The problem is that the comparison still involves different time spans, since both IMBIE studies do not cover 2019. Both Greenland and Antarctica mass changes vary considerably from year to year, so the comparison is still a bit unfair. However, I'd argue that this manuscript does not actually need this explicit comparison between the relative size of the glacier and ice sheet contribution, since I don't think that anybody in the sea-level community would consider glaciers as 'not important' or so.

One other point that has not been addressed yet is the embedding of the observed acceleration over 2000-2019 in longer-term estimates. I think the authors have mis-understood my question in round one. The question I had, was how to interpret the observed acceleration in the short record, given the large multi-decadal variations in glacier mass change. Parkles&Marzeion (2019), Marzeion et al. (2015), for example, argue that the current rate of glacier melt is smaller than the glacier melt in the ~1930s. That implies that the observed acceleration could very well be a result of long-term variability, and that it does not necessarily point at a forced response to anthropogenic forcing. In other words, multi-decadal variability in glacier mass loss can show up as a (significant) acceleration in a 20-year record. The authors do not make the explicit claim that this acceleration is due to anthropogenic forcing or beyond natural variations, but it may be a good idea to spell this out and mention that variability, which is thoroughly discussed in the manuscript, can have a serious impact on the observed accelerations, and that multi-decadal variations in global glacier mass loss have occurred before.

Figure 1: add to the legend that the numbers in each circle refer to the mass change.

Data supplement: I really like the completeness of the new data supplement. My only comment is that units are missing. A list of the used units in either the column header or in a readme file would be helpful.

References:

Marzeion, B., Leclercq, P. W., Cogley, J. G., & Jarosch, A. H. (2015). Brief Communication: Global reconstructions of glacier mass change during the 20th century are consistent. The Cryosphere, 9(6), 2399–2404. https://doi.org/10.5194/tc-9-2399-2015

Parkes, D., & Marzeion, B. (2018). Twentieth-century contribution to sea-level rise from uncharted glaciers. Nature, 563(7732), 551–554. https://doi.org/10.1038/s41586-018-0687-9

Referee #3 (Remarks to the Author):

The authors have done an excellent job in addressing the questions and issues raised by the reviewers. I have already reviewed the first version of this manuscript, and therefore I don't repeat my evaluation of the key results. In summary, this is an important and timely study that uses a state-of-the-art method to improve and fuse the information from the best remote sensing observations of Earth's glaciers to derive a 20-year long time series of elevation and mass balance change history spanning 2000-2019. The resolution of the new data set and its global nature will enable the investigation of the relationship between climate forcing and glacier changes globally, a critically important goal for improving the understanding of underlying processing and predictions. The description of the methodology, particularly the spatiotemporal interpolation, is now detailed and transparent enough to enable reproducing the results, and the study provides robust estimates of the errors.

However, the lack of clarity about spatiotemporal scales at which grid cell/pixel, glacier, and regional-scale elevation changes are computed and will be distributed remains a concern. According to the manuscript, grid cell-scale elevation change time series are calculated for 30 by 30-meter grid cells with monthly resolution (lines 407 in Method and 281 in Supplementary information). However, extended data Fig. 3 provides an example for a 100 by 100 m grid cell. I guess that glacier and regional scale estimates are also calculated monthly, or at least with finer resolution than annual (see Fig. S7). However, I could not find a definitive description. Although elevation and mass changes are derived in a fine spatiotemporal scale, most results reported in the manuscript are 5-year or longer average rates (e.g., Figs. 4, Extended Table 1). I appreciate the authors' argument that short-term changes are prone to errors and spatial correlations (lines 597-605). However, their detailed results, such as comparing region-scale estimates with the published results shown in Fig. S7, indicate that the 5-year averaging might be too conservative. Some time series reveal short-term variations that seem to agree with results from other methods (e.g., Iceland (06)). Unfortunately, it is hard to judge the agreement's robustness, as the annual uncertainties for the geodetic mass change, although mentioned in the figure caption, are not shown in the figure. I strongly recommend that the authors include monthly resolution time series in the data distribution, enabling users to derive their aggregated (longer-term) products or apply time-varying firn-compaction model estimates, for example.

Unfortunately, its current structure makes it complicated to follow the manuscript. There are three sets of figures, in the main article, in the extended data, and in the supplementary information. Moreover, understanding the Method section requires to read the Supplementary information first. It is especially true for the sections about the Elevation time series and the Validation of elevation time series (lines 428-496). The method section (including figures) should be a standalone description with the supplementary information providing details for those interested in the study's intricacies. Based on Nature's instructions to authors, the Supplementary information should not include figures, and a maximum of 10 figures/tables are allowed in the Method. Complying with these requirements (if they apply) would require the authors to reduce the number of figures by eight, resulting in a more concise presentation. Finally, the extended data figures do not appear in the text in the order they are numbered, and some, e.g., Extended Data Fig. 1, are not even mentioned in the text.

Detailed comments:

Line 40: individually-resolved estimates – does it refer to time series for individual pixels or individual glaciers or something else?

Line 67: this statement is misleading. All glaciers have been frequently covered by optical and surface elevation imaging. However, accurate elevation changes were only derived from the stereo imaging and radar data for a relatively small percentage of the glaciers.

Line 77: there is no reference to Extended Data Fig. 1. The first figure that is referred to is Extended Data Figure 2. I expect that all figures are described in the text, and they are labeled in the order they appear.

Line 78: Changes in glacier elevations are traditionally determined not only from DEMs but also from altimetry and in situ measurements. The statement should reflect this.

Line 101: consider replacing "exceeding five years" with "five years or longer".

Line 115-118: contrary to the statement here, estimates of mass loss or thinning rate accelerations are not given in Fig. 1 or Table 1. The calculation of glacier accelerations is mentioned in the Methods, so a reference to the Methods would be most helpful.

Line 118-119: was the temporal evolution of mass loss/thinning rate acceleration determined from the 5-year averages in Extended Table 1? Does the statement here refer to the global glacier results or without ice sheet peripheral glaciers? A short explanation could be included in the Methods section.

Line 120: the expression "beyond the periphery of ice sheets" is ambiguous. Does it refer to the "Total excluding regions 05 and 19" in Extended Table 1? If yes, why are the rates and errors different? Are the rates annual or 5-year averages or, as described in the text, annual rates?

Line 124: Please correct the estimate from Ciraci et al., 2020. It is 50±20 Gt/yr and not 50±40 Gt/yr as quoted here.

Line 125-129: As this manuscript (and lots of other work) show, ice-sheet and peripheral mass losses change in a complicated way in time, rather than being simple quadratic functions. While I value the attempt to untangle the peripheral glaciers and the ice-sheet mass-loss histories by comparing published results with the new estimates, the interpretation presented here needs improvement.

Line 136: Extended Fig. 7 doesn't seem to provide useful information about the surging glaciers in Svalbard. Was the intention to refer to Table 3, region 7?

Lines 150 and 152: According to Fig. 2, the northernmost Arctic region appears to include the northern Greenland periphery, but not the southern Greenland peripheral glaciers, which are listed in the southern Arctic region (lines 152, 207). However, Extended Table 1 includes only "Greenland Periphery, Region 5" without a division to north and south. Is it intentional or an oversight? A division to northern and southern Greenland peripheries could strengthen the discussion about northern and southern Arctic regions.

Line 176: I assume that the "peculiar" surface elevation change pattern refers to the slow-down of thinning in East Greenland. A more specific description might be helpful. The radar altimetry results in the IMBIE-2 (Shepherd et al., 2020) are from Sørensen, L.S., Simonsen, S.B., Forsberg, R., Khvorostovsky, K., Meister, R., Engdahl, M.E., 2018. 25 years of elevation changes of the Greenland Ice Sheet from ERS, Envisat, and CryoSat-2 radar altimetry. Earth and Planetary Science Letters 495, 234–241. I assume that the authors did not include the original reference because of the limit on references in the main paper. However, it would be nice if they could acknowledge the Sørensen paper – perhaps in the supplemental material?

Lines 178-183: The description of the Antarctic Ice Sheet mass loss is confusing. The East Antarctic Ice Sheet (EAIS) mass gain is mostly due to the mass-gain of the slow-moving high elevation region of EAIS, while the marine-terminating glaciers are mainly using mass. Many recent papers describe this pattern, and Schröder et al., 2019 could be a good start. Unfortunately, it is difficult to discern the Antarctic elevation change pattern in Fig. 2 due to the need to present the results in a familiar global projection. Did the authors detect glacier thickening in East Antarctica? It might be of value to present results for the different regions of Antarctica in the Supplement. This would allow the authors to show the results around the WAIS and EAIS as well as the decelerating thinning around the Antarctic Peninsula.

Lines 184: please refer to Extended Table 1 for thinning rates in Iceland and Scandinavia. The thinning rate errors in the text are slightly different from the errors in the table, a discrepancy that needs to be reconciled.

Line 189: the relevance of Bevis et al., 2019 to the North Atlantic anomaly would need more explanation.

Line 191-192: the meaning of an "accelerating region" should be defined. Are these the regions that show higher thinning rates in 2015-2019 than 2000-2004 in Extended Table 1?

Line 195: is northwestern America the same as Western Canada and US, region 02?
Line 207: as mentioned before, the southern Greenland periphery is not presented as a separate region. Maybe a distinction between northern and southern Greenland peripheries would be helpful?

Figure 2: there seems to be a labeling issue for the Low Latitudes (16), and Antarctic and subantarctic (19) as the labeled regions' descriptions do not agree with the map.

Figure 3: In addition to 5-year rates, the chart in the right presents annual reconstructions. However, the authors stated (lines 598-599) that periods shorter than five years are affected by temporal autocorrelation and argue against using short-term estimates. So, are the annual estimates robust enough to be considered? Moreover, results from other studies, e.g., Zemp et al., 2019, which included global solutions with a high temporal resolution, are only represented with a

long-term average rate here. Are those higher temporal resolution global data sets not available for comparison? Please clarify. Finally, for citations in this figure, I suggest using both the reference number (according to the reference list in the manuscript) and the citation (first author et al., date). This would help the reader connect the references in the text (numbered) and figures (first author et al., date).

Lines 391-392: the doubling of the glacier area (from 1888 to 3516) only increased the area by 30 square km. Is this correct?

Line 407: does the error refer to the error map of TanDEM-X?

Line 463: what is a vertically close elevation observation?

Line 476: ILAKD1B refers to elevations collected in Alaska by OIB, while the text implies laser elevations worldwide. Were OIB laser altimetry elevations outside Alaska used in the study?

Line 546: incomplete sentence – please correct.

Line 579: The supplementary table that compares regional mass change results with regional studies is missing from the manuscript.

Extended data fig. 1: not mentioned in the text

Extended data fig. 2: the original number of the different DEM types is only shown in the figure. I suggest mentioning these numbers in the figure caption or in the manuscript text.

Lines 840-842: I assume that Fig. 3(a) is a global result, while the rest of the figure appear to refer to a specific time series. Please clarify.

Extended data fig. 6: this figure is supposed to illustrate the great spatial resolution of the new reconstruction. However, the current image resolution is very course and doesn't serve this purpose. Rather than relying on image resolution in the final version of the manuscript, I recommend to include a few zoomed in regions to demonstrate the quality of the results.

Extended data fig. 7: the arrangement in this map is different from Fig. 2 that makes it difficult to navigate the different regional maps. I suggest to include an overview map with the tile outlines.

Line 921: please clarify the meaning of the time-evolving regional glacier areas, decreasing linearly for all regions.

Supplemental information:
Line 110: the referred figure should be extended data fig. 3(a) instead of fig. 2(A), I think.
Line 118: see comment for line 110
Line 121: the numbering of the equations is not continuous.
Line 288: should it refer to Fig S2?
Line 290: reference to Fig S3?
Line 469: the first bin in extended data fig.5 is 0.1 km, not 0.15 km
Line 528: what is an "unmapped tongue" and why are the results consistent with the corresponding glaciers?

References in review:

Schröder, L., Horwath, M., Dietrich, R., Helm, V., Van Den Broeke, M.R., Ligtenberg, S.R.M., 2019. Four decades of Antarctic surface elevation changes from multi-mission satellite altimetry. The Cryosphere 13, 427–449.

Referee #4 (Remarks to the Author):

Paper Overview

This paper provides a study into the rates and accelerations of mass loss in glaciers. As glaciers form a core constituent of global fresh water reserves, understanding their decay is central to understanding water availability as well as sea level rise.

Precise observations of mass loss in glaciers are scarce both spatially and temporally: glaciers are in remote regions with harsh climates. By leveraging high resolution elevation data primarily taken from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) the authors are able to estimate Digital Elevation Models (DEMs) covering glaciers over a 20 year period from the year 2000 through 2019. These elevation estimates can then be used to determine time series of glacial mass at locations for which no standard observations are available. The elevation estimates are validated using ICESat and ICEBridge observations, verifying the absence of spatiotemporal biases. Temporal differences of glacial mass estimates are then used to estimate glacial mass loss in terms of water-equivalent mass change (this is standard in hydrology).

The glacier mass loss estimates are then given a detailed treatment with special attention given to not only the mass loss, or the mass loss rate, but the acceleration of glacial mass loss. The impacts of glacial mass loss on trends in sea level rise, and the correlation of trends in precipitation and glacial mass loss are also covered in detail. One alarming takeaway and possibly most significant result is the finding that the mass loss in glaciers that are distinct from the Greenland and Antarctic ice sheets (GIS, AIS) is significantly outpacing the loss seen in either the GIS or AIS independently.

Strengths

The dominant strength in this paper is the detail given to the empirical study. Global scale trends are considered, such as contribution to sea level rise, but arguably more impressive is the detailed spatial analysis that examines distinct geographic regions providing estimates of mass loss, mass loss rates, and mass loss acceleration for each of them.

Additionally a sort of causal analysis is performed over these specifically examined geographic regions, connecting the estimated glacial mass trends with macro-level climate trends like temperature and precipitation. The inclusion of these considerations aids the paper in both expanding the paper’s breadth (by considering relevant climatological variables beyond glacial mass) and helping verify the glacial mass estimates made in the first place (by showing the estimates are consistent with other well studied trends in climate science). There are also detailed validation steps for the elevation estimates and uncertainties. The inclusion of multiple secondary sources to validate the elevation estimates helps somewhat in trusting the predictions made by the GP regression. Most compelling in the consideration of measurements outside of glacial mass is Figure 4, which makes immediately clear how the estimates of rate of glacial mass change line up changes in precipitation and temperature. The figures are well made and contain a rich amount of information.

Concerns

Covariance function selection:

My main concern is with the selection and estimation of the covariance function (kernel) for the Gaussian process, which will profoundly affect the elevation estimates and their uncertainties. Given the importance of the covariance function, and the otherwise detailed validation of procedures, this part of the paper seems relatively lacking and ad-hoc. The kernel used is PL+ESS+RQ+RBF+(PL*RQ). This is a very particular design choice, largely motivated to match a variogram. What other choices were tried, and how were they rejected? The linear kernel will enforce a continued long-range trend, which we know cannot persist. Why not another RBF component with a longer length-scale? Or why not use other less-smooth kernels like the Matern kernel? What hyperparameters did you learn (such as the length-scales for the various components) and are they interpretable in this application?

In general, estimation of covariance functions by variograms tends to be much less precise than using the marginal likelihood of the Gaussian process. Why not try this approach for learning the hyperparameters of your kernel composition, and compare? Why not try automatic kernel learning approaches, such as spectral mixture kernels (with appropriate initialization techniques), in conjunction with marginal likelihood estimation? In the supplement, the reason given for preferring the variogram approach is "Our objective is to model variograms with characteristics representative of many pixels at once, and to apply these variograms directly in the regression. The rationale behind this approach is to mitigate the sparse sampling of elevations in time at the pixel scale by utilizing the repeat spatial coverage of our observations." But this objective can also be achieved (and likely improved upon) by following a marginal likelihood approach.

Since uncertainty estimation is crucial in this application, why not also place distributions over the kernel hyperparameters, and then perform Bayesian marginalization over these distributions? Surely by not doing this, you will be underestimating your uncertainty? What does the marginal likelihood look like a function of these parameters? How well determined are they by the data?

What considerations do you put into selecting a mean function for the Gaussian process, and how does that interact with your choice of covariance function? Why not consider more flexible or more informative prior mean functions?

Implementing the above suggestions would significantly increase my confidence in the Gaussian process estimates. At the minimum, it would be good to see a detailed study of how this kernel was decided upon, including a description of what other configurations were tried and why they were rejected.

Scalability:
In many cases downsampling is applied for computational reasons with the Gaussian processes. But this is not necessary with modern advances in GP scalability, and loses information that could be very helpful in using the marginal likelihood for kernel estimation (as described above). It is now common to use Krylov subspace methods, such as linear conjugate gradients, and stochastic Lanczos quadrature, to
scale even exact Gaussian process methods to problems with millions of points, through GPU parallelization. These methods are implemented in popular packages such as GPyTorch. Given the spatiotemporal structure of your problem, you can also likely use Kronecker methods to significantly scale your approach without making approximations, since your inputs are probably on close to a multidimensional (fully connected) lattice (can be expressed as a Cartesian product of vectors). The idea would be to create a kernel which is a product of 1D kernels (for example with the form you have chosen already) operating separately on each input dimension. Given the input structure of your problem, the covariance matrix would then be a Kronecker product of much smaller matrices. More information can be found (e.g., http://mlg.eng.cam.ac.uk/pub/pdf/Saa11.pdf), and there are many extensions to these techniques for missing points, or using virtual grids that cover the range of your raw inputs.

In any case, the downsampling is likely not necessary, and will be hurting the performance of your approach, and limiting your ability to do automatic kernel learning, which would be a very compelling alternative to the hand crafted kernel you are using.

Outliers:
In the supplementary material there is extended discussion of filtering and diminishing the effects

of outliers. Outside of high fidelity to the secondary data sources such as ICESat and ICEBridge, how can we be sure that important structure in the data is not being removed in enforcing insensitivity to outliers? It seems possible that in procedures such as fitting variograms based on median residuals (rather than mean residuals) would produce overly smooth estimates of elevation, which is not uncommon in other spatial estimates of climatological quantities.
Key High-Level Concern:
My high level primary concern is the treatment of the glacial mass estimates as data. Historical data are used to make and verify Gaussian process based estimates of glacier mass using elevation, but in many sections the estimates are treated interchangeably with recorded data. The results are at least not explic- itly stated as estimates throughout, and are presented as though the estimated glacial mass is just a noisy measurement.
One place in the paper this becomes a concern is in figure 3. In some cases there is substantial disagreement between the estimates in this work and the estimates made in other works which is mentioned in the discussion. However, discussing the estimates as fact in other cases does lead to concern when points in figure 3 indicates other sources may disagree with the findings.

Minor:

It would be more accurate to refer to your uncertainty intervals are “95% credible sets” rather than “confidence intervals”.

Author Rebuttals to First Revision (please note that the authors have quoted the reviewers in black text and responded in blue text):

Referee #1 (Remarks to the Author):

I apologize that this review is a little late.
The flow of the paper is now much more understandable.

I see no large revisions needed for publication of this manuscript.

We thank again referee #1 for his/her valuable contributions, especially to improve the introduction, conclusions, and the general flow of the paper during the first round.

Please explain the rationale for splitting RGI regions in the diagrams but not the tables?

The rationale for splitting some regions (on Figure 1, only) is detailed in the figure caption: “*Regions 2, 5, 9, 17 are further divided to illustrate contrasted temporal patterns.*”.

We report estimates for the 19 first-order RGI regions in Extended Data Table 1 for consistency with the inventory and previous studies.

Please be explicit in the caption of figure 1 about what the numbers in the circles are.

We have added to the caption of Figure 1: “*Mass change rates larger than 4 Gt yr*$^{-1}$ *are printed in blue inside the disk*.”.

Referee #2 (Remarks to the Author):

Review of “Accelerated global glacier mass loss in the early twenty-first century” by Hugonnet et al.

This is the second round of review, and I’d like to thank the authors for the extensive work they’ve put into revising this paper. The authors have satisfactorily addressed most of my comments, and have put together a very extensive online dataset. I only have some minor points left, which are listed below.

We thank referee #2 for their additional comments that further improved the comparison with the ice sheets and the physical interpretation of the observed glacier acceleration.

Page 3: ‘Intergovernmental Panel on Climate Change (IPCC)’: I’d refer to the specific report (SROCC) in this case, instead of the IPCC as a whole.

We now refer specifically to the SROCC report: “*...identified as a critical research gap by the Special Report on the Ocean and Cryosphere in a Changing Climate (SROCC) of the Intergovernmental Panel on Climate Change (IPCC)*[4].

Page 4: While many of the time series have been extended, I still don’t fully agree with the comparisons with the Greenland and Antarctic Ice Sheet contributions. The problem is that the comparison still involves different time spans, since both IMBIE studies do not cover 2019. Both Greenland and Antarctica mass changes vary considerably from year to year, so the comparison is still a bit unfair. However, I’d argue that this manuscript does not actually need this explicit comparison between the relative size of the glacier and ice sheet contribution, since I don’t think that anybody in the sea-level community would consider glaciers as ‘not important’ or so.

With agreement from the editor, we decided to retain this comparison. It is our view that separating the losses from the Antarctic and Greenland ice sheets and those of their peripheral glaciers is one of importance that has been overlooked in many recent studies. The high resolution of our study provides an opportunity to deal with this inconsistency that hampers more rigorous assessments for glaciers, ice sheets and sea-level rise.
To further improve the comparison, we changed the end date to 01/01/2019 to exactly match the period considered by the IMBIE team for the Greenland Ice Sheet estimate. We specify in the caption of Table 1 that the IMBIE Antarctic Ice Sheet estimate has an earlier end date of 06/2017. Note that this should not be problematic as recent studies show little variability in the Antarctic ice sheet mass loss rates during the period 06/2017-01/2019 (Velicogna et al., 2020).

One other point that has not been addressed yet is the embedding of the observed acceleration over 2000-2019 in longer-term estimates. I think the authors have mis-understood my question in round one. The question I had, was how to interpret the observed acceleration in the short record, given the large multi-decadal variations in glacier mass change. Parkles&Marzeion (2019), Marzeion et al. (2015), for example, argue that the current rate of glacier melt is smaller than the glacier melt in the ~1930s. That implies that the observed acceleration could very well be a result of long-term variability, and that it does not necessarily point at a forced response to anthropogenic forcing. In other words, multi-decadal variability in glacier mass loss can show up as a (significant) acceleration in a 20-year record. The authors do not make the explicit claim that this acceleration is due to anthropogenic forcing or beyond natural variations, but it may be a good idea to spell this out and mention that variability, which is thoroughly discussed in the manuscript, can have a serious impact on the observed accelerations, and that multi-decadal variations in global glacier mass loss have occurred before.

The referee raises a good point. We added a sentence at the end of the Discussion (section "Drivers of temporal variability") considering that the observed acceleration of glacier mass loss might arise from multiple sources. :
"*Previous studies*[35] *have indicated large multi-decadal variation in rates of glacier mass change across the 20th century, implying that some of the acceleration we observe could fall within the range of natural variability. Nonetheless, the strong concordance to the increase in global surface temperatures suggests, indirectly, a considerable response to anthropogenic forcing.*" Here we cited: Parkes and Marzeion (2018).

Figure 1: add to the legend that the numbers in each circle refer to the mass change.

We have added to the caption of Figure 1: "*Mass change rates larger than 4 Gt yr*$^{-1}$ *are printed in blue inside the disk*.".

Data supplement: I really like the completeness of the new data supplement. My only comment is that units are missing. A list of the used units in either the column header or in a readme file would be helpful.

We have added a README file specifying the units of the provided datasets.

References:

Marzeion, B., Leclercq, P. W., Cogley, J. G., & Jarosch, A. H. (2015). Brief Communication: Global reconstructions of glacier mass change during the 20th century are consistent. The Cryosphere, 9(6), 2399–2404. https://doi.org/10.5194/tc-9-2399-2015

Parkes, D., & Marzeion, B. (2018). Twentieth-century contribution to sea-level rise from uncharted glaciers. Nature, 563(7732), 551–554. https://doi.org/10.1038/s41586-018-0687-9

Referee #3 (Remarks to the Author):

The authors have done an excellent job in addressing the questions and issues raised by the reviewers. I have already reviewed the first version of this manuscript, and therefore I don't repeat my evaluation of the key results. In summary, this is an important and timely study that uses a state-of-the-art method to improve and fuse the information from the best remote sensing observations of Earth's glaciers to derive a 20-year long time series of elevation and mass balance change history spanning 2000-2019. The resolution of the new data set and its global nature will enable the investigation of the relationship between climate forcing and glacier changes globally, a critically important goal for improving the understanding of underlying processing and predictions. The description of the methodology, particularly the spatiotemporal interpolation, is now detailed and transparent enough to enable reproducing the results, and the study provides robust estimates of the errors.

We thank referee #3 for his/her thorough 2nd review which allowed us to further clarify several statements and improve the presentation of methods throughout the paper and in relation to the Supplementary Information.

However, the lack of clarity about spatiotemporal scales at which grid cell/pixel, glacier, and regional-scale elevation changes are computed and will be distributed remains a concern. According to the manuscript, grid cell-scale elevation change time series are calculated for 30 by 30-meter grid cells with monthly resolution (lines 407 in Method and 281 in Supplementary information). However, extended data Fig. 3 provides an example for a 100 by 100 m grid cell. I guess that glacier and regional scale estimates are also calculated monthly, or at least with finer resolution than annual (see Fig. S7). However, I could not find a definitive description.

Thank you for pointing this out. To improve clarity, we added the spatial and temporal resolution for the different data and estimates directly on Extended Data Fig. 1.
For the DEMs, the information has been added to the main text: "*We use modern photogrammetry techniques and specifically-developed statistical methods to generate and bias-correct nearly half a million Digital Elevation Models (DEMs) at 30 m horizontal resolution.*".
It is also present both in the Methods (line 407: "*For all DEMs bilinearly resampled to 30 m, ...*", line 432: "*... downsampled to 100 m to decrease computing time*", line 438: "*... derive continuous elevation time series at a monthly time step independently for each of the 400 millions pixels*") and in the Supplementary Methods (Section 2.6-2.7: "*at a monthly temporal resolution from January 1st, 2000 to January 1st, 2020 for each pixel at a posting of 100 m*").
We have clarified the information at each step by simplifying and relocating the sentences.
Additionally, this information along with other useful knowledge for exploiting the dataset, are described in the README available through the Data and Code availability statements (https://github.com/rhugonnet/ww_tvol_study).

Although elevation and mass changes are derived in a fine spatiotemporal scale, most results reported in the manuscript are 5-year or longer average rates (e.g., Figs. 4, Extended

Table 1). I appreciate the authors' argument that short-term changes are prone to errors and spatial correlations (lines 597-605). However, their detailed results, such as comparing region-scale estimates with the published results shown in Fig. S7, indicate that the 5-year averaging might be too conservative. Some time series reveal short-term variations that seem to agree with results from other methods (e.g., Iceland (06)). Unfortunately, it is hard to judge the agreement's robustness, as the annual uncertainties for the geodetic mass change, although mentioned in the figure caption, are not shown in the figure.

We increased the width of error bars on Fig. S7, as those were indeed difficult to distinguish. We agree with referee #3 that the temporal resolution is sometimes better than five years. This is especially true in Iceland, where our data coverage is the highest (Extended Data Table 2). Thanks to this good repeat coverage, we can isolate part of the “atypical” extreme annual glacier changes independently observed in Iceland in the years 2010 (strong loss) and 2015 (small gain), reported for example by Belart et al. (2020) or Aðalgeirsdóttir et al. (2020). We however cannot fully capture the interannual variability due to a smoothing of our elevation changes in the GP regression, caused by the limited vertical precision of our DEMs.
This effect, and the varying temporal resolution of our results, are discussed in the Supplementary Discussion section “Time series comparison and temporal resolution”, that includes the following statement: “*Based on our volume change uncertainties, the temporal resolution at which volume changes are statistically significant at the regional scale (95% confidence interval <0.2 m yr^{-1}) is of 3-7 years depending on the spatial domain and temporal coverage.*”
Additionally, one should note that, under the assumptions of volume-to-mass conversion described by Huss (2013), we cannot confidently report mass change uncertainties for periods shorter than about five years (as noted in the caption of Fig. S7). An “exception” to this assumption is the global mass change signal which, as a sum of globally distributed and largely climatically independent regions, is less susceptible to the temporal autocorrelation of this density conversion factor.

I strongly recommend that the authors include monthly resolution time series in the data distribution, enabling users to derive their aggregated (longer-term) products or apply time-varying firn-compaction model estimates, for example.

We now provide time series with monthly resolution through the Data availability statement. In the Methods, Supplementary Methods, and the README of the Data availability statement, we stress that the seasonal estimates are subject to snow-cover-induced systematic errors (Section 3.3 of Supplementary Information, Fig S4-S5).

Unfortunately, its current structure makes it complicated to follow the manuscript. There are three sets of figures, in the main article, in the extended data, and in the supplementary information. Moreover, understanding the Method section requires to read the Supplementary information first. It is especially true for the sections about the Elevation time series and the Validation of elevation time series (lines 428-496). The method section (including figures) should be a standalone description with the supplementary information providing details for those interested in the study's intricacies. Based on Nature's instructions to authors, the Supplementary information should not include figures, and a maximum of 10 figures/tables are allowed in the Method. Complying with these requirements (if they apply) would require the authors to reduce the number of figures by eight, resulting in a more concise presentation. Finally, the extended data figures do not appear in the text in the order they are numbered, and some, e.g., Extended Data Fig. 1, are not even mentioned in the text.

We have further streamlined the Methods and SI to ensure the Methods section is "stand-alone" and independent from the SI. We have encapsulated all methodological steps within the main Methods and provide more details in the SI, aimed at expert readers only. We have changed the location of information between Methods and SI, notably added a paragraph on the filtering methods, improved the elevation time series section, and modified several linkages to improve the presentation.
For what the comments on the Figures are concerned, we followed the editor's advice and kept Figures and Tables in the Supplementary Information for expert readers.
Finally, Extended Data Fig. 1 is now referenced in the main text section "A spatiotemporally resolved estimation", in addition to another reference at the start of the Methods section.

Detailed comments:

Line 40: individually-resolved estimates – does it refer to time series for individual pixels or individual glaciers or something else?

We have changed to: "*highly-resolved*" for clarity.

Line 67: this statement is misleading. All glaciers have been frequently covered by optical and surface elevation imaging. However, accurate elevation changes were only derived from the stereo imaging and radar data for a relatively small percentage of the glaciers.

Good catch. We have reformulated: "*the uneven coverage of optical and radar surface elevation change estimations that account at most for 10% of the world's glaciers*[21]".

Line 77: there is no reference to Extended Data Fig. 1. The first figure that is referred to is Extended Data Figure 2. I expect that all figures are described in the text, and they are labeled in the order they appear.

As stated above, Extended Data Fig. 1 is now referenced in the main text, in addition to another reference that was at the start of the Methods section.

Line 78: Changes in glacier elevations are traditionally determined not only from DEMs but also from altimetry and in situ measurements. The statement should reflect this.

We modified this statement that aimed at focusing on the temporal approach: "*Changes in glacier elevation based on DEMs are traditionally quantified by differencing pairs of acquisitions from two distinct epochs.*"

Line 101: consider replacing "exceeding five years" with "five years or longer".

We decided to keep "*exceeding five years*" to avoid the use of three comparative adjectives in the same sentence ("*longer*", "*larger*", "*shorter*"), which could negatively impact clarity.

Line 115-118: contrary to the statement here, estimates of mass loss or thinning rate accelerations are not given in Fig. 1 or Table 1. The calculation of glacier accelerations is mentioned in the Methods, so a reference to the Methods would be most helpful.

We have removed the reference "*(Fig. 1, Table 1)*" and replaced it with "*(see Methods)*".

Line 118-119: was the temporal evolution of mass loss/thinning rate acceleration determined from the 5-year averages in Extended Table 1? Does the statement here refer to the global glacier results or without ice sheet peripheral glaciers? A short explanation could be included in the Methods section.

This information can be found in the Methods section entitled "Acceleration", which contains the following statement: "*Glacier mass change acceleration and its uncertainties were derived from weighted least-squares on the 5-year elevation and mass change rates (i.e., 2000-2004, 2005-2009, 2010-2014 and 2015-2019), propagating their related uncertainties as independent.*"

Line 120: the expression "beyond the periphery of ice sheets" is ambiguous. Does it refer to the "Total excluding regions 05 and 19" in Extended Table 1? If yes, why are the rates and errors different? Are the rates annual or 5-year averages or, as described in the text, annual rates?

We have changed "*beyond the periphery of ice sheets*" into "*excluding peripheral glaciers*" for clarity. The term "*peripheral glacier*" was defined in the previous paragraph.
The rates indeed refer to annual rates for the year 2000 and the year 2019, thus differing from the 5-year averages reported in Extended Data Table 1: "*... thinning rates nearly doubled from 0.36 ± 0.21 m yr-1 in 2000 to 0.69 ± 0.15 m yr-1 in 2019*".

Line 124: Please correct the estimate from Ciraci et al., 2020. It is 50±20 Gt/yr and not 50±40 Gt/yr as quoted here.

The study by Ciracì et al. (2020) reports 1-sigma uncertainties, we thus multiplied their uncertainties by a factor of two in the text and Figures (e.g., Fig. 3) for consistency with studies that use 2-sigma uncertainties (including our own study, as well as the ones of Wouters et al. (2019), Gardner et al. (2013), etc..).
The 2-sigma level of our uncertainties is specified at the start of the main text and, based on a comment from the editor, is now as well included in all relevant Figure and Table captions.

Line 125-129: As this manuscript (and lots of other work) show, ice-sheet and peripheral mass losses change in a complicated way in time, rather than being simple quadratic functions. While I value the attempt to untangle the peripheral glaciers and the ice-sheet mass-loss histories by comparing published results with the new estimates, the interpretation presented here needs improvement.

We agree with referee #3 that a complex temporal evolution cannot be represented by simple quadratic functions. When estimating the acceleration, time series for which quadratic functions poorly match the temporal evolution will create very large uncertainties in the trend (due to least-square uncertainty propagation), making acceleration estimates statistically insignificant. This is why recent estimates of acceleration of the Greenland Ice Sheet mass loss are not statistically significant (Velicogna et al., 2020), and why our estimate of mass loss acceleration for all glaciers (48 ± 16 Gt yr^{-1} per decade) has an uncertainty that is twice as large as the one that excludes peripheral glaciers (62 ± 8 Gt yr^{-1} per decade) despite being based on 5-year mass loss rates with similar uncertainties. The temporal evolution of the global estimate (including peripheral glaciers) doesn't match a quadratic evolution as well as the one excluding peripheral glaciers (partly due to a slight slowdown in 2015-2019), resulting in a larger acceleration uncertainty of ± 16 Gt yr^{-1}.
In summary, our results account for these effects and show statistically significant trends for glaciers. This, in turn, allows us to conclude that recent glacier mass loss accelerated, thus contributing to accelerated sea-level rise.

Line 136: Extended Fig. 7 doesn't seem to provide useful information about the surging glaciers in Svalbard. Was the intention to refer to Table 3, region 7?

Thank you for spotting this. We corrected the numbering to Extended Data Fig. 6 (instead of 7). Svalbard surges are visible on this figure and pointed out by arrows.

Lines 150 and 152: According to Fig. 2, the northernmost Arctic region appears to include the northern Greenland periphery, but not the southern Greenland peripheral glaciers, which are listed in the southern Arctic region (lines 152, 207). However, Extended Table 1 includes only "Greenland Periphery, Region 5" without a division to north and south. Is it intentional or an oversight? A division to northern and southern Greenland peripheries could strengthen the discussion about northern and southern Arctic regions.

For consistency with the used glacier inventory and previous studies, we chose to report results in Extended Data Table 1 only for the 19 first-order RGI regions. In the text and on Fig. 1, we further subdivide region 5 (Greenland Periphery) to better illustrate its contrasted spatial and temporal patterns. Although the corresponding values for northern/southern Greenland Periphery are not directly available in a Table, the reader can appreciate those contrasted changes in Fig. 1, Fig.2 and Extended Data Fig. 7.
We have added the data for subregions of Fig. 1 into a table available through the Data availability statement.

Line 176: I assume that the "peculiar" surface elevation change pattern refers to the slow-down of thinning in East Greenland. A more specific description might be helpful. The radar altimetry results in the IMBIE-2 (Shepherd et al., 2020) are from Sørensen, L.S., Simonsen, S.B., Forsberg, R., Khvorostovsky, K., Meister, R., Engdahl, M.E., 2018. 25 years of elevation changes of the Greenland Ice Sheet from ERS, Envisat, and CryoSat-2 radar altimetry. Earth and Planetary Science Letters 495, 234–241. I assume that the authors did not include the original reference because of the limit on references in the main paper. However, it would be nice if they could acknowledge the Sørensen paper – perhaps in the supplemental material?

We have added a more specific description focusing on eastern Greenland with a reference to Extended Data Fig. 7: “*..., particularly notable around the eastern Greenland subregions of mass gain in 2015-2019 (Extended Data Fig. 7).*”. We have added the reference to Sørensen et al. (2018) at this point of the main text.

Lines 178-183: The description of the Antarctic Ice Sheet mass loss is confusing. The East Antarctic Ice Sheet (EAIS) mass gain is mostly due to the mass-gain of the slow-moving high elevation region of EAIS, while the marine-terminating glaciers are mainly using mass. Many recent papers describe this pattern, and Schröder et al., 2019 could be a good start. Unfortunately, it is difficult to discern the Antarctic elevation change pattern in Fig. 2 due to the need to present the results in a familiar global projection. Did the authors detect glacier thickening in East Antarctica? It might be of value to present results for the different regions of Antarctica in the Supplement. This would allow the authors to show the results around the WAIS and EAIS as well as the decelerating thinning around the Antarctic Peninsula.

We detected slow thickening or stable mass loss (within uncertainty bounds) in East Antarctica. We added values within parentheses at this point of the text to clarify our statements: “*Western Antarctic glaciers substantially lost mass (-0.31 ± 0.07 m yr^{-1}) while those of East Antarctica slowly thickened (0.04 ± 0.04 m yr^{-1}). Ice masses surrounding the Antarctic Peninsula, representing 63% of the glacier area in the Antarctic and Subantarctic, experienced slow, decelerating thinning (-0.18 ± 0.05 m yr^{-1}) also captured by recent gravimetric surveys of the entire Peninsula*[25]*.*”.
We now provide additional estimates through the Data availability statement, including subregions of Antarctica, Greenland, all RGI second-order regions and also HiMAP (Hindu Kush Himalayan Monitoring and Assessment Program) subregions for High Mountain Asia, which are particularly of interest to many readers.

Lines 184: please refer to Extended Table 1 for thinning rates in Iceland and Scandinavia. The thinning rate errors in the text are slightly different from the errors in the table, a discrepancy that needs to be reconciled.

Now corrected. Thank you for catching this discrepancy.

Line 189: the relevance of Bevis et al., 2019 to the North Atlantic anomaly would need more explanation.

Bevis et al. (2019) describes the pattern of slowdown of the Greenland ice sheet mass loss after 2013, especially in the southwest area. We thus think it is a relevant reference at this location (suggested by referee #1). Note that, in this sentence, the reference is not referring to the “*North Atlantic anomaly*”, but to the Greenland ice sheet mass loss: “*Taken together, the slowdown in mass loss from these two regions, in addition to the one of peripheral glaciers of southeast Greenland Periphery*[30] *define a regional pattern that we refer to as the North Atlantic anomaly.*”. The reference number 30 cited here corresponds to Bevis et al. (2019).

Line 191-192: the meaning of an "accelerating region" should be defined. Are these the regions that show higher thinning rates in 2015-2019 than 2000-2004 in Extended Table 1?

Yes. We clarified by rephrasing as: “*Elsewhere on Earth, glacier thinning accelerated. The combined mass loss of these regions with increased loss escalated from...*”.

Line 195: is northwestern America the same as Western Canada and US, region 02?

Yes. We clarified by adding this information within parenthesis: “*Most notably, glaciers in Northwestern America (Alaska, Western Canada and US) are responsible for…*”.

Line 207: as mentioned before, the southern Greenland periphery is not presented as a separate region. Maybe a distinction between northern and southern Greenland peripheries would be helpful?

Following previous answers, we think that Fig. 1, Fig. 2, and Extended Data Fig. 7 are sufficient to illustrate the differences mentioned in the text, and that Extended Data Table 1 should be restricted to RGI first-order regions only for the sake of consistency. Estimates for these specific subregions are now made available through the Data availability statement.

Figure 2: there seems to be a labeling issue for the Low Latitudes (16), and Antarctic and subantarctic (19) as the labeled regions' descriptions do not agree with the map.

We have corrected the labeling and legend of Fig. 2.

Figure 3: In addition to 5-year rates, the chart in the right presents annual reconstructions. However, the authors stated (lines 598-599) that periods shorter than five years are affected by temporal autocorrelation and argue against using short-term estimates. So, are the annual estimates robust enough to be considered? Moreover, results from other studies, e.g., Zemp et al., 2019, which included global solutions with a high temporal resolution, are only represented with a long-term average rate here. Are those higher temporal resolution global data sets not available for comparison? Please clarify.

The global estimate is not affected much by temporal autocorrelation. Uncertainties in annual rates are thus provided for global estimates only. We have added a statement in the Methods section “Aggregation to global”: “*We report uncertainties in mass changes for periods shorter than five years solely for the global or near-global estimates (e.g. Fig. 3b) by assuming that the aggregation of largely independent RGI regions leaves limited temporal autocorrelation of density conversion factors.*”.
Unfortunately, the annual rates from Zemp et al. (2019) or Wouters et al. (2019) have large uncertainties (larger than the current panel size of Fig. 3b) which prevents a useful comparison to them (i.e. all mass change estimates would agree within these very large error bars). We thus do not compare to these time series. We have added a statement in the caption of Fig. 3: “*Annual rates of earlier studies are not shown due to large uncertainties.*”.

Finally, for citations in this figure, I suggest using both the reference number (according to the reference list in the manuscript) and the citation (first author et al., date). This would help the reader connect the references in the text (numbered) and figures (first author et al., date).

We have added the reference number for the citations directly in the legend of Fig. 3.

Lines 391-392: the doubling of the glacier area (from 1888 to 3516) only increased the area by 30 square km. Is this correct?

Yes, this is correct. This is due to the mapping of new very small glaciers (<1 km²) but also to the subdivision of existing glaciers into smaller entities.

Line 407: does the error refer to the error map of TanDEM-X?

Yes. We repeated “*TanDEM-X*” before “*height error map*” for clarity.

Line 463: what is a vertically close elevation observation?

We clarified by re-phrasing this sentence: “*This sum decomposes the differences of elevation observations with varying time lags into:…*”

Line 476: ILAKD1B refers to elevations collected in Alaska by OIB, while the text implies laser elevations worldwide. Were OIB laser altimetry elevations outside Alaska used in the study?

Yes, the IODEM3 dataset from Operation IceBridge, referenced at the same point in the text, covers many regions other than Alaska (see Table S1 in the Supplementary Information with the precise data coverage). Note that the ICESat dataset is also mentioned at the point, which is why the term “*worldwide*” is used: “*We retrieved all ICESat*[58] *(GLAH14) and IceBridge*[59,60] *(IODEM3 and ILAKS1B) laser and optical elevations intersecting glaciers worldwide...*”.

Line 546: incomplete sentence – please correct.

The conversion from Word to PDF in Nature’s Manuscript Center seems to have cropped out 3-4 lines of text here. Fortunately, they have not changed since the first version of the manuscript: “*For each glacier, we estimated an uncertainty in the area A based on a buffer*[79] *of 15 m corresponding to the typical resolution of the optical imagery*[33,80–82] *used to derive these outlines. Uncertainties in the area vary from about 0.1% of the area for large icefields (>1000 km²) to 50% of the area and above for small isolated glaciers (<0.1 km²).*”.

Line 579: The supplementary table that compares regional mass change results with regional studies is missing from the manuscript.

Following the editor’s recommendations, this Table was removed from the Supplementary Information and placed into a separate file provided as Supplementary Table 1.

Extended data fig. 1: not mentioned in the text

We added a reference to Extended Data Fig. 1 in the main text, in addition to the one at the start of Methods.

Extended data fig. 2: the original number of the different DEM types is only shown in the figure. I suggest mentioning these numbers in the figure caption or in the manuscript text.

We added the number of ArcticDEM and REMA DEM strips to the caption of Extended Data Fig. 2: "*67,986 ArcticDEM and 9,369 REMA strips are counted before co-registration to TanDEM-X.*".

Lines 840-842: I assume that Fig. 3(a) is a global result, while the rest of the figure appear to refer to a specific time series. Please clarify.

We added a clarification in the caption of Extended Data Fig. 3a,b : "*...estimated globally*". The subpanels c), d) and e) were specified as a single pixel example, as stated in the caption: "*..., illustrated here for a 100 m by 100 m pixel on the ablation area of Upsala*".

Extended data fig. 6: this figure is supposed to illustrate the great spatial resolution of the new reconstruction. However, the current image resolution is very course and doesn't serve this purpose. Rather than relying on image resolution in the final version of the manuscript, I recommend to include a few zoomed in regions to demonstrate the quality of the results.

We increased the zoom of several panels on Extended Data Fig. 6, except for Iceland and Svalbard, which we have displayed in full. To give a sense of the high resolution of our estimate, we have also added a zoomed panel for small glaciers in Coropuna, Peru. Note that Extended Data Figures are available online at full resolution.

Extended data fig. 7: the arrangement in this map is different from Fig. 2 that makes it difficult to navigate the different regional maps. I suggest to include an overview map with the tile outlines.

We added the following in the caption of Extended Data Fig. 7: "*Region labeling refers to that of Fig. 2.*", as was done for Fig. 4. We here chose to repeat the arrangement of Fig. 4 as it leaves less unused space, thus allowing larger visualization of the tiles within each region and for each successive 5-year period.

Line 921: please clarify the meaning of the time-evolving regional glacier areas, decreasing linearly for all regions.

We removed "*decreasing linearly for all regions*" and now refer to Methods for further details on glacier time-evolving areas.

Supplemental information:
Line 110: the referred figure should be extended data fig. 3(a) instead of fig. 2(A), I think.
Line 118: see comment for line 110

Both are now corrected.

Line 121: the numbering of the equations is not continuous.

We repeated Equation (1) of the main Methods thus keeping its original numbering. All other equations in the Supplementary Information are referred to as Equations S1-S15.

Line 288: should it refer to Fig S2?
Line 290: reference to Fig S3?

Both are now corrected.

Line 469: the first bin in extended data fig.5 is 0.1 km, not 0.15 km

Good catch. Extended Data Fig. 5 value was incorrectly rounded, this is now corrected.

Line 528: what is an “unmapped tongue” and why are the results consistent with the corresponding glaciers?

We have clarified our statement referring to the consistency between 20-year elevation changes observed on “stable terrain” and “glacier terrain” despite using different Gaussian Process kernels: *“Nonetheless, the linear estimation allows similar 20-year changes to be captured at the boundary of glacierized and stable terrain. For instance, unmapped debris-covered tongues treated as stable terrain show elevation changes consistent with the rest of the glacier.”*

References in review:
Schröder, L., Horwath, M., Dietrich, R., Helm, V., Van Den Broeke, M.R., Ligtenberg, S.R.M., 2019. Four decades of Antarctic surface elevation changes from multi-mission satellite altimetry. The Cryosphere 13, 427–449.

Referee #4 (Remarks to the Author):

Paper Overview

This paper provides a study into the rates and accelerations of mass loss in glaciers. As glaciers form a core constituent of global fresh water reserves, understanding their decay is central to understanding water availability as well as sea level rise.
Precise observations of mass loss in glaciers are scarce both spatially and temporally: glaciers are in remote regions with harsh climates. By leveraging high resolution elevation data primarily taken from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) the authors are able to estimate Digital Elevation Models (DEMs) covering glaciers over a 20 year period from the year 2000 through 2019. These elevation estimates can then be used to determine time series of glacial mass at locations for which no standard observations are available. The elevation estimates are validated using ICESat and ICEBridge observations, verifying the absence of spatiotemporal biases. Temporal differences of glacial mass estimates are then used to estimate glacial mass loss in terms of water-equivalent mass change (this is standard in hydrology).
The glacier mass loss estimates are then given a detailed treatment with special attention given to not only the mass loss, or the mass loss rate, but the acceleration of glacial mass loss. The impacts of glacial mass loss on trends in sea level rise, and the correlation of trends in precipitation and glacial mass loss are also covered in detail. One alarming takeaway and possibly most significant result is the finding that the mass loss in glaciers that are distinct from the Greenland and Antarctic ice sheets (GIS, AIS) is significantly outpacing the loss seen in either the GIS or AIS independently.

Strengths

The dominant strength in this paper is the detail given to the empirical study. Global scale trends are considered, such as contribution to sea level rise, but arguably more impressive is the detailed spatial analysis that examines distinct geographic regions providing estimates of mass loss, mass loss rates, and mass loss acceleration for each of them.
Additionally a sort of causal analysis is performed over these specifically examined geographic regions, connecting the estimated glacial mass trends with macro-level climate trends like temperature and precipitation. The inclusion of these considerations aids the paper in both expanding the paper’s breadth (by considering relevant climatological variables beyond glacial mass) and helping verify the glacial mass estimates made in the first place (by showing the estimates are consistent with other well studied trends in climate science). There are also detailed validation steps for the elevation estimates and uncertainties. The inclusion of multiple secondary sources to validate the elevation estimates helps somewhat in trusting the predictions made by the GP regression. Most compelling in the consideration of measurements outside of glacial mass is Figure 4, which makes immediately clear how the estimates of rate of glacial mass change line up changes in precipitation and
temperature. The figures are well made and contain a rich amount of information.

We thank referee #4 for the detailed and constructive comments on our paper. Below, we provide answers and additional information to his/her inquiries on the current approach for Gaussian Process regression, the treatment of outliers and the comparison to other studies.

Concerns
Covariance function selection:

My main concern is with the selection and estimation of the covariance function (kernel) for the Gaussian process, which will profoundly affect the elevation estimates and their uncertainties. Given the importance of the covariance function, and the otherwise detailed validation of procedures, this part of the paper seems relatively lacking and ad-hoc.

We agree with referee #4 that the covariance function is important. In the following, we provide point-by-point responses on how we proceeded in selecting this function.

First, we address the referee's remark on the sensitivity of the estimates to the covariance function and present an analysis demonstrating that this selection does not significantly affect our estimates of volume change and mass change. This low sensitivity is directly due to the dense and repeated observational coverage (on average, 39 independent elevation observations per pixel in 20 years) which leads to a limited influence of the covariance function on the interpolated 5-year, 10-year and 20-year estimates.

Sensitivity of the kernel parametrization of the covariance function
We have added Fig. S9 on the sensitivity to the kernel hyperparameters in the Supplementary Information (also shown on the next page).
We computed elevation time series for all glacierized pixels of two regions: Iceland, which has the densest temporal sampling, and Scandinavia, which has the sparsest temporal sampling (Antarctic and Subantarctic excluded). These regions were chosen as they are potentially the most sensitive of all regions due to (i) their nonlinear evolution during 2000-2019 (Extended Data Table 1), and (ii) their small size which implies strongly spatially correlated signals. To quantify this sensitivity, we varied the hyperparameters by an order of magnitude and computed the deviation to our reported estimates (Extended Data Table 1). Further details are available in the figure caption.
We interpret the sensitivity to the parameter Δt_{nl} (temporal scale of local linear trend) as the result of a slightly different "extrapolation" near the temporal boundaries of our period (2000 and 2019). This is why the difference is largest for longer periods (10- and 20-year). Although the "extrapolation" at the temporal boundaries is limited due to the dense temporal sampling of our data, they have been considered with attention and selected in relation to a physical interpretation (further discussed two answers down, in response to a specific comment from the referee). The sensitivity of the parameter ${\sigma_l}^2$ (variance of local, inter-annual change) shows less of an effect in Scandinavia than Iceland, likely due to a less dense coverage of interannual observations, especially near the temporal boundaries.

In summary, the sensitivity of our estimates to the Gaussian Process hyperparameters within an order of magnitude is very low (<3% absolute deviation) and well within our uncertainties (<30% of uncertainty range for volume change, <5% for mass change).

We added several of the preceding statements to a new section in the Supplementary Discussion entitled “*Sensitivity to the Gaussian Process hyperparameters*”, where Fig. S10 is referenced. We also added related statement on sensitivity in Section 2.5 of the SI.

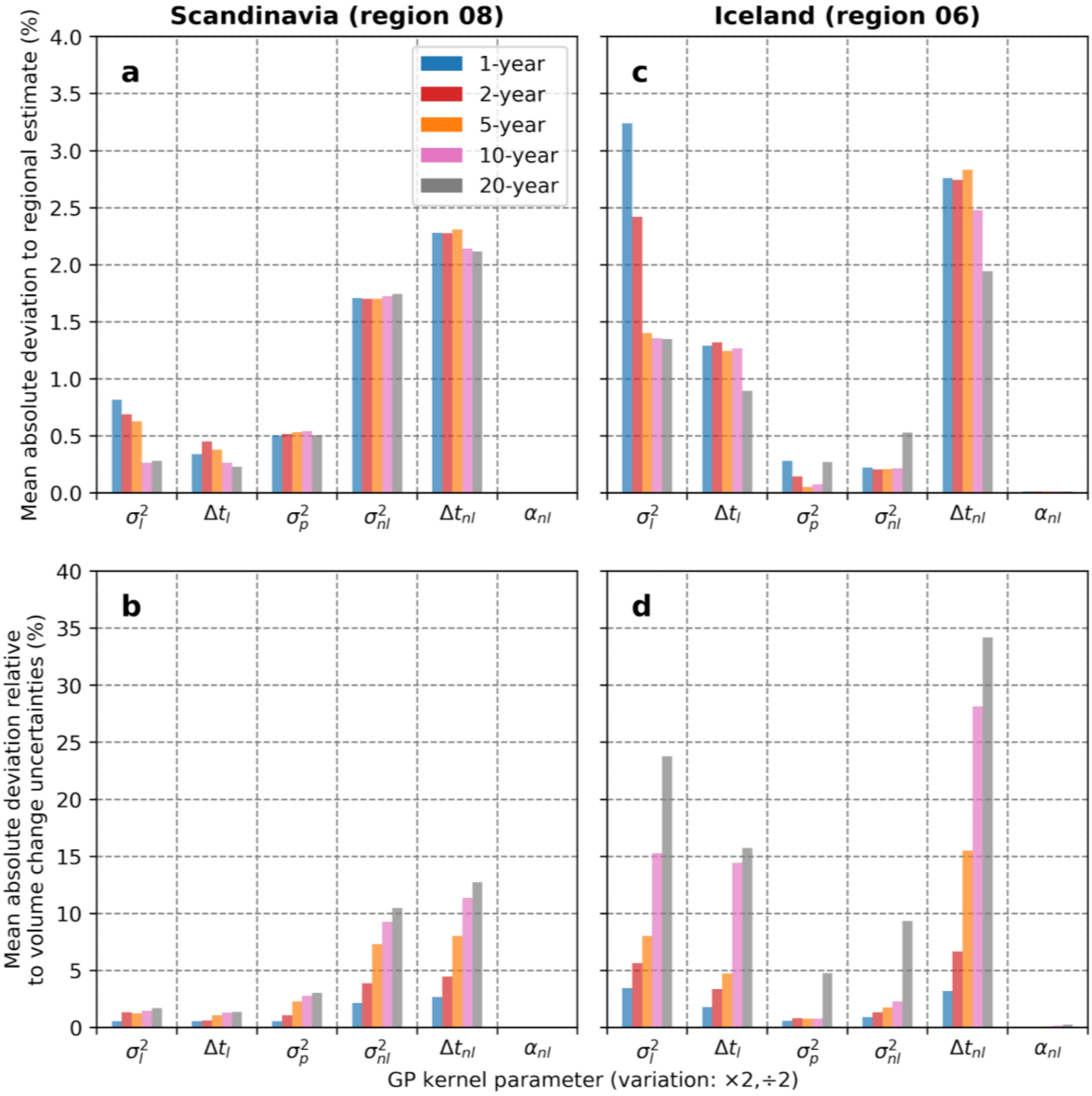

Fig. S10: Sensitivity to Gaussian Process kernel parameters.

Sensitivity of the regional estimates of glacier volume changes to the Gaussian Process kernels parameters for Scandinavia **(a,b)** and Iceland **(c,d)**. The kernel parameters are varied by multiplying and dividing the value used in this study by 2, and refer to Equation 2 (see Methods, or Supplementary Methods section 2.5). Iceland and Scandinavia were selected as they are potentially the most sensitive to Gaussian Process kernel parameters. This is due to both their small size (spatially correlated signal) and the fact that they show strong nonlinear changes during the past two decades (Extended Data Table 1). Additionally, they include a wide spectrum of temporal coverage, as Iceland is the region with the largest repeat coverage (~66 observations in 20 years per pixel) while Scandinavia is the region with the lowest repeat coverage (~27 observations in 20 years per pixel), excluding Antarctic and Subantarctic. Panels **(a)** and **(c)** show the mean absolute deviation relative to the regional estimate and panels **(b)** and **(d)** the mean absolute deviation relative to the estimated volume change uncertainty. The mean absolute deviation is computed from all possible successive time periods of a certain length in 2000-2019 (e.g., 5-year periods indicate 2000-2004, 2005-2009, 2010-2014 and 2015-2019) and varied parameters (x2,÷2). Overall, varying all Gaussian Process kernel parameters within this order of magnitude impacts the estimates less than 3%, which is well within estimated volume change uncertainties (at most 30% of uncertainty range) and estimated mass change uncertainties (at most 5% of uncertainty range). The maximum absolute deviation is within the same range and does not exceed 1.5 times the mean absolute deviation.

The kernel used is PL+ESS+RQ+RBF+(PL*RQ). This is a very particular design choice, largely motivated to match a variogram.

We decided to rely on a variogram for the following reasons:

- We required a covariance estimation method robust to outliers, since the latter are extremely frequent in our elevation data (see Extended Data Fig. 3, panels c and d). Our understanding is that maximum marginal likelihood methods can be significantly negatively affected by outliers.
- We aimed at identifying the form of the kernels that composed the covariance of the data, and identify possible nonstationarities in the covariance before applying GP regression. For this, our methodology is based on traditional spatiotemporal statistics (e.g., kriging), where the covariance model is user-defined and fitted to match the empirical covariance, and possible nonstationarities are to be identified by the user.
- The procedure is less computationally expensive than GP optimization of the parameters, which has been known to have difficulty scaling with big data (e.g., Liu et al. (2020)). Those aspects are further discussed below in the answers to the referee remark on “Scalability”.

When estimating our variograms, we did not find significant nonstationarities of the covariance with the elevation change trend, the terrain slope, or the accumulation and ablation areas of glaciers (normalized elevation). We found good consistency between the variograms of different glacierized regions, and different locations, with similar form and hyperparameters within close ranges, justifying the use of the same covariance function at the global scale (Figure R1, below).
These points have been further clarified in the related section of main Methods and SI.

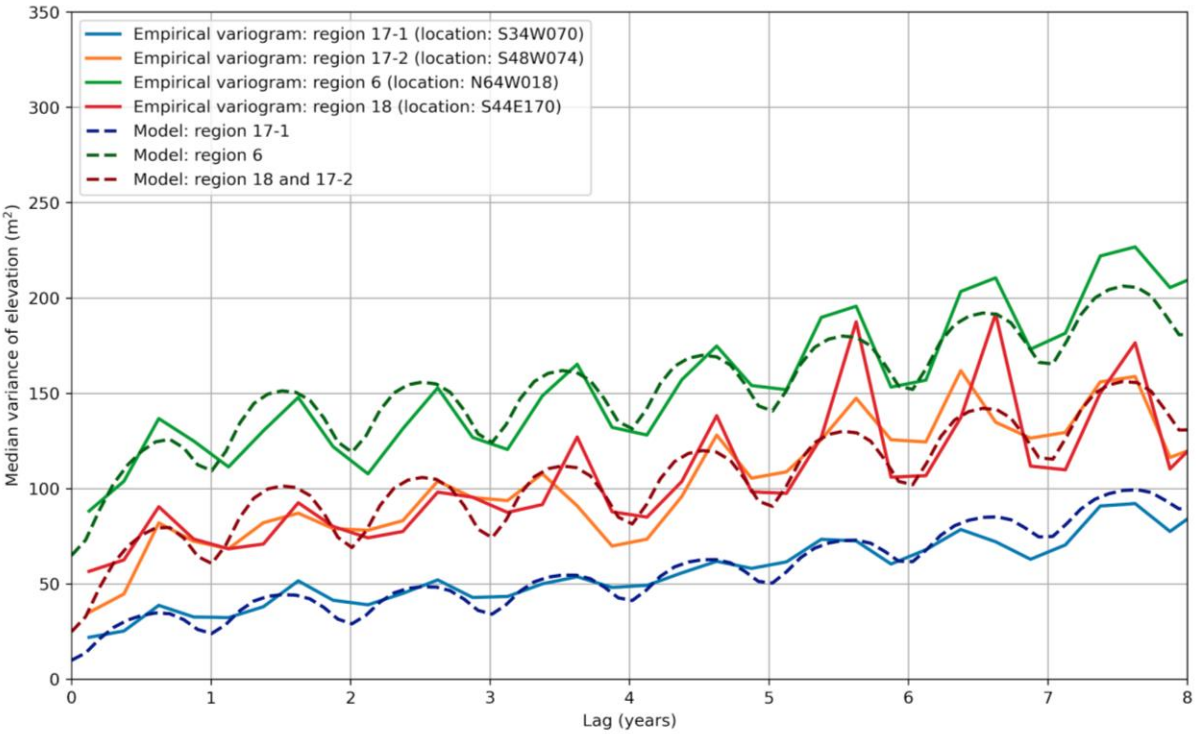

Fig. R1. Median variogram estimation in different regions.
The different levels of white noise (intercept) constitute the only major difference between variograms, and are due to varying elevation measurement errors at the locations sampled (independent from the covariance form). The underlying linear, local and seasonal trends are consistent between regions, with little variability. The amplitude of the seasonal signal is the parameter found to vary the most (2 - 7 m). For our study, testing showed little influence of constraining this parameter by area/region (see sensitivity of σ_p^2 on Fig. S10 above). This is because we generally have sufficient repeat data to estimate the periodicity from a mean variance value (~5 m). Additionally, DEMs are inevitably subject to systematic snow-cover biases (Fig. S4-S5) due to the current lack of a globally timestamped DEM. Consequently, estimating the seasonal cycle is not an objective of this study, and the periodic signal serves at mitigating seasonal effects for capturing short and long-term changes (see Methods, and SI Section 2.5).

What other choices were tried, and how were they rejected? The linear kernel will enforce a continued long-range trend, which we know cannot persist. Why not another RBF component with a longer length-scale?

The ESS and short-range RBF kernels were determined without ambiguity from the form of the variogram (Extended Data Fig. 3b, Fig. R1). For the long-term linear trend and local long-range trend, we considered using either (PL + RQ) or (PL + RQ*PL).

We studied each kernel's influence on the estimates and related those to a physical interpretation:

- Using (PL + RQ) implies that the elevation change can locally vary within a certain range from a "mean" long-term (20 years) linear trend. When no observation is available at the temporal boundaries (in this case near 2000 or 2020), the extrapolated trend falls back towards the "mean" linear trend (Figure R2, below). The usage of (PL + RQ) therefore implies that a short-time acceleration would have to be compensated by a deceleration towards the mean trend, and vice versa at sub-decadal scales. We know from long-term field observations that such a "rebound" response does not happen, and indeed it would also be difficult to interpret that in terms of glacier dynamics at these time scales.

- Using (PL + RQ*PL) implies that the "mean" linear trend can vary locally. When no observation is available at the boundaries, the extrapolated trend falls back towards the local "mean" linear trend (Figure R2, below). Physically, this is more consistent with existing observations and with known decadal and sub-decadal climatic oscillations that influence glacier change. As highlighted by the referee, the long-term, continued linear trend cannot persist, which is why we instead capture the local linear trend from the temporally closest observations.

We added a related paragraph justifying the choice mentioned above and the related physical interpretation in the SI, Section 2.5.

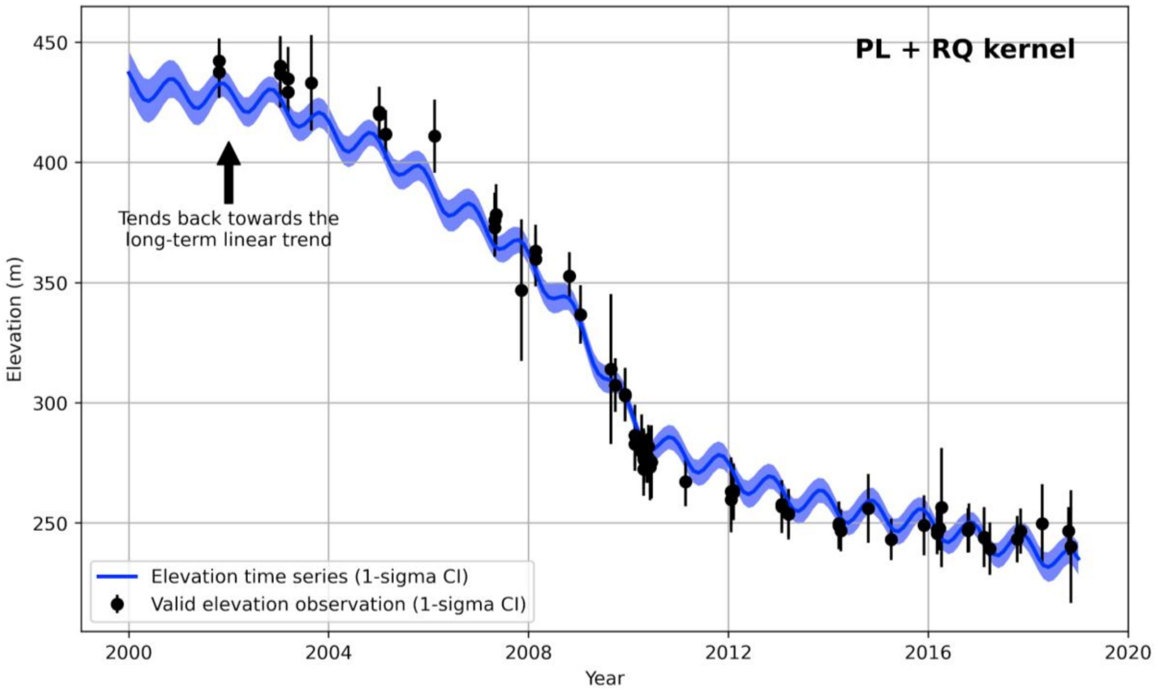

Fig. R2. Example of elevation time series with (PL + RQ) kernel.
Pixel on the tongue of Upsala glacier, Southern Patagonian Icefield. The effect at the time boundary is amplified for illustration purposes by reducing the variance of the RQ kernel.

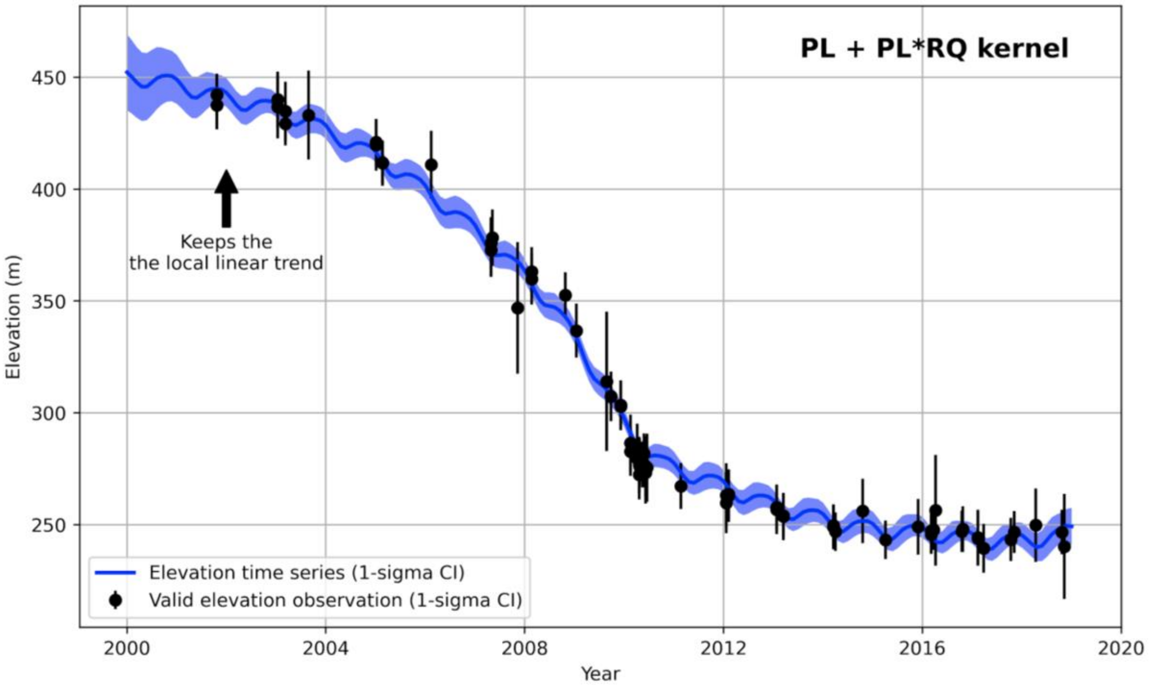

Fig. R3. Example of elevation time series with (PL + PL*RQ) kernel
Same example as Fig. R1. The estimated trend is consistent with the observations and the physical knowledge of short- to long-term glacier elevation changes.

We performed pixel-based and region-based testing. The first option (RQ) was rapidly discarded at the pixel-scale from Leave-One-Out Cross Validation. While the difference between (RQ) and (RQ*PL) was noticeable for pixels with limited temporal coverage (<10 observations in 20 years per pixel), once integrated over the millions of pixels of an entire region, the differences between 5-year estimates derived from (RQ) and (RQ*PL) were minor. This is because the elevation change trends are primarily constrained by the elevation observations, which have a repeat coverage that is dense enough to mitigate most of the undesired effects of the (RQ).

We thus chose the (RQ*PL), which better resolves the temporal boundaries of our study and matches a physical interpretation. Note, however, that the results which we present require little temporal extrapolation given the dense coverage of ASTER. As shown by the blue ribbons of Fig. 1, most regions have a yearly observational coverage of at least 60-70% of the surface area until 2019. Only the year 2000 stands out with a systematically reduced coverage (fewer acquisitions were made by ASTER durings its early stages), meaning that the trends we estimate for 2000 is most often extrapolated from observations starting in 2001. For the rest, the estimates are constrained by the data itself, with very little extrapolation.

Or why not use other less-smooth kernels like the Matern kernel?

We considered using a Matern kernel (a RBF kernel generalized with smoothing factor). Again, due to our dense temporal sampling, we assumed the sensitivity of an additional smoothing parameter would likely have very limited influence, which is indeed the case for the α_{nl} parameter in the sensitivity analysis of Fig. 10. The fact that, for a given pixel, our data does not have sub-meter precision (2-3 m on low slopes) would make it hard to identify this level of detail.

What hyperparameters did you learn (such as the length-scales for the various components) and are they interpretable in this application?

We have interpreted in more details the hyperparameters, and added a related statement in the Supplementary Information Section 2.5: "*We found* $\phi_p = 1$ *yr,* $\sigma_p \approx 5$ *m, implying a seasonal periodicity component of 5 m on average. We found that the local signal was best decomposed into a sum of three RBF kernels with* $\Delta t_{l1} \approx 0.75$ *yr,* $\sigma_{l1} \approx 5$ *m,* $\Delta t_{l2} \approx 1.5$ *yr,* $\sigma_{l2} \approx 4$ *m and* $\Delta t_{l3} \approx 3$ *yr,* $\sigma_{l3} \approx 2$ *m, which suggests that, once the underlying linear trend and periodicity is removed, inter-annual glacier elevations are on average within 5 m of each other within a year, within 9 m within 1.5 year and within 11 m within 3 years. Finally, based on pixel-scale testing (for filtering purposes) and the temporal range of the underlying linear trend observed in our empirical variograms, we constrained the local linear values to* $\sigma_{nl} \approx 10$ *m,* $\alpha_{nl} \approx 10$ *and* $\frac{1}{2\alpha_{nl}} \cdot \Delta t_{nl} \approx 5$ *yr. Those values mean that, on average, local linearity lasts around 5 years and within 10 m of the underlying linear trend.*"

In general, estimation of covariance functions by variograms tends to be much less precise than using the marginal likelihood of the Gaussian process. Why not try this approach for learning the hyperparameters of your kernel composition, and compare? Why not try automatic kernel learning approaches, such as spectral mixture kernels (with appropriate initialization techniques), in conjunction with marginal likelihood estimation? In the supplement, the reason given for preferring the variogram approach is "Our objective is to model variograms with characteristics representative of many pixels at once, and to apply these variograms directly in the regression. The rationale behind this approach is to mitigate the sparse sampling of elevations in time at the pixel scale by utilizing the repeat spatial coverage of our observations." But this objective can also be achieved (and likely improved upon) by following a marginal likelihood approach.

We agree with referee #4 that marginal likelihood methods and automated kernel learning approaches hold great promise to refine the approach for studying precise elevation change. Better conditioning the kernels and the hyperparameters, for example, could be a compelling way of study local changes (e.g., a specific part of a glacier) with higher-precision data that would have less dense temporal coverage than ours, and would rely significantly on the Gaussian Process prediction for the interpolation.

Our results, however, show a low sensitivity to kernel hyperparameters due to the dense repeat coverage. For the same reason, and as previously discussed, early testing showed limited sensitivity to the use of different kernels (e.g., (PL + RQ) or (PL + RQ*PL)). Therefore, we did not implement any automated kernel learning approach. Those choices were also motivated by scalability and the strong influence of outliers, described in following answers to the referee inquiries on those aspects.

Since uncertainty estimation is crucial in this application, why not also place distributions over the kernel hyperparameters, and then perform Bayesian marginalization over these distributions? Surely by not doing this, you will be underestimating your uncertainty?

We appreciate referee #4's point if our objective was to more precisely quantify the uncertainties in elevation time series. In the frame of our study, however, we focus on quantifying the uncertainties in volume change, for which pixel-wise elevation uncertainties play a limited role. We note that, despite this, we have strived to provide more rigorous elevation change uncertainties than preceding studies. Those two aspects are detailed below.

On the uncertainty estimation
We have added Fig. S9 to the Supplementary Information (also available two pages down), that illustrates the propagation of the different sources of uncertainties to the volume and mass change uncertainty.
While we agree with referee #4 that the uncertainties are crucial in this application, the uncertainties that are of concern are those of glacier volume change, later converted to mass change. The uncertainties in elevation, derived from the Gaussian Process method, have virtually no impact on the uncertainties of the estimated volume changes (Fig S9a). This is because the uncertainties in elevation changes are dominated by short- to long-range spatial correlations (2 km - 200 km) and not by the pixel-wise values. We quantified these long-range spatial correlations using the difference to millions of precise ICESat measurements (Extended Data Fig. 5a,b). By doing this, we account for (i) the limits of statistical interpolation for glacier elevation changes (in other words, what physical glacier signal might be missed when performing a temporal interpolation) and (ii) the limits of the parametrization of our Gaussian Process regression (or, stated, differently, the possible room for improvement).
Importantly, using high-precision volume change data derived from high-resolution DEMs, we validated the reliability of our volume change uncertainties estimated from spatial correlations, and this at various scales (Extended Data Fig. 5c-e). We find no bias (0.03 ±0.03 m yr^{-1}) and that our uncertainties match empirical uncertainties (and are even conservative for small glaciers). The refined uncertainties for volume change represent a step forward compared to existing work (discussed in the section "*Uncertainty analysis of volume changes*" of Methods).

On Gaussian Process-based elevation uncertainties
While the elevation uncertainties of each pixel have little impact on our results, we nonetheless strived to provide a more rigorous analysis than those found in existing studies. First, by conditioning pixel-wise elevation measurement errors before interpolation (Extended Data Fig. 3a, Section 2.3 of Supplementary). Then, by comparing the interpolated values to all available high-precision ICESat and IceBridge measurements. Importantly, we have shown that the pixel uncertainties derived by our Gaussian Process regression are on average about twice as large as they would need to be (Methods section "Validation of elevation time series", Extended Data Fig. 4), which differs from the reviewer's hypothesis that our uncertainties might be too small. Whilst there is no doubt that, at local scales, there is room for method improvement, we are confident that our glacier- and regional-scale uncertainty estimates are not only honest but even conservative.

Implementing the above suggestions would significantly increase my confidence in the Gaussian process estimates. At the minimum, it would be good to see a detailed study of how this kernel was decided upon, including a description of what other configurations were tried and why they were rejected.

Further above, we provided a summary of how this kernel was decided upon. Additional information has been added to the related Section 2.5 of the SI.
In a nutshell, we have focused on ensuring that the predictions at the extrapolation boundaries of the Gaussian Process regression were consistent with the physical processes that govern glacier elevation changes. Again, we stress that our results are only marginally using a temporal extrapolation at all, thus greatly reducing the relevance of the corresponding methodological choices.
We would also like to reiterate that our study benefits from the large-scale validation with ICESat, that ensures the absence of both temporal or spatial biases, as well as conservative elevation uncertainties. Most of all, we used the difference between the interpolated values and ICESat measurements to assess spatial correlations, which we have shown to represent 99% of volume change uncertainty sources (Fig S9a, glaciers > 10 km²). By being derived directly from the difference with our GP estimates, our uncertainty analysis accounts for possible improvement in parametrization. It currently yields small elevation change uncertainties compared to other sources of uncertainties (Fig. S9b), and compared to preceding studies (Zemp et al. (2019), Extended Data Fig. 4: geodetic uncertainties). Finally, our volume change uncertainties are validated against a large number of high-resolution volume changes. This shows that, with the current method, no possibly omitted structure of the data (e.g., nonstationarities) is impacting the global-, regional- or glacier-scale estimates.

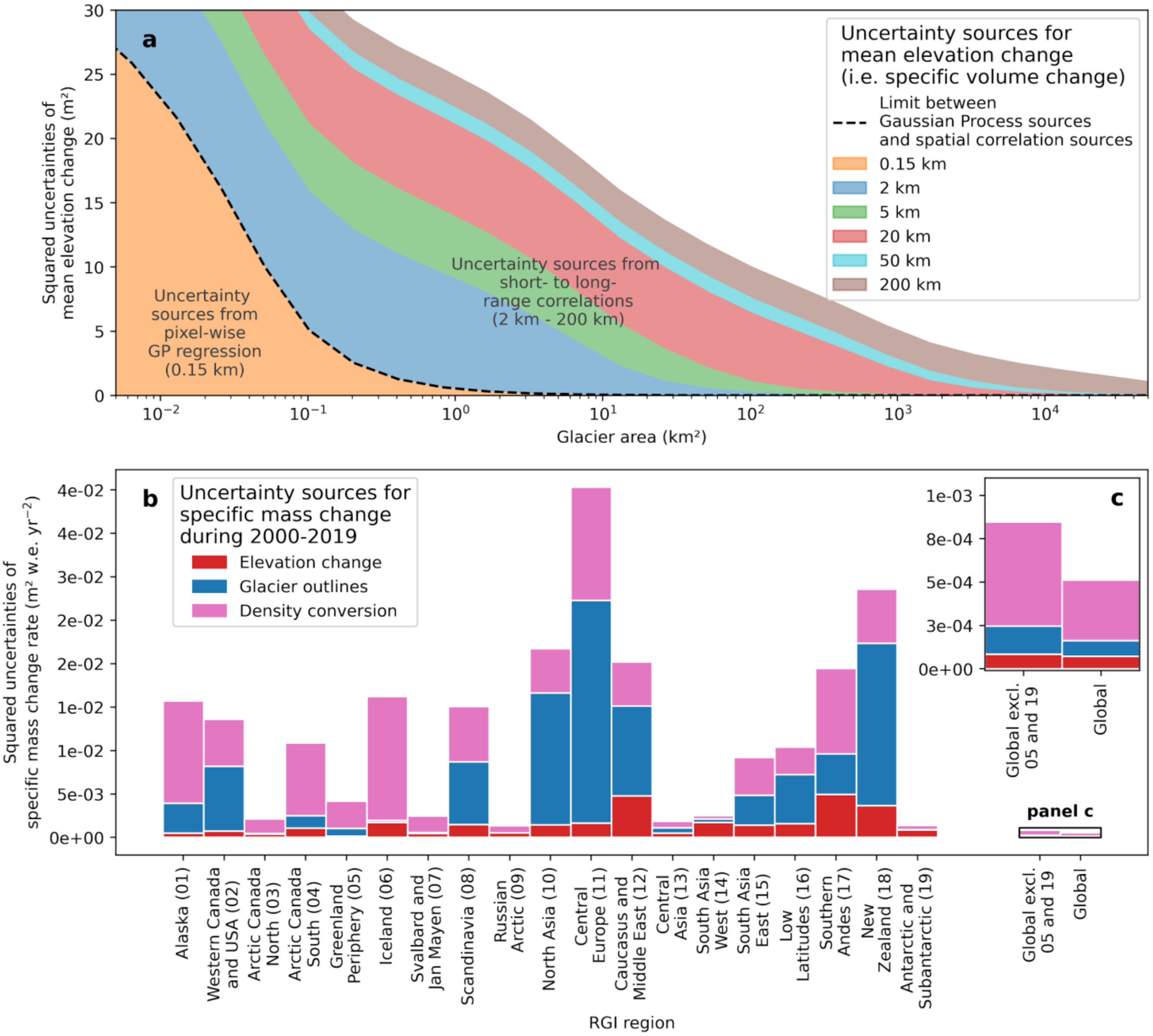

Fig. S9: Uncertainty sources.
Uncertainty sources for estimation of volume changes **(a)** and mass changes **(b,c)**. **a**, Propagation of elevation change uncertainties to volume change uncertainties with varying glacier area. As this computation is specific to the time lag of each pixel to the closest observation, for each glacier, at each time step, panel (**a)** refers to an example. The spatial correlations are computed from a time lag to the closest observation (representing the average of our study) of 0-1 year for 50% of observations, 1-2-years for 20% of observations, 2-3 years for 20% of observations and 3-4 years for 10% of observations (see Extended Data Fig. 5 a,b). We assume a mean pixel-wise error of 10 m and simplify by considering only the first step of integration over a continuous glacierized area (Equation 5). This assumption leads to slightly larger contributions from short-range correlations than with further propagation to the second propagation step between discontinuous glaciers (Equation 6). Uncertainties are largely dominated by short- to long-range spatial correlations. **b**, Propagation of uncertainty sources to specific mass changes for each RGI region, and all glaciers with and without the Greenland Periphery and the Antarctic and Subantarctic which are zoomed on panel **(c)**. Uncertainties are largely dominated by the volume-to-mass conversion uncertainties globally, and also by uncertainties in glacier outlines for regions with a relevant share of small glaciers.

Scalability:
In many cases downsampling is applied for computational reasons with the Gaussian processes. But this is not necessary with modern advances in GP scalability, and loses information that could be very helpful in using the marginal likelihood for kernel estimation (as described above). It is now common to use Krylov subspace methods, such as linear conjugate gradients, and stochastic Lanczos quadrature, to
scale even exact Gaussian process methods to problems with millions of points, through GPU parallelization. These methods are implemented in popular packages such as GPyTorch. Given the spatiotemporal structure of your problem, you can also likely use Kronecker methods to significantly scale your approach without making approximations, since your inputs are probably on close to a multidimensional (fully connected) lattice (can be expressed as a Cartesian product of vectors). The idea would be to create a kernel which is a product of 1D kernels (for example with the form you have chosen already) operating separately on each input dimension. Given the input structure of your problem, the covariance matrix would then be a Kronecker product of much smaller matrices. More information can be found (e.g., http://mlg.eng.cam.ac.uk/pub/pdf/Saa11.pdf), and there are many extensions to these techniques for missing points, or using virtual grids that cover the range of your raw inputs.

In any case, the downsampling is likely not necessary, and will be hurting the performance of your approach, and limiting your ability to do automatic kernel learning, which would be a very compelling alternative to the hand crafted kernel you are using.

We thank the referee for this valuable information. As our work was started in 2018, and performed mostly during 2019, we did not know about the concomitant advances in GP scalability (Gardner et al. (2018), Pleiss et al. (2018)).
We are unfamiliar with how these recent approaches can scale past millions of points and note that our 30 m DEMs comprise some 200 billion points. Most importantly, however, we do not feel that re-designing our entire approach to be compatible with GPU fits with the current scope and stage of our study. While these advances are of interest to us, the following considerations might help in relativizing the necessity of such advanced procedures in the frame of this study:

On the downsampling:
If the referee is referring to the DEMs downsampled from 30 m to 100 m:
We do not expect this procedure to have any impact on the results. This is demonstrated by the strong spatial correlation at short spatial ranges of 150 m of ASTER data (Section 4.1 of the Supplementary). It implies that 30 m pixels are strongly correlated. In other words, a 1x1 100 m downsampled pixel will hold almost as much information as a 3x3 grid of 30 m pixels. Consequently, there is almost no information loss.

If the referee is referring to the random sampling of 10,000 samples by tile and by category of external variable (to test possible nonstationarities) during the variogram estimation:
We have performed this operation in many regions, with similar results (Figure R1). Hence, this downsampling (a common approach for variogram estimation with gridded data) did not seem to affect the performance of the covariance estimation significantly. We therefore did not feel the need to sample more data at once.

On general scalability:
It is not only a concern of CPU but also disk storage. The choices of downsampling to 100 m, and estimating covariance based on 10,000 samples, were mainly motivated by limitations in manipulating the whole dataset at once. The chosen methods made it easier to partition the data during covariance estimation, or Gaussian Process inference, for technical reasons:
The entire dataset, uncompressed, amounts to about 100 TB of RAM. The generation of a monthly time series of elevation at 30 m resolution on all of Earth's glaciers and within a 10 km buffer, with all related metadata, would amount to ~20 TB of compressed files on disk, solely for the elevation time series estimates. With the means that we currently have at our disposal, this is not feasible.

Because of these reasons, we estimate that adapting to the novel methods mentioned by the referee would take a long time not only in implementation, but also computing (at least half a year) - and this without any significant impact on our estimates and uncertainties for glacier mass change estimation, as previously demonstrated. Consequently, while we are grateful towards referee #4 for this valuable information on recent statistical advances, and keen to explore these new means of computations and methods in the future, we think such an additional effort would be disproportionate to the present study. The large computational effort we required was the main motivation to perform several validation procedures of the current implementation step by step, region by region with independent high-precision data, before running our final computations globally.

Outliers:
In the supplementary material there is extended discussion of filtering and diminishing the effects of outliers. Outside of high fidelity to the secondary data sources such as ICESat and ICEBridge, how can we be sure that important structure in the data is not being removed in enforcing insensitivity to outliers? It seems possible that in procedures such as fitting variograms based on median residuals (rather than mean residuals) would produce overly smooth estimates of elevation, which is not uncommon in other spatial estimates of climatological quantities.

These are credible points raised by the referee, related to the intricacies of glacier elevation data, which we detail below.

On the filtering of outliers
Glacier elevation change has been long-studied, and we know that the most "extreme" elevation changes that can be observed are those caused by glacier surges. For the largest ones, it can result in hundreds of meters of elevation gain/loss in less than a year (e.g. Nathorstbreen in Svalbard: Extended Data Fig. 6, Fig. S3c,d). As we cover all of Earth's glaciers, we identified the Nathorsbreen surge as the largest of the past two decades, and performed our testing at that site to ensure its data was not filtered as outliers, and that no other surges in the world would be filtered either. Because surges are generally rare, we initially left that specific information in the structure of our code only (following remarks from the editor in the 1st round). We have added a short statement in the Supplementary Information, Section 2.5: "*Our primary objective was to ensure a low sensitivity to outliers, which were not effectively filtered out when using shorter-time scale parameters. In order to*

avoid removing glacier surges, we included a conditional loop in our procedure, which was calibrated on Nathorsbreen Glacier, Svalbard (the largest surge observed during our period of study, Fig S3).”.
Additionally, we do not expect our filtering procedure to remove actual elevation data, as we leave a large interval of 4 times the Gaussian Process credible intervals in our filtering procedure (Supplementary Section 2.6). Based on inspection of our results, we have identified quite the opposite, in fact, with some small areas still slightly affected by unfiltered cloud outliers (e.g. a few pixels around the highest peaks of certain subregions, for instance in Kunlun, Mount Elbrus or Aletsch glacier).
An improvement of this aspect would require a global classification of surging glaciers to refine parameters, which was not manageable in the current study. This point is now detailed in Supplementary Section 2.7: “*Improving these aspects would require a classification of glacierized terrain for extreme events prior to constraining the temporal covariance and performing temporal interpolation, which was not feasible at a global scale.*”

On the effect of median variograms
We agree with referee #4 that a median variogram would produce overly smooth estimates of elevation. To some extent, this is inevitable, due to several factors:

- The limited precision of individual elevation observations from ASTER (2-3 m on low slopes and up to 20-30 m on high slopes) makes it hard to confidently deconvolve elevation changes during short periods.
- The current difficulty of filtering the large amount of elevation outliers in elevation data (due, for example, to photogrammetric artefacts or clouds) can rapidly create false trends if the covariance function allows for rapid, short-term local changes.

A smoothing of elevation estimates is thus preferable over creating false trends. We feel that it does not constitute a limitation for our application due to the following reasons:

- Thanks to both the newly generated data and the mitigation of seasonality in the time series (which was the greatest limit of linear methods implemented by previous studies), we are now able to capture significant 5-year elevation trends. This is a significant step forward on the study of recent glacier surface elevation changes.
- We quantify the bias created by this overly smooth estimation for shorter periods within our spatial correlation analysis based on ICESat measurements. Therefore, this effect is reliably represented by our uncertainties, which can be large for short periods.
- Even if we were able to better capture elevation changes over shorter periods, we currently do not have the knowledge to reliably convert the related volume changes into mass change. This is because processes such as firn compaction, snow redistribution, or refreezing can importantly affect the evolution of glacier surface elevation - without implying any mass changes.

Those aspects are discussed in the Supplementary Discussion sections “Improved elevation change estimation” and “Time series comparison and temporal resolution”.

Key High-Level Concern:
My high level primary concern is the treatment of the glacial mass estimates as data. Historical data are used to make and verify Gaussian process based estimates of glacier mass using elevation, but in many sections the estimates are treated interchangeably with

recorded data. The results are at least not explic- itly stated as estimates throughout, and are presented as though the estimated glacial mass is just a noisy measurement.
One place in the paper this becomes a concern is in figure 3. In some cases there is substantial disagreement between the estimates in this work and the estimates made in other works which is mentioned in the discussion. However, discussing the estimates as fact in other cases does lead to concern when points in figure 3 indicates other sources may disagree with the findings.

We understand from the referee's comment two distinct points:

- The nomenclature of "data" and "estimate" have a certain philosophical difference that must be clearly differentiated.
- The comparison with previous studies is sometimes problematic because the uncertainty bounds do not overlap with our own.

We clarify those two aspects below.

On the terminology of "data" and "estimate":
We have modified the text in the Main and Supplementary to ensure that the elevation observations are always referred to as "data" while all variables inferred from them are referred to as "estimates".

On the concern between our uncertainties and those of other studies on Fig. 3:
During the comparison to other studies, an important factor to keep in mind is that the uncertainties computed by these other studies were generally based on empirical approximations and much simpler approaches, and that the uncertainties were not validated against any independent data -- primarily because such data are hard to come by. The range of these estimated uncertainties is therefore not systematically trustworthy, and at times hardly comparable to ours.
The most striking example might be the estimate for region 2 (Western Canada & USA) from the study of Gardner et al. (2013): -0.93 ± 0.23 m w.e. yr^{-1} (Fig. 3a, red). This estimate is based on a few in-situ measurements available for a dozen small glaciers (less than 1% of the surface area), and extrapolated to the 20,000 glaciers in the regions. In our study, the estimate for this region and for the same period is of -0.35 ± 0.09 m w.e. yr^{-1}, and is based on a sample comprising 28 observations per pixel over 98% of the surface area. Our uncertainties are validated with independent data, and are likely too large (i.e. conservative) due to the limited knowledge about the density for volume-to-mass conversion (this is mentioned in the main, and further discussed in the Supplementary Discussion "Uncertainty propagation and limitation of density-based mass change uncertainties": "*This effect likely provides uncertainties that are too large, especially for regions with strong mass losses (e.g., Fig. 3).*"). This thus indicates that the uncertainties by Gardner et al. (2013) were largely under-estimated.
Similar considerations can be made for many other regions and studies. Those discrepancies come from either less rigorous approaches (no validation), or from limited data available for estimating the corresponding uncertainties. As we cannot resolve or change the issue for these independent studies, we highlight again the improvements made by our own study. In the main text, we do that with the following statements:
"*We further utilize ICESat data to constrain the spatiotemporal correlations that are either structural to our interpolated elevation time series or that emerge due to latent, uncorrected*

ASTER instrument noise, and we propagate our elevation errors into volume change errors accordingly. We validate the reliability of our uncertainty estimates down to the scale of individual glaciers by comparison to independent, high-precision DEM differences for 588 glaciers around the globe (Extended Data Fig. 5)."

Minor:
It would be more accurate to refer to your uncertainty intervals are "95% credible sets" rather than "confidence intervals".

We have changed "confidence intervals" to "credible intervals" for the elevation uncertainties derived from the Gaussian Process regression. As the rest of the uncertainty analysis is mostly based on frequentist inference, and in order to avoid confusion with this term not widely used in glaciological literature, we have kept "confidence intervals".

We thank again referee #4 for his/her valuable comments, which have resulted in producing a clearer presentation of both the Gaussian Process covariance function selection, the analysis of current parametrization sensitivity and the global impact of uncertainty propagation within the study. Those are now available in our paper. This review also brings to our attention some recent improvements of great interest for future implementations of the Gaussian Process approach.

References in the answers to editor and referees

Aðalgeirsdóttir, G. *et al.* Glacier Changes in Iceland From ~1890 to 2019. *Front Earth Sci.* **8**, 520 (2020)

Pleiss, G., Gardner, J. R., Weinberger, K. Q. & Wilson, A. G. Constant-Time Predictive Distributions for Gaussian Processes. *arXiv [cs.LG]* (2018)

Sandberg Sørensen, L. *et al.* 25 years of elevation changes of the Greenland Ice Sheet from ERS, Envisat, and CryoSat-2 radar altimetry. *Earth Planet. Sci. Lett.* **495**, 234–241 (2018)

Gardner, J. R., Pleiss, G., Bindel, D., Weinberger, K. Q. & Wilson, A. G. GPyTorch: Blackbox Matrix-Matrix Gaussian Process Inference with GPU Acceleration. *arXiv [cs.LG]* (2018)

Liu, H., Ong, Y.-S., Shen, X. & Cai, J. When Gaussian Process Meets Big Data: A Review of Scalable GPs. *IEEE Trans Neural Netw Learn Syst* **31**, 4405–4423 (2020)

Bevis, M. *et al.* Accelerating changes in ice mass within Greenland, and the ice sheet's sensitivity to atmospheric forcing. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 1934–1939 (2019)

Belart, J. M. C. *et al.* Mass Balance of 14 Icelandic Glaciers, 1945–2017: Spatial Variations and Links With Climate. *Front Earth Sci.* **8**, 163 (2020)

Ciracì, E., Velicogna, I. & Swenson, S. Continuity of the Mass Loss of the World's Glaciers and Ice Caps From the GRACE and GRACE Follow-On Missions. *Geophys. Res. Lett.* **47**, 226 (2020)

Velicogna, I. *et al.* Continuity of Ice Sheet Mass Loss in Greenland and Antarctica From the GRACE and GRACE Follow-On Missions. *Geophys. Res. Lett.* **47**, L11501 (2020)

Gardner, A. *et al.* A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009. *Science* **340**, 852–857 (2013)

Huss, M. Density assumptions for converting geodetic glacier volume change to mass change. *The Cryosphere* **7**, 877–887 (2013)

Parkes, D. & Marzeion, B. Twentieth-century contribution to sea-level rise from uncharted glaciers. *Nature* **563**, 551–554 (2018)

Wouters, B., Gardner, A. S. & Moholdt, G. Global Glacier Mass Loss During the GRACE Satellite Mission (2002-2016). *Front. Earth Sci.* **7**, 535 (2019)

Reviewer Reports on the Second Revision:

Referee #3 (Remarks to the Author):

I have only a few minor comments for the revised article, described below.

lines 34-35: Can you specify what you mean by "glaciers presently lose more mass"? The rest of the sentence refers to annual mass-loss rates and their acceleration. Does the first part of the sentence refer to a specific period? Otherwise, it seems to be redundant.

line 56: is the impacts on ecosystems refers to the impact of mass changes on the ecosystems? Changes in ecosystems have a complicated relationship to cryospheric changes. Mass-change is probably not the best proxy for characterizing this relationship, e.g., the extent of glaciated terrain, permafrost, etc., might be more relevant.

lines 76-77: Referencing the use of "modern" photogrammetry techniques does not seem appropriate here. The reprocessing of stereo ASTER images is a remarkable achievement, but it can mostly be attributed to a better modeling of sensor geometry rather than modern photogrammetry techniques. The statistical methods are relatively standard and not specifically-developed. However, the authors used them skillfully and appropriately parameterized to solve the DEM time series's spatiotemporal interpolation.

lines 92-93: it might be appropriate to mention NASA here, NASA's ICESat and Operation IceBridge campaigns. Note that the name of Operation IceBridge is abbreviated OIB and I suggest using the abbreviation for the rest of the paper. For example, instead of IceBridge data, the authors can refer to OIB data.

Line 124, 127, 164, etc.: the authors should be clear to which study they refer to. Frequent use of "this" can result in ambiguous statements.

Line 171: please mention the source of the radar DEM. Was it interferometry, altimetry, something else?

Line 181: I'm afraid I have to disagree with the statement about Antarctica's scattered peripheral glaciers. This statement combines glaciers in the peninsula, EAIS, and WAIS. The peninsula has only outlet glaciers and ice shelves, WAIS has large outlet glaciers, and EAIS has a large ice sheet surrounded by scattered peripheral glaciers. These different scenarios should be clearly distinguished.

Line 193: remove extra "one of" from the sentence

Line 213: should it be coastal mountains rather than Coast Mountains?

Line 221: I don't see any evidence showing that the warming of the atmosphere controls the global acceleration glacier mass loss. All I see is a correlation -- in my opinion, more is need to prove a causative relationship.

Lines 274-275: the disk scale needs to be explained better

Figure 4 caption: I can't follow the explanation of the tiles' size -- please review and revise.

Lines 498-499: where is the explanation of the annotations in this equation?

Line 508: Please include the full name AND the abbreviation of the ICESat ad OIB (IceBridge) product names.

Referee #4 (Remarks to the Author):

I believe the authors have done a good job of checking their approach, especially with the new updates to the supplementary material.

I do think the Gaussian process methods could be improved. It could be possible, for instance, to exploit Kronecker methods for significantly improved scalability, even without requiring GPU parallelization. Although more common in machine learning than statistics, I also do not believe exploiting GPU hardware for acceleration would be a major undertaking. Using the marginal likelihood would also be possible, and would not be sensitive to outliers with a reasonable observation model (such as heavy-tailed noise), or pre-processing approach.

However, I also agree these changes in this instance are not likely to have a major effect on the glacier mass volume change predictions, especially since the Gaussian process is being used for interpolation rather than extrapolation. I also think the authors did a good job of validating the variogram approach they used, and that this approach does have some benefits of interpretability. I particularly appreciate the new content that was added to the supplementary material.

Overall, this is a thoughtful, exciting, and well-validated study.

Author Rebuttals to Second Revision (please note that the authors have quoted the reviewers in black text and responded in blue text):

Referees' comments:

Referee #3 (Remarks to the Author):

I have only a few minor comments for the revised article, described below.

lines 34-35: Can you specify what you mean by "glaciers presently lose more mass"? The rest of the sentence refers to annual mass-loss rates and their acceleration. Does the first part of the sentence refer to a specific period? Otherwise, it seems to be redundant.

This statement is valid for both the full period of study 2000-2019 or any 5-year subperiods considered including the most recent one (Table 1). We thus leave this sentence to avoid repeating the time period of our study in the abstract, which has stringent limits of space.

line 56: is the impacts on ecosystems refers to the impact of mass changes on the ecosystems? Changes in ecosystems have a complicated relationship to cryospheric changes. Mass-change is probably not the best proxy for characterizing this relationship, e.g., the extent of glaciated terrain, permafrost, etc., might be more relevant.

We agree with referee #3 that the extent of glaciated terrain or permafrost can be useful for assessing change in ecosystems close to glaciers. Quantifying mass change, however, allows better assessment of changes in river runoff affecting all downstream ecosystems, from neighbouring ones down to the oceans. All are included in this introductory sentence focused on glaciers (and not permafrost), and based on Cauvy-Fraunié et al. (2019).

lines 76-77: Referencing the use of "modern" photogrammetry techniques does not seem appropriate here. The reprocessing of stereo ASTER images is a remarkable achievement, but it can mostly be attributed to a better modeling of sensor geometry rather than modern photogrammetry techniques. The statistical methods are relatively standard and not specifically-developed. However, the authors used them skillfully and appropriately parameterized to solve the DEM time series's spatiotemporal interpolation.

We respectfully disagree with the reviewer on this point. The improved quality of the ASTER DEMs compared to existing NASA products (e.g., ASTER_14DMO) arises from both improved correction methods and recent advances in photogrammetric techniques (e.g., improved correlation schemes: Rupnik et al. (2017), Beyer et al. (2018)). We feel that the term “modern photogrammetry techniques” best describes this aspect of our study.

Additionally, we note to the referee that the statistical methods mentioned in this sentence are related to DEM generation and correction, and not to the elevation time series: “*We use modern photogrammetry techniques and specifically-developed statistical methods to generate and bias-correct nearly half a million Digital Elevation Models (DEMs)*”. The empirical statistical corrections of cross- and along-track biases were specifically-developed for ASTER DEMs during the PhD work of Luc Girod (Girod et al. (2017)) and further refined in our study.

lines 92-93: it might be appropriate to mention NASA here, NASA's ICESat and Operation IceBridge campaigns. Note that the name of Operation IceBridge is abbreviated OIB and I suggest using the abbreviation for the rest of the paper. For example, instead of IceBridge data, the authors can refer to OIB data.

We added NASA at this point of the text: "*...with 25 million high-precision measurements from NASA's Ice, Cloud, and land Elevation Satellite (ICESat) and Operation IceBridge campaigns, ...*".
As suggested by Nature's guidelines, we believe it is prudent to use as few acronyms as possible. This is especially true given that Operation IceBridge is often referred to simply as "*IceBridge*", for instance in the official product releases used in this study (Alexandrov et al. (2018) and Larsen et al. (2010)). We thus keep the abbreviation "*IceBridge*", more transparent than "*OIB*" suggested by the referee.

Line 124, 127, 164, etc.: the authors should be clear to which study they refer to. Frequent use of "this" can result in ambiguous statements.

For the statements lines 124-127, all studies are referenced at their mention in the text. We found a single use of "*this*" which we changed to "*the latter*" for clarity: "*Observational studies were yet unable to discern significant accelerated glacier mass loss*[19,21]*, with the exception of a recent gravimetric study*[20] *that estimated an acceleration of 50 ± 40 Gt yr-1 per decade excluding peripheral glaciers. Despite its large uncertainties, the latter estimate*[20] *is in agreement with our results.*".

For line 164, all references to studies are repeated and we again modified "*this*" into "*the latter*": "*We note, however, large discrepancies between the latter gravimetric study*[19] *and a more recent one*[20] *in both Iceland and the Russian Arctic.*".

Line 171: please mention the source of the radar DEM. Was it interferometry, altimetry, something else?

We have clarified the statement by adding "*interferometric*" and removing "*DEM-based*": "*...twice as negative as that of a recent interferometric radar study*[29]*, ...*".

Line 181: I'm afraid I have to disagree with the statement about Antarctica's scattered peripheral glaciers. This statement combines glaciers in the peninsula, EAIS, and WAIS. The peninsula has only outlet glaciers and ice shelves, WAIS has large outlet glaciers, and EAIS has a large ice sheet surrounded by scattered peripheral glaciers. These different scenarios should be clearly distinguished.

Our aim here is not to enter into the details of the processes of mass loss for different parts of the AIS. Rather we want to draw the attention of the reader toward the similarities between the changes of scattered peripheral glaciers and the changes of the Antarctic ice sheet. We agree with referee #3 that a more detailed interpretation would bring in more information. Unfortunately, due to stringent limits of space, such additional information would not fit in the present main text and would have limited relevance as our study focuses on glaciers and not on the Antarctic ice sheet.
Our general statements refer to mass loss observations and its comparison to patterns observed by Velicogna et al. (2020). Those are simply observational and therefore valid independently of the underlying processes driving these patterns.

Line 193: remove extra "one of" from the sentence

Here the subject of the sentence is “*the slowdown of mass loss*”, thus requiring the addition of “*one of*” before mentioning that of the Greenland periphery region: “*Taken together, the slowdown in mass loss from these two regions, in addition to the one of peripheral glaciers of southeast Greenland Periphery*[32]*, define a regional pattern that we refer to as the North Atlantic anomaly.*”.

Line 213: should it be coastal mountains rather than Coast Mountains?

It should be the Coast Mountains: “The Coast Mountains are a major mountain range in the Pacific Coast Ranges of western North America, extending from southwestern Yukon through the Alaska Panhandle and virtually all of the Coast of British Columbia south to the Fraser River” (ref https://en.wikipedia.org/wiki/Coast_Mountains or see Bostock (1948)).

Line 221: I don't see any evidence showing that the warming of the atmosphere controls the global acceleration glacier mass loss. All I see is a correlation -- in my opinion, more is need to prove a causative relationship.

We agree and modified “*controlled*” into “*mirrors*”: “*While decadal changes in precipitation explain some of the observed regional anomalies, the global acceleration of glacier mass loss mirrors the global warming of the atmosphere (Fig. 4).*”.

Lines 274-275: the disk scale needs to be explained better

We have clarified the caption of Fig. 2: “*Disks scale with the glacierized area of each tile and are colored according to the mean elevation change rate (colored in grey if less than 50% surface is covered by observations or if the 95% confidence interval larger than 1 m yr^{-1}; only applies to 0.4% of the glacierized area).*”.

Figure 4 caption: I can't follow the explanation of the tiles' size -- please review and revise.

We have clarified the related sentence in the caption of Fig. 4: “Tiles are 1°x1° with size inversely scaled to uncertainties in the mean elevation change difference (full size: 95% confidence intervals of less than 0.2 m yr^{-1}; minimum size of 10%: 95% confidence intervals larger than 1 m yr^{-1}).”.

Lines 498-499: where is the explanation of the annotations in this equation?

We have added a parenthesis at the mention of the kernel “*parameters*” in the sentence preceding Equation (2) to ensure all annotations are mentioned in the main Methods text: “*We thus conditioned the parameters of the ESS, RBF and RQ kernels (*ϕ_p*,*σ_p*,*Δt_l*,*σ_l*,*Δt_{nl}*,*σ_{nl} *and* α_{nl}*) at the global scale based on our empirical variograms...*”.

Line 508: Please include the full name AND the abbreviation of the ICESat ad OIB (IceBridge) product names.

The full product names are directly available in the reference section, where the datasets are cited in accord with their reference at this point of the text: ”*We retrieved all ICESat (GLAH14*[63]*) and IceBridge (IODEM3*[64] *and ILAKS1B*[65]*) laser and optical elevations intersecting glaciers worldwide*”.
Adding the full names directly in the Methods section would introduce several long sensor and region acronyms that would need to be detailed, significantly lengthening the main Methods without adding any information of importance.

We thank referee #3 for those additional comments for clarification.

Referee #4 (Remarks to the Author):

I believe the authors have done a good job of checking their approach, especially with the new updates to the supplementary material.

I do think the Gaussian process methods could be improved. It could be possible, for instance, to exploit Kronecker methods for significantly improved scalability, even without requiring GPU parallelization. Although more common in machine learning than statistics, I also do not believe exploiting GPU hardware for acceleration would be a major undertaking. Using the marginal likelihood would also be possible, and would not be sensitive to outliers with a reasonable observation model (such as heavy-tailed noise), or pre-processing approach.

However, I also agree these changes in this instance are not likely to have a major effect on the glacier mass volume change predictions, especially since the Gaussian process is being used for interpolation rather than extrapolation. I also think the authors did a good job of validating the variogram approach they used, and that this approach does have some benefits of interpretability. I particularly appreciate the new content that was added to the supplementary material.

Overall, this is a thoughtful, exciting, and well-validated study.

We thank referee #4 again for their thoughtful comments on our study, which have helped to clarify and enrich the description of the methods and uncertainty analysis.

References for answers to referees

Velicogna, I. *et al.* Continuity of Ice Sheet Mass Loss in Greenland and Antarctica From the GRACE and GRACE Follow-On Missions. *Geophys. Res. Lett.* **47**, L11501 (2020)

Larsen, C. IceBridge UAF Lidar Scanner L1B Geolocated Surface Elevation Triplets, Version 1. (2010, updated 2020)

Beyer, R. A., Alexandrov, O. & McMichael, S. The Ames Stereo Pipeline: NASA’s Open Source Software for Deriving and Processing Terrain Data. *Earth and Space Science* **5**, 537–548 (2018)

Cauvy-Fraunié, S. & Dangles, O. A global synthesis of biodiversity responses to glacier retreat. *Nat Ecol Evol* **3**, 1675–1685 (2019)

Alexandrov, O., McMichael, S. & Beyer., R. A. IceBridge DMS L3 Ames Stereo Pipeline Photogrammetric DEM, Version 1. (2018)

Girod, L., Nuth, C., Kääb, A., McNabb, R. & Galland, O. MMASTER: Improved ASTER DEMs for Elevation Change Monitoring. *Remote Sensing* **9**, 704 (2017)

Rupnik, E., Daakir, M. & Pierrot Deseilligny, M. MicMac – a free, open-source solution for photogrammetry. *Open Geospatial Data, Software and Standards* **2**, 14 (2017)

Bostock, Hugh Samuel, *Physiography of the Canadian Cordillera. E. Cloutier, King's Printer* (1948)