

Review “New Hydrometeorological Observations from Inglefield Land and Thule, NW Greenland” by Esenther et al.

The authors present new field instruments to measure river stage (water level) at three locations in northwest Greenland over 2019-2021. Instruments include a CF Bubbler, laser rangefinder units, pressure transducers and time-lapse cameras, recording stage values at different temporal resolution (15-min, hourly, 3-hourly) with various uncertainties and data gaps. To address this, the authors introduce a hybrid product that merges the most accurate sensor estimates from the extensively monitored Minturn River location. The authors highlight the added value of the hybrid product which yields the highest accuracy, while having the shortest data gaps among all other products. Using remote sensing (precipitation and albedo) and meteorological data (shortwave radiations, air temperature) from an automatic weather station installed at Minturn River, the authors identify the drivers of river stage changes.

I enjoyed reading the paper, which is very well written, and thoroughly describes the instruments used to estimate river stage, including uncertainties, pros and cons of each technique. The presented data set will undoubtedly be of high interest to the community. However, I missed a more elaborate analysis on the potential drivers of river stage variability in section 3.4. I also have some concerns about the “suggested purpose” of this new data set, i.e., whether it can be used for climate model evaluation. Below, the authors will find my General, Point and Stylistic comments. Based on these, I recommend **major revisions** before acceptance.

General comments

1. The abstract, introduction and conclusions highlight the need for new surface runoff estimates to evaluate (SMB) climate models. However, the actual river stage as measured in this study is not simulated by climate models, which estimate runoff fluxes, preventing a direct comparison. If possible, the authors could attempt to estimate runoff fluxes combining the hybrid stage product with the riverbed level and section. The resulting runoff flux data set could then be directly used for climate model evaluation. If this cannot be achieved with available measurements, I recommend reformulating the statements about model evaluation (e.g., L17-19, L39-41, L45-46, L50-51, L71-73, L408-409), and explicitly state that river stage is measured, but not the runoff flux, to avoid confusions. The manuscript title could be reformulated as: “River stage observations from ...”
2. The main results section 3.4 is relatively short. The correlation analysis between the hybrid product and meteorological variables to identify drivers of river stage variability is interesting. However, no figures support the analysis, making it hard for the reader to interpret the results. The authors could consider adding scatterplots between each tested meteorological variables (x-axis) and the hybrid stage product (y-axis) and provide corresponding R^2 and p-values within each graph. The authors could also elaborate on the impact of e.g., precipitation (rain), air temperature, cloudiness and shortwave radiation, ice albedo on the recorded stage variability at Minturn River. This is initiated in L345-353 and L363-366, but a broader analysis would be beneficial.
3. The discussion section currently repeats some information presented in the results section (e.g., L368-386), suggesting that these sections could be merged.

Point comments

L36: Runoff and solid ice discharge are the two major contributors to GrIS mass loss, I suggest reformulating as: “Besides solid ice discharge, climate change-induced meltwater runoff is a dominant driver of Greenland ice sheet (GrIS) mass loss that is projected to increase ...” or something equivalent. The authors could consider providing more recent references including e.g., Mouginot et al. (2019), King et al. (2020) on the contributors to GrIS mass loss, and Trusel et al. (2018), Noël et al. (2020) or Hofer et al. (2020) on the meltwater runoff increase in the 21st century.

L38-39: What do you mean by mass budget residual? In climate models, runoff is estimated using tipping bucket snow models that represent surface melt (solving the surface energy budget (SEB)), meltwater retention and refreezing in firn layers, and subsequent runoff. Could you clarify the statement?

L39-41: I am confused here, as the river stage measurements cannot be directly compared to modeled runoff fluxes from climate models. See also General comment #1.

L45-47: Mankoff et al. (2020) compared modeled and measured runoff from different locations, including Qaanaaq in NW Greenland and Zackenberg in NE Greenland. Please, clarify.

L61-62: Runoff discharge measurements are available from Qaanaaq in NW Greenland in Mankoff et al. (2020). Please, clarify.

L123: I am not sure about the journal policy on unpublished references. Either remove or use "personal communication". The same holds across the manuscript.

L182-183: Do you mean that the two other watersheds are too small to derive significant statistics? Please, clarify.

L209-210: For clarity, the authors could state that laser rangefinder M1, M2, N1 and F1 located in the three regions will be referred to as Lidar M1, M2, N1 and F1 in Figures 5-7. This should be done after introducing each of these sensors for the first time.

L216-219: Could you refer to the terms of equation 1 in the main text: e.g. "surface below the laser box ($Z_{\text{Lidar Box}}$) ... distance to water surface (Median Lidar Distance) ... vertical angle of the laser range finder ($\theta_{\text{Lidar Box}}$)"

L235: Could you highlight these lowest 75% stage values in Fig. 5 e.g., using colored shades? The reader can thus directly notice when the camera data set has the largest uncertainty.

L246: What do you mean by "predict CF Bubbler"? Fill the gap in that data set?

L255-265: The assessment described here is not supported by any figure/table and is only briefly discussed in section 3.4. See also General comment #2.

L289: Do you mean "July 15" instead of "June 15"?

L343: In section 3.4, could you refer to Figures 5-7 where appropriate, and explicitly state the range of discussed "days of the year".

L345: The authors should state that the AWS data are only available at Minturn River.

L348: "diurnal cycle of river stage"?

Style

L20: "SMB model evaluation" instead of "validation". L40: "evaluating", L46: "evaluation" and L51: "evaluate".

L32-33: For conciseness: "hydrological observations that are freely available ..."

L45: What do you mean by "SMB runoff", "modeled runoff"? Please, clarify.

L50: "quantify runoff drainage to the ocean and evaluate SMB models."

L60: "thus require"

L64: Maybe: "This paper describes new hydrometric sensor installations, and the resulting 3-year time series (2019-2021) of river stage (water level) at three proglacial gauging sites in NW Greenland."

L66: Define the acronym AWS here, and use it across the paper e.g., L90, L109 "AWS".

L73: As the acronym GrIS was defined earlier, you can now use it across the manuscript.

L95: It would be good to introduce the acronym PT for "pressure transducer" somewhere in the text.

L113: “pressure transducer that measures ...”
L100: Remove the dot after “riverbank)”. **L116:** Add a dot after “1.5%”. **L187:** Add a dot after “2023)”
L145: “is achieved via”
L184: “IR obtained from the integrated”
L195-196: Maybe: “due to turbulent flow as confirmed by the time-lapse camera imagery.”
L213: “used to calculate”
L258: “meteorological station”
L259: “difference between shortwave/solar downward and upward radiation.”
L268: “two bank-mounted laser ...”
L315: Remove “producing”.
L316: “low quality” instead of “suspect quality”?
L336: “yields the most complete”
L339: Remove “timescale” after 15-min.
L374: Replace “Having said that,” by “However,”
L409: “predict surface mass loss from runoff in an understudied”.
L412: “upward/downward shortwave/solar radiation”

Figures and Tables

L255: Add a reference to Table 1 after “hybrid product)”.
L285: Refer to Fig. 6 after “beginning in 2021”.
L295: Refer to Fig. 7 after “June,7 2021”.
L316: Refer to Fig. A1 after 75%.
Figure 1: Explain what the yellow star represents in the caption.
Figure 2: L438: “bedrock-dominated ... to study surface runoff without ...”
Figure 4: Spell out what PT means.
Figure 5: The authors could highlight when the camera data are low quality with e.g. colored shades. Please, also explain what Lidar M1 and M2 refer to in the caption. The same holds for Figs. 6 and 7 with Lidar N1 and F1.
Table 1: The caption should briefly describe the information listed in the Table, not provide an analysis of its content. Could you reformulate? Also explicitly state that these statistics only refer to the Minturn River location.
Figure 8: Use the same x-axis in all subfigures (see Net shortwave/solar radiation).
L484: “net shortwave/solar radiation”.
Appendix A: For clarity, “by -6 m or -9 m” instead of “by one of the two values”.
L500: “were observed in the laser rangefinder ...”

References

Mouginot et al. (2019): <https://www.pnas.org/doi/10.1073/pnas.1904242116>
King et al. (2020): <https://www.nature.com/articles/s43247-020-0001-2>
Trusel et al. (2018): <https://www.nature.com/articles/s41586-018-0752-4>
Noël et al. (2020): <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020GL090471>
Hofer et al. (2020): <https://www.nature.com/articles/s41467-020-20011-8>
Mankoff et al. (2020): <https://essd.copernicus.org/articles/12/2811/2020/>