Arctic supraglacial lake derived bathymetry combining ICESat-2 and spectral stratification of satellite imagery

Jinhao Lv et al.

Review by Ian Willis

Intro

This is quite an interesting, small-scale study – deriving bathymetries of 4 lakes on the GrIS using a combination of Sentinel-2 and ICESat-2 data. The paper compares what the authors refer to as a "classic bathymetry inversion model", notably the "log-transformation linear regression model (Lyzenga model)" with a modified version of it, which calculates depths separately for different spectral bands and then combines the results to produce an overall bathymetry. It appears as though the ICESat-2 data are used to calibrate parameters in the 'classic' and 'stratified' versions of the Lyzenga model, and then the two bathymetries for the 4 lakes are compared (validated) against ArcticDEM bathymetries (strips collected 2-4 months before the Sentinel data). The 'stratified' version of the Lyzenga model does a modestly better job at capturing bathymetries, as measured using R^2, RMSE and MAE, when compared against the ArcticDEM data.

I have a few main issues with the paper and recommendations for improving it, and a long list of small line by line comments, which I suggest would need to be addressed too before the paper is ready to be published.

Main Comments

1. The paper does not adequately acknowledge all the work that has been done by many people over the last decade +, which uses a physically-based algorithm to calculate water depths on the GrIS or Antarctic ice shelves from optical imagery (MODIS, Landsat, ASTER, Sentinel-2). For example, none of the following papers is mentioned or acknowledged in this context.

Philpot, W.D. 1989. - Bathymetric mapping with passive multispectral imagery. Appl. Opt., 28(8), 1569–1578. Original algorithm.

Sneed, W.A. & Hamilton, G.S. (2007) – First time algorithm used on ice mass using ASTER to estimate supraglacial lake depth/volume on the Greenland Ice Sheet. AGU Publicationsdigitalcommons.library.umaine.edu

Sneed, W.A. & Hamilton, G.S. (2011) – Validation of the method with satellite imagery over glacial melt ponds (Greenland), confirming the approach's accuracy. Cambridge University Press & Assessment+1

Georgiou, S. et al. (2009) – Applies and extends the Sneed & Hamilton method to track seasonal evolution of a West Greenland supraglacial lake with ASTER. Cambridge University Press & Assessment+1

Banwell, A.F. et al. (2014) – Uses the Sneed & Hamilton method to estimate supraglacial lake depths on **Larsen B Ice Shelf (Antarctica)** and in Paakitsoq, Greenland. Aberystwyth Universityrepository.cam.ac.uk

Pope, A. et al. (2016) – Implements/assesses the radiative-transfer (Sneed & Hamilton–type) depth retrieval with **Landsat-8** in West Greenland and compares with other multispectral methods. Copernicus TC+1

Moussavi, M.S. et al. (2016) – Derives and validates supraglacial lake volumes over West Greenland using **WorldView-2** with the RTE/Sneed-Hamilton framework. ScienceDirect+1NOAA Institutional Repository

Williamson, A.G. et al. (2017) – "FASTER" workflow applies the Sneed & Hamilton physically based algorithm to MODIS for automated lake area/volume tracking on the Greenland Ice Sheet. ScienceDirect+1

Williamson, A.G. et al. (2018) – Dual-satellite (Sentinel-2 & Landsat-8) analysis over Greenland; depth estimates follow the Sneed-&-Hamilton-style RTE with recommended parameters. Copernicus TCstatic.cambridge.org

Moussavi, M.S. et al. (2020) – Detects **Antarctic ice-shelf** lakes (Amery, Roi Baudouin, Nivlisen, Riiser-Larsen) and computes depths/volumes from **Landsat-8/Sentinel-2** using the Sneed & Hamilton-derived approach. MDPI

Dell, R., et al (2020). Lateral meltwater transfer across an Antarctic ice shelf. ***The Cryosphere, 14***(7), 2313–2330. https://doi.org/10.5194/tc-14-2313-2020. Uses the algorithm to calculate water depths from Landsat imagery across an Antarctic ice shelf.

The paper below is mentioned, but not explicitly in the context of using the 'Philpot' approach:

Melling, L. et al. (2024) – Intercompares RTE (using literature coefficients from Sneed & Hamilton and others) with ICESat-2/ArcticDEM across lakes in **southwest Greenland**. Copernicus TC+

The well used 'Philpot' approach should be acknowledged and it would also be useful to mention (or show) how the 'Philpot' equation varies from the classic and modified Lyzenga models. Of course, it would also be valuable to actually use the 'Philpot' equation (e.g. with the parameters advocated by Pope et al 2016) to compare against the 2 calibrated versions of the Lyzenga model and the ArcticDEM.

2. The paper remains rather limited in scope. In essence it has applied two versions of an empirical model: one, the 'classic' Lyzenga model, that has already been applied to a GrIS lake by Lv et al 2024; and a modified version of it that uses

different spectral bands, that has so far only been applied outside of the context of GrIS lakes. The models were calibrated using ICESat data and validated against ArcticDEM data. There appears to be quite marginal improvement in using the modified version of the model. It is a methods paper, and it doesn't tell us anything about the glaciology or hydrology of GrIS surface lakes. It would be a more compelling paper if it had gone further.

It could have remained methods based but compared also the 'Stumpf' model and the 'Philpot' model for the 4 lakes. It could have told us how transferable the model is – it appears as though the model was calibrated separately for each lake? What were the different parameter values associated with these calibrations? What happens if you apply the model calibrated for one lake to the other lakes? What errors occur? How universal is the model? What happens if you combine all the ICESat-2 data with all the corresponding spectral data to produce an 'all lake' model? How transferable is that? The model needs testing against data that are outside of the lakes that were used for calibration.

It could also have been expanded to tell us about the hydrology of the lakes by calculating their volumes over summer seasons and between years.

3. The work uses the NDWI (which uses the Green and NIR bands) to separate water from ice. This is unusual as virtually all the papers I know about (see e.g., those listed above) use the variant of NDWI tailored for glacier ice and water separation that uses the blue and red bands instead of green and NIR:

$$NDWIice = (Blue-Red)/(Blue+Red)$$

Did the authors consider using this? Why did they opt for the Green and NIR bands? Could we be shown whether it makes a difference to the results?

4. There is confusion throughout the paper about the definition of what the authors refer to as "a satellite-derived bathymetry (SDB) method". See lines 12-15 where it is first mentioned and defined, but then all subsequent references to it as well. Need to clarify that this is the *complete* method that is being used in this paper, **i.e. the Lyzenga model applied to optical satellite data (Sentinel-2), calibrated using altimetry data (ICESat-2).** And the paper uses two versions of the Lyzenga equation. It is not just the use of optical data alone as is sometimes implied in the manuscript.

5. The methodology is not clearly articulated in a consistent way throughout the paper, and parts of the methods appear in the results (see line by line comments below). A key aspect of the methodology that is not articulated clearly up front is that the ICESat-2 data are being used to calibrate the parameters of the two versions of the Lyzenga model. The equation for the 'traditional' version of the Lyzenga model and how it is applied are not explained. Does the traditional version use just one waveband from the optical data? If so which one?

Line by line comments

15 ‘verify’ => ‘validate’ [as ‘validate’ is used subsequently in the paper].

17. say you’re using ‘time stamped’ DEM strips here.

19-21. Do you need a sentence before this one explaining what you did to enable you to make this statement? How did you prove it is 'scalable' and what do you mean by that? I’m not convinced you’ve proved it’s scalable, have you?

26. Is Luthje et al a good reference here? Does it explicitly consider ‘ecological’ issues?

28-30. Beckmann and Winkelmann, 2023 do not consider lakes in their paper as implied. Clarify what you mean by this sentence and use an appropriate ref.

32. **You can't sail ships on ice sheet surface lakes so delete this!** Your 3 refs all relate to lidar. Could you quote studies using bathymetry from small boats, e.g.:

1. Box, J. E., & Ski, K. (2007). *Remote sounding of Greenland supraglacial melt lakes: implications for subglacial hydraulics*. Journal of Glaciology, 53(181), 257–265. https://doi.org/10.3189/172756507782202883. Cambridge University Press & Assessment
2. Tedesco, M., & Steiner, N. (2011). *In-situ multispectral and bathymetric measurements over a supraglacial lake in western Greenland using a remotely controlled watercraft*. The Cryosphere, 5, 445–452. https://doi.org/10.5194/tc-5-445-2011. Copernicus TC

Both papers include in-situ depth sounding (Box & Ski with a raft + depth sounder; Tedesco & Steiner with a remotely controlled boat equipped with GPS and sonar).

32. You say ‘in polar regions’ here. But I’d suggest focussing on the GrIS earlier in your introduction and stop referencing Arctic / Polar regions after that.

41. These references are not on the GrIS so make that clear in the sentence. You could reference other papers relevant for the GrIS and AIS.

42-3. “…integrated multispectral technology with ICESat-2 to conduct bathymetric detection and inversion, leveraging both active and passive remote sensing…” Here and elsewhere in the paper (e.g. 47-8, 64) it would be useful to stick to one order and not switch the order you mention these two types of satellite data. So here you could write “integrated multispectral technology with ICESat-2 to conduct bathymetric detection and inversion, leveraging both passive and active remote sensing”. It’s a trivial point but it makes for a clearer, more logical read.

44. You refer to ‘the mainland’ but you need to be clearer in this section that the methods you’re using were first applied in settings outside the GrIS. Tell us what mainland here as it reads like you’re talking about mainland Greenalnd!

46. ”…leverages active and passive remote sensing techniques…” You told us that 3 lines up so delete. More useful to tell us what sensors .

50. Full stop after ‘Bermuda.”

50. ‘proposed’ is the wrong word. Do you mean ‘divided’? Or ‘classified’?

54. You say “improving the inversion accuracy” but compared to what?

56-7. Before the sentence spanning these lines, I'd specify that you're focussing on surface lakes on ice masses. Note it's not just for 'Arctic regions’ as you include Fricker et als work on Antarctic surface lakes in your referneces.

61. “Watta algorithm”. This is the 3rd algorithm that’s been introduced now with no details. It’d be more important to tell us the basis of this (and earlier) algorithms not what Datta and Wouters called it. I assume they called it this because it is an amalgamation of their names, but this is not really crucial information here.

61. YOU say ‘arctic lakes’ but it was just for GrIS lakes. Having already focussed down on the GrIS in your review, I’d stick to referencing work relevant to the GrIS here and not keep mentioning ‘Arctic’ or ‘Polar regions’ etc.

63. Delete ‘some’ and delete ‘from 2019 to 2023’ This latter not relevant.

65. ‘polar’ See my comment above for line 61.

65-7. Is this just the Lv et al (2024) paper that’s relevant here? If so say so. But you should also acknowledge Mousavi et al, Pope et al, Williamson et al, Melling et al, etc. who did consider different bands using the ‘Philpot’ algorithm.

68 Suggest “…Chu et al. (2023) to offshore islands…”

68-70. So you're building on Lv et al 2024 which so far is the only paper to apply what you're calling the SDB method to GrIS lakes, but you're adding spectral stratification as used by Chu et al in a different context? Could state this more clearly

72-3. Ok so you’re 'calibrating' using ICESat-2 and validating with ArcticDEM. This could be expressed more clearly.

74 ‘Arctic’ => ‘GriS’

74. “…offers effective technical support for predicting Arctic glacier melt and global climate change.” This is too grand. I’d delete this. You can’t do this based on your work.

75-7. Suggest delete.

80-2. Suggest delete “in the Arctic, the second-largest ice sheet in the world, surpassed only by the Antarctic Ice Sheet. However, the GrIS is more fragile and sensitive to temperature changes than the Antarctic Ice Sheet (Robinson et al., 2012).” Everybody knows the first statement and the 2nd is a bit vague.

85. ‘aimed => ‘aims’ and “verify” => ‘validate’

86 “bathymetry data …were used” [data are plural]

91. Its blue green red yellow respectively

Fig 1. Your Arctic inset is rather ugly. I suggest use another inset - just for the GrIS - and remove the words 'study area' from the map.

100-1. Suggest change to “Sentinel-2 imagery was obtained for lakes A and B on 4 July 2020, lake C on 17 July 2022, and lake D on 15 July 2021.”

102. ‘can be’ => ‘was’ [say what you did not what is possible to do]

110-11. Suggest “The left and right points of each beam pair are approximately 90 m apart in the transverse track direction and about 2.5 km apart in the along-track direction.”

112-13. Suggest : “The ICESat-2 data used in this study were acquired on 6 July 2020 (lakes A and B), 15 July 2021 (lake C), and 14 July 2022 (lake D).”

115 “data, and”

115-16. “Please note that this study assumes” => “We assume”

118 “can be” => “data were”

119 delete “and has the characteristics of a large coverage area and high spatial resolution.”

121. Delete “high-resolution Arctic digital elevation model ArcticDEM”

122-4. Ambiguous. Tell the reader what you did - what data you used.

124. Delete “It should be pointed out that”

124-5. Explain why these dates. Presumably the first data to cf. with your derived bathymetry along the ICESat-2 line for lakes A and B, 2nd for bathymetry of lake C, 3rd for lake D?

130-1. You say “To address the challenges and the limitations of traditional bathymetric methods, which do not consider the varying penetration of electromagnetic waves of different wavelengths into water…” But this is not true as many previous studies using the ‘Philpot’ method have done this.

Figure 2. This implies you’re producing two versions of the Lyzenga model - traditional and spectral stratified. And you'll compare them both against Arctic DEM data? This was not mentioned as part of your methodology earlier in the abstract or in your brief overview of the methodology for the work on lines 68-73. It should be clearer earlier that this is part of your work.

140. Delete “provided by the ESA” and say “…data are…”

142. ‘…can be processed…’ Again, tell us what you did not what can be done.

144-6. Why did you not use the version for separating water from ice? NDWI_ice? there’s another variant of NDWI tailored for glacier ice detection that uses the blue and red bands instead of green and NIR.
That formulation is:
NDWIice =(Blue-Red)/(Blue+Red)

- Blue: reflectance in the blue band (~0.45 μm, e.g., Sentinel-2 Band 2)
- Red: reflectance in the red band (~0.65 μm, e.g., Sentinel-2 Band 4)

151. ‘data are’

152. Delete “In this study”

154 and 155 “should say ‘the four’

155 “Due to the fact that” => ‘Because”

157-8. “Finally, more accurate bathymetric photons were obtained for constructing a bathymetry inversion model.” I don’t understand this sentence and it splits up the two either side of it that are related. I suggest delete this sentence (or clarify what it means if it’s important).

169. Delete “It should be noticed that”

176-7. Suggest “ This study applied the spectral stratification method using the Otsu algorithm, which automatically determines thresholds without requiring input parameters (Otsu, 1975).”

177-8. This is not a sentence. The main clause is missing a finite verb. Right now, “multispectral images of water stratified into four layers” is written as if it’s a complete statement, but it lacks a clear subject performing an action. Should it say '...water were stratified..." ?

But 'stratified into 4 layers' implies lake depth layers, but you don’t mean this - you’re giving 4 wavebands. This is confusing.

179. Could delete ‘layer’

189. It’d be useful to show the equation for this “traditional Lyzenga model” too.

191. You say “...combining the traditional Lyzenga model...” But with what? This is unclear.

192-3. “the near-infrared layer, the red layer, and the green layer were combined for processing. In other words, Arctic SGLs were divided into green and blue layers.” This is contradictory. Did you use the NIR and Red or just the Green? Explain more precisely what you did.

197-8. So what values do these parameters take? You haven’t really explicitly stated you’re fitting these equations using the ICESat-2 data to derive Z, then you’re using these calibrated equations to derive bathymetry for the whole lake. Please in your methods explain the calibration process explicitly. In the results it would be useful to know what values these parameters take for the two equations for the 4 lakes. It’d also be useful to know what the equation for the original Lyzenga model is, how that works, and what the parameter values are for that and how they vary between lakes.

210 ‘performed’ => ‘performs’

211 ‘was’ => ‘is’

213. It’s only here that it becomes apparent that you're constructing 2 models. The original Lyzenga and a spectral stratified version of Lyzenga? This has not been clear throughout your methodology sections so far. For example, you do not mention this in Section 3.

216-22. This should all have been stated above in methods. Not here. Start the results section with the results!

229. Where have you used the active (ICESat-2) data to derive Fig 5? This was not adequately explained.

231. You say “…underscoring the reliability and feasibility of the spectral stratified model”. But you cannot conclude this until you’ve compared with the Arctic DEM. Delete this.

232. You don’t need any of these phrases ending with the word ‘that’, e.g. ‘It should be noted that…’. Just delete this and all such phrases.

232-3. You say “ the ArcticDEM data only contains spatial information of the lake bottom, and lacks water surface elevation information when obtaining bathymetry benchmark data”. Is that true? Esp as you use early season strips. How do you know the lakes were empty at that time rather than containing water and frozen over?

236-7. You say “However, the sediment in the lakes of the experimental area primarily consists of bedrock, a type of material that remains stable and does not undergo significant changes over short periods.” This makes no sense to me at all.

239. Having read to here, I’m not convinced you needed any of the previous section 4.1 on the qualitative analysis. Consider deleting it as the quantitative analysis is what is needed.

242. “spectral stratified Lyzenga model” . Check entire paper and refer to this model in the same way throughout. This is quite clear here but you’ve not so far ever referred to it in this way.

246. ‘visually demonstrate’ => “illustrate”

247-9. This sentence is obvious and could be deleted.

249-50. Do you mean to refer to Fig 7 here as it doesn’t show ArcticDEM validation. Do you mean Fig 8? If so refer to Fig 7 earlier.

251. “ spectral stratified model” See my comment for line 242. Try to refer to this model consistently throughout the paper.

256-7. You say “…are primarily concentrated in the 2-6 m range, with more pronounced differences at the transitions between different spectral layers. This is hard to see, esp. the last point as you don't mark on where the different spectral layer data were used to construct the bathymetry.

Fig 8. I'd assumed this would be a subset of Fig 6a, but for the data in the 2 sub-areas shown in Fig 7. But it's 'density' . Why are you showing this? It needs explaining.

276-7. Delete this sentence from here. This belongs in the Abstract and the start of the Conclusion.

278. You say “volume changes” but this implies you’re going to determine volume change through time, i.e. applied your final model over more time periods. But you don’t do this, which is a shame. Clarify what you mean by ‘volume changes’ here.

Figure 9. I’m not fully convinced these add anything although they're nice visually. Suggest put in Supplementary Materials. But what would be more useful to know is what are the volumes of the 4 lakes derived by the traditional Lyzenga method and from the Arctic DEM? How do the 3 estimates compare? Which overestimates vs. underestimates cf. others?

289-90. You say “While the spectral stratification-based method enhances the accuracy of bathymetric inversion compared to traditional approaches, such as the Lyzenga model…” But you can’t say this. You can't generalise. You can only refer to the traditional Lyzenga model here, not ALL 'traditional approaches' by which I assume you also mean the Stumpf model? As that is the only other one you mentioned in your Intro. Did you ever consider applying this one and comparing it? And of course it’d have been valuable and interesting to have applied the ‘Philpot’ method too and evaluated that.

291. ‘amount’ => ‘number’

291. What do you mean ‘strip-shaped data’? Do you mean “ orbital track” ?

291-2. You say there were not enough training sample data. But each lake seems well sampled along the tracks according to Fig 3. What was the total number of sampling points for each lake? Are you saying you need more? Is that really the case?

292. You say the lack of training data “…hindered further improvement in model accuracy” which is poorly phrased but more importantly, you can't conclude that. Perhaps extra data would not have improved accuracy (when cf. ArcticDEM). It MAY have improved the estimation of the model parameters and thereby led to greater accuracy, but you don't know this.

293-4. You say “…where bathymetric variations are minimal, the impact of spectral penetration differences is limited, resulting in only marginal improvements in accuracy.” Did you point out such areas? Looking at Fig 6 I do not see greater 'improvements' at high depths cf. shallow depths.

294. ‘Second,…’