Review 2:

We thank Dr. Lauren Vargo very much for her time in reviewing our manuscript, and for her comments and suggestions. We have responded with comments interspersed into her review.

This manuscript investigates the influence of anthropogenic forcing on glacier mass loss over the industrial era. The authors use a 3-stage model to simulate glacier length and mass balance changes for glaciers with different response times. I'm less familiar with this model compared with other glacier models, but it was well-described and the original paper was cited to better understand the model. The authors use different climate inputs: synthetic temperature, proxy temperature, and model temperature and precipitation. These climate inputs cover the last millennium to ensure that glaciers, even those with response times of 400 years, have reached climatic equilibrium over the industrial period for which anthropogenic forcing is calculated. Running these longer simulations is where this paper builds on previous attribution studies, particularly Marzeion et al., 2014.

This paper is well-written, the structure makes it easy to read, the figures are clear and add to the manuscript, and the paper provides insight into an important topic: investigating the influence of anthropogenic forcing on glacier mass loss, in a way that includes inherited conditions.

The main question/concern I have is if the anthropogenic contributions are calculated in an appropriate way. I understand, mathematically, why presented anthropogenic forcing values can exceed 100% (if a glacier losing mass in full forcing scenario would have gained mass in natural forcing). But, if you're looking to calculate the anthropogenic contribution to glacier mass loss, should you do so in a way that calculates that contribution to be between 0 and 100%? As negative values and values over 100% don't realistically make sense. I think calculating the influence of anthropogenic forcing in another way so that values are between 0 and 100% would make the paper and results clearer and less confusing, giving it a higher impact. I wonder if there is a better way to do this using cumulative mass change from _1880 through the end of the simulation. It also seems important that a small change in temperature between that shown in Figures 2 & 3 to Figures 4 & 5 do impact the contribution of anthropogenic forcing percentages by 1/3 to almost 1/2 for different response times.

Both reviewers raised the issue of this fractional metric. We've specifically chosen it because it is used extensively by the IPCC; and it was used by Marzeion et al (2014, M14), whose analysis we parallel. Moreover, the headline result of M14 was given in terms of this metric, and the result has been cited in subsequent IPCC reports and elsewhere in the literature. The word 'contribution' causes a strong reaction in some, especially when the number is negative or exceeds 100%, even if the equation hasn't changed. We note that the original M14 analysis allowed for a negative value in the uncertainty range and that, for anthropogenic warming, the current IPCC uncertainty range allows for it to exceed 100% of the observed. Apparently, there were divergent opinions in IPCC discussions over the use of word "contribution" and reports now avoid the word; and cast the assessment as the magnitude of anthropogenic warming relative to the observed. In our revisions we adopt this: we have removed the word contribution from the manuscript anywhere it might have this connotation. And we describe the analysis everywhere as the magnitude of the anthropogenic mass loss relative to the observed (or equivalently, relative to the total). In addition to these changes throughout the manuscript, when we introduce the metric, we now give readers two substantial paragraphs detailing the reasons for the metric, guiding them about how to interpret it, and now tell them ahead of time that we also analyze another metric in a later section. We hope that this sets a reader up better to interpret our results.

The two paragraphs are reproduced here: *"We here take a moment to discuss various possible metrics that can be used to characterize the influence of anthropogenic factors on glacier mass balance. Each have different merits. One approach is to set up a null hypothesis of a counterfactual, purely natural climate with no anthropogenic forcing, and to pose the question "how likely is the observed change in the natural case?". If the likelihood of the observed change falls below some stated level of significance, then the null hypothesis is rejected, and a "signal" can be declared detected. This was the approach applied to glacial retreat in Roe et al., 2017, and we apply it to mass balance in Section 4.4 of this paper. Another approach is to define a measure of the variance explained among an ensemble of climate models where natural variability, different anthropogenic forcing scenarios, and different models all contribute to the spread (e.g., Hawkins and Sutton, 2009; Marzeion et al., 2020), which is helpful for partitioning sources of uncertainty. A third approach, one that has been used by the Intergovernmental Panel on Climate Change (IPCC) over several past reports, is to estimate the magnitude of the anthropogenic component relative to the observed change. For instance, as of the 2018 IPCC Special Report, the central estimate of the magnitude of the anthropogenic change in global-mean surface temperature over the industrial era is that it is equal to 100% of the observed warming, with an assessed likely (i.e., a 2-in-3 chance) range of ±20% (Allen et al., 2018). Note this uncertainty range allows for the possibility that the magnitude of the anthropogenic warming exceeds 100% of the observed, because it is possible that the natural climate would have otherwise been cooling. M14 followed the IPCC in using this approach. It is useful, for instance, in the context of interpreting sea-level rise, where other contributions from thermal expansion and large ice sheets need to be understood. We want our results here to parallel those of M14, and we adopt the same approach as M14 in Sections 4.1 and 4.3 of this study.*

M14 defined the following metric for comparing the magnitude of the anthropogenic mass loss to the FULL mass balance: $F_{anth} = 100\% \times (ANTH/FULL) = 100\% \times (1 - NAT/FULL)$. *M14 calculated the 20-yr running mean of* F_{anth}*, in order to assess how anthropogenic mass loss has evolved over time at decadal scales. For assessing the total change in mass over the industrial era, ANTH and FULL can be integrated over time before calculating their ratio. Note that this cumulative value will not, in general, be equal to the time-average of* F_{anth}*. In our analyses, we present both the decadal and cumulative values. Both provide useful information: the 20-yr running mean provides insight into the causes of decadal-scale mass-balance observations, whereas the cumulative anthropogenic mass loss can be more directly tied to glacier retreat (see Section 4.4) and sea-level rise."*

Specific comments (intermediate):

L49-50: This line makes it feel like this is an important part of this paper- assessing all glaciers for the contribution to sea level rise. Should this go in the title, or at least abstract? But then I was waiting for the results to be tied in with sea level rise later on, but it wasn't discussed.

We specifically don't tie into sea-level rise with our main results: we can make strong statements about the relative importance of natural variability and anthropogenic climate change in mass loss, but we do not estimate absolute mass loss. However, our results should affect interpretations of what has caused the sea-level rise associated with mass loss from ice caps and glaciers, and we now include a specific sentence about this in the discussion to clarify.

Overall, I wonder if it's important to emphasize that these are idealized scenarios: all glaciers globally are represented by just five different response times, and using a model that doesn't include

The end of this comment got clipped out of the review, unfortunately. In one limit, all models are idealized scenarios. Our central result is showing that the magnitude of the anthropogenic mass dominates the total mass loss for all glacier response times. That is arguably a more powerful result than highly detailed simulations of a specific setting: it is more general, and shows that those very specific details (which have their own uncertainties of course), don't actually matter to understand the overall role of anthropogenic climate change in glacier mass loss.

L308-316: Are the solar irradiance and orbital changes ensembles not analyzed at all? If you only use one and volcanic-forcing has the biggest impact, using that one makes

sense. But why not include all (besides too much data/run time)?
There are not many millennium-scale GCM ensembles (more are gradually becoming available). The group that made these integrations did consider irradiance and orbital changes, but only individually, and then all together when they also included the anthropogenic CO_2. That is, there is not a run with all natural forcings and no anthropogenic forcing. It also is a GCM that warms less than observations over the 20th century. That is less than ideal for us, but doesn't prevent the essential point being made. We've included an expanded discussion of the model to guide a reader more helpfully.

L419: This goes along with my main comment, but results show anthropogenic contribution of over 100%.
Hopefully our reframing of the language around 'contribution' helps in the revised manuscript.

Specific comments (minor/technical):

L35-36: Do we know this for sure/are there studies that have shown this? Is there a citation to use?
We tweaked the language to make it clear it comes from Zemp et al. (2015).

L41: "An alternative approach", referring to an alternative to Roe et al., 2017?
Thank you, we've tweaked the language to clarify.

L46: Vargo et al 2020 shows a method that applies the method discussed here but to glaciers with limited observations (but still needs some observations, so still a subset of the total).
Thanks. Included as a reference.

L58: Anthropogenic forcing would account for some observed retreat? 25% of mass loss due to anthropogenic forcing isn't nothing.
Agreed, but if that was all it was (25%), it would stand in stark contrast to the use of glacier imagery to educate about the impact of anthropogenic climate change. Obviously, we hope our study resolves that apparent contradiction.

L61-70: Good setup/description for part of the problem you're addressing
Thank you!

L159-174: I think these assumptions are reasonable (some need to be made). But for Arctic glaciers, if 10% have a response time of over 400 years, those are probably some with the most ice mass? With larger contributions to sea level rise than smaller glaciers with quicker response times? Maybe just something to note.
Thanks, yes, we note in the discussion that for sea-level change, the upper-bound of the uncertainty estimates might need to be considered. Our central result doesn't change because a 400-yr timescale is already effectively at the large-tau limit.

L197: I wonder if/how much changing the 0.5C of white noise influences results? Maybe discussed in Roe et al., 2017.
Our equation is linear, so the amplitude of the noise leaves the impact of the underlying trend untouched. In other work (including in Roe et al., 2017), we have considered the impact of persistence in interannual variability, which can drive larger natural variability of the glaciers. Observed interannual variability in summer temperature and winter precipitation is quite close to white noise (see, e.g., Burke and Roe, Climate Dynamics, 2013)

L245: This paragraph is a nice overview.
Thank you!

L249: 'zThe' typo
Thank you!

Section 4.2: I know papers have shown that for many glaciers around the world, summer temperatures are especially important in overall mass balance. But was a summer temperature reconstruction used because that's what is available, or because it's the best value to use?
We chose it in preference to annual-mean reconstructions because it was a relatively recent publication, and because of summer's importance for melt. For the purpose here, we could have used several others equally well.

L319: Why is precipitation considered now but not with idealized scenarios? And how is it incorporated here versus in equations for the idealized scenario?
We omit precipitation for the idealized scenarios because precipitation mainly acts as a noise maker, rather than affecting trends. We included a section in the supplementary explaining why temperature trends dominate mass-balance trends. We now point readers to this section, and discussion why we omit precipitation when introducing the synthetic scenarios. For these experiments based on climate model output, we included precipitation for completeness and we now describe in the main text how precipitation is included.

L360-364: I haven't seen this done before, but I think it makes sense.

L378 - 379: Seems odd to say it could have been done a completely different way (different input data) but results would obviously be the same. Is it obvious?
The magnitude of interannual variability is sufficiently well represented by models, which is all the results depend on. We changed to use the word "conclusion" rather than answer

Fig 9: Is it useful to compare where measured mass loss falls/plots?
It is a nice idea. However, we haven't tried to simulate any specific glacier. In principle we certainly could, but then we'd really just be calibrating parameters to get agreement. We don't have direct mass-balance measurements going back to 1880, although reconstructions from 1865 for four glaciers in the Alps give similar magnitudes of cumulative mass balance (see Huss et al., 2008, JGR, Fig. 6).

Figure 10: This took me reading a couple of times to understand, but it helps show that a problem with M14 is that glaciers with long response times (100, 200, 400 years) haven't actually reached climatic equilibrium.