Review of “Ice volume estimates for the Himalaya-Karakoram region: evaluating different methods”

Summary:

I very much looked forward to reading this paper. I’ve used scaling techniques in many publications, but I’m the first to stand up and celebrate a good engineering solution. The physics of scaling is fascinating, revealing, and informative; but when accuracy is key, I see no reason that we shouldn’t try to tease more information out of our data by using a well-designed numerical model. Accurate volume estimates seem like the perfect application for a numerical model. To that end, the authors have compared several volume-area scaling solutions to other popular and successful models of the type pioneered by Farinotti, Huss, Clarke, and others. I’m a fan of these numerical models, and I looked forward to the comparison.

The topic is especially timely. As the authors note, the calculation of total ice volume for a region is important for sea level rise estimates and for general water resource planning. Most previous analyses have used volume-area scaling, so a thorough comparison to the newer models is overdue. I think this paper does an admirable job on several fronts, but it derails significantly on a few others, and I am not convinced that the comparison has been definitive.

The treatment of the conceptual aspects of volume-area scaling is wildly inaccurate, and a real comparison between volume-area scaling and numerical models is not possible until the best scaling parameters have been identified for the Himalaya-Karakoram. Therefore, I’d like to see some significant improvements outlined below.

Details:

Pg. 4818, line 4: The historical popularity of volume-area scaling is not just due to the simplicity of the application. It’s also because area data has been historically easy to measure and to compile.

Pg. 4820: Some disadvantages of modeling are mentioned, but the inherent dangers of a numerical inversion are not discussed. Huss and Farinotti (2012), for example, average over long wavelengths to avoid large calculation errors at short spatial wavelengths. This should probably be discussed or briefly mentioned (more than in the cursory nod given to slope averaging on pg. 4822, lines 14-15). Improper averaging could lead to incorrect volume estimates or to volume precision that is not warranted. In the conclusion of the paper, the authors also mention thickness distributions as one reason to prefer numerical models over scaling. I agree, but the limitations (or at least inherent dangers) of an inversion should be acknowledged earlier in the paper, particularly when the thickness distribution could potentially be calculated on too fine of a grid.

Pg. 4824: The sensitivity analysis seems appropriate. However, I have one note of concern. The scaling techniques have three notable sources of sensitivity/error – errors in the area A, errors in the scaling exponent gamma, and errors in the multiplicative scaling factor c. (Of these, it could be argued that the scaling exponent is fixed by the scaling physics and technically cannot be a source of error, but that's not too important here.)

On the other hand, my understanding of the numerical models is that they have many more free parameters. But page 4825 mentions only f and tau. What about all of those parameters mentioned on page 4822 (like hmin, hga, r, and n)? I know they are optimized, but they are still potential sources of error. And what about the potential variability introduced by the selection of the 50m elevation intervals? Ditto for the parameters in Huss and Farinotti (2012); to count the free parameters, I just reread this paper, and there are very many. I do recognize that each of these model parameters has been selected to give a good fit to available data, but then the same could be argued about the scaling parameters of Arendt et al (2006) and Chen and Ohmura (1990). So if you are allowing these "fit" quantities to vary in the scaling approaches, then shouldn't we consider the potential variability of "fit" parameters in the models? If not, explain why.

In other words, it might be worth noting that an advantage of the scaling technique is that very few parameters are necessary, and that they can be fit to a particular region (as done in Arendt et al, 2006); and therefore the sources of error are more easily controlled and understood in a scaling approach. By the same token, an advantage of the numerical approach is that the extra free parameters give the solution more flexibility.

Pg. 4825, line 21: Ah, wonderful. My faith in glaciology is restored by the sentence

> "*Furthermore, V-A [scaling] relations are designed to estimate the volume of a larger glacier ensemble, but are not suitable to assess the [precise] volume of individual glaciers, which further hampers their comparison with [individual] measurements*."

Excellent! Very few publications acknowledge this important shortcoming of scaling approaches, and I am very pleased to see it acknowledged here.

By the way, I recommend inserting the word "scaling" as I have done above – there could be other types of volume-area relationships that do not involve power laws. And I recommend inserting the word "precise" as I have done above; we can use volume-area scaling to give an order of magnitude estimate of a single glacier's volume, even if it is entirely inappropriate to use scaling to get a more precise estimate for a single glacier. I would also insert the word "individual" as I have done above; because we could still compare sets of volume calculations to sets of measurements. For example, comparisons could be done between probability distributions or between sums of volumes (total or mean volume of all the glaciers in each set).

This last point is probably worth mentioning in this paper. If there were more volume measurements, then a comparison between the scaling approach and the data would be possible by looking at probability distributions or moments of the distributions (mean volume, variance, etc.). Regressions of the data would also give useful information for comparisons. So even if we cannot apply scaling to individual glaciers, a useful comparison with lots of volume measurements is possible, but unfortunately the data is not yet available.

Pg. 4829, line 23: Again, another excellent point. The separation of ice masses into separate entities creates a problem particularly for volume-area scaling approaches. I'm glad to see this mentioned, both here and on the next page.

Pg. 4830, line 1: Do you mean "Conceptual aspects" or do you mean "Conceptual shortcomings"? You only present shortcomings, but I'm sure there are some conceptual advantages which might be included under the title "Conceptual aspects".

Pg. 4830, section 5.2: Everything had been going so well in this paper, but at this point, the paper's logic really derails. Section 5.2 needs a major rewrite in order for this paper to be publishable. To be blunt, the paragraph starting on line 2 has so many conceptual errors, that I wonder if the authors have an agenda beyond a detailed comparison of the different techniques. Some of these statements can't be made with a straight face, and it's almost like the authors are winking or goading. I hope and assume that is unintentional.

Again, I'm sorry to be so blunt with this part of the review, but it is not possible for this paper to provide a fair comparison between volume-area scaling and numerical modeling techniques if such erroneous statements are believed (or at least written) by the authors. Keep in mind, that I'm inclined to agree with the authors' general conclusions that the numerical modeling approaches offer significant advantages and possibly even better accuracy and precision when compared to volume-area scaling. But you can't make that point with false statements about volume-area scaling. False statements will only weaken your argument, even if the final results are unchanged.

For example:

> *"Estimating a glacier's volume based on its area has three major shortcomings: (i) mean ice thickness and area have a weak correlation."*

Say what? Since when? (And why would that matter for volume-area scaling?)

No, Figure 8 does nothing to prove this point. Figure 8 shows only one order of magnitude in thickness! I'm amazed that you can see any power law correlation over such a small range of values. In fact, that must be surprisingly good data. Establishing a power law relationship from data (we're not talking theory at the moment) requires MANY orders of magnitude on each axis; for a discussion of this point, see the many excellent references in the literature, including Clauset et al, 2009. This has been discussed repeatedly in disciplines ranging from seismology to hydrology to planetary geology. With only one order of magnitude, this plot is entirely insufficient. For example, I can take any relationship (non-linear power law, exponential, linear, other) between any two measured variables, and if I choose a small enough segment of the plot, then the noise will appear to dominate.

Figure 8 only demonstrates that the thickness-area scaling relationship is very noisy in this region of the world (and yes, probably other regions too). The thickness-area relationship has a solid physical underpinning (i.e., physics) and is a *theoretically* valid relationship, but there are not sufficiently large glaciers in the Himalaya (and probably elsewhere) to make it a *practical* relationship in this region. Without larger glaciers, this plot can't establish a power law trend that rises above the noise. Now *that* might be worth stating in this paper, if you insist on talking about area-thickness scaling ,but it is not at all the same as saying area-thickness is an inherently noisy relationship (you don't have appropriate data to support that claim).

By the way, there is an obvious alternative to thickness-area scaling. Use a volume-area relationship instead! After all, you are trying to derive volume, not thickness. This paper's fixation on thickness scaling is baffling. Glacier volumes will cover many more orders of magnitude, so that the power law relationship can be reliably assessed. This might make it more difficult to compare scaling techniques to the existing thickness data (pg. 4825, line 26), but again, as the authors already imply (pg. 4825), that is a shortcoming in our ability to validate the technique, but it is not a shortcoming of the technique itself (as erroneously implied by section 5.2).

Pg. 4830, line 4. Continuing with the errors in section 5.2:

> *"The high correlation of volume and area comes from the self-correlation in the relationship of area and volume, as volume is the product of area and mean thickness."*

Oh no, not this nonsense again. This is just plain wrong with no room for interpretation. It has been refuted at least twice before in the literature by Lüthi et al (2008) and by Raper and Braithwaite (2009). So if you are going to resurrect this tired idea, then you will need to be back it up with solid arguments.

This statement betrays confusion about empirically measured volumes versus calculated volumes. We can *measure* the volume of an object by its displacement of water in a beaker, by using gravimetric techniques, by using GPR, or by a dozen other methods. In a completely separate measurement, I can pull out a ruler and measure the surface area of the object. These independent measurements have absolutely nothing to do with each other, and no calculation has been used to derive the volume from the measured area. Now let's suppose we want to find out if there is a relationship between the measured volume and the measured area. We can plot them both and do an appropriate regression. The regression will be significant if there is a relationship. That's basic statistics.

Alternatively, we can *calculate* the volume of the object by multiplying the area by the mean thickness. But then we would obviously have a tautological relationship between the two variables. There would be no point in doing a regression because we already know the relationship a priori. A regression in this case would have less noise, but that is entirely irrelevant; the decreased noise is because the volume data was manufactured from the area data, not measured.

When the authors talk about an autocorrelation, they are claiming that the volume has been *calculated* from the product of the area and the mean thickness. Well, if that is the case, then as noted above, it would be ludicrous to even attempt a regression. It would not be a test of the relationship between glacier volume and glacier area because the measurements would have been manufactured and would not be independent.

If the authors want to claim that there is an autocorrelation between volume and area (because volume has been calculated from area), then they don't have actual measurements of volume, and they only have calculated/manufactured volumes. In that case, they don't have appropriate data! And in that case they most certainly should NOT be claiming that their (less noisy) volume-area plots indicate a shortcoming of volume-area scaling.

To emphasize this point, here are two examples of what an erroneous "autocorrelation" statement would mean if it were applied to other disciplines.

(1) Hacks Law would be invalid. This law states that a river basin's area scales as a power law with the length of the river basin. Hack's Law has been studied, derived, proved, and routinely used by hydrologists for over half a century. Note however that river basin area equals basin length times basin width, so by this paper's logic, Hack's Law has a serious shortcoming and is an autocorrelation that artificially suppresses noise. Well, you will have a very, very hard time convincing hydrologists of this.

(2) The Stefan-Boltzmann law describes the energy flux radiated from a black body as a function of temperature. The energy flux scales as the fourth power of the temperature. But following the logic of this paper, this would mean the temperature is autocorrelated with the flux (because the flux is proportional to temperature times the temperature cubed). Again, you will have a very difficult time convincing physicists. They might even laugh. It is, after all, a fundamental law of statistical mechanics, and it has been beautifully validated with real world data.

In summary, when the authors claim that an autocorrelation is a conceptual shortcoming of volume-area scaling, this is tantamount to admitting they don't know the difference between an appropriate regression of real volume data and an inappropriate regression of calculated (manufactured) volumes. I really doubt that's the case, so please get rid of this wildly inaccurate claim.

Pg. 4830, line 6.

> *"(ii) The correlation of glacier area and thickness is rather weak (cf. Fig. 8), and modified by the surface slope of the glacier."*

What? Can you back that up with data? Can you back that up with a derivation from the physics? At the very least, can you provide a reference? You already discussed the correlation in point (i) on line 2, so that part of the sentence can be deleted. And the second part relating area-thickness scaling to slope is intriguing but baffling. Even if it is true (reference please) what does that have to do with volume-area scaling? Why is this (potential) aspect of thickness-area scaling a shortcoming for volume-area scaling? I know that there is a slope closure condition that is (sometimes) used to derive the volume-area scaling exponent, but that closure condition usually delineates glaciers from ice caps. If there is more to this, then please explain.

Pg. 4830, line 11.

"(iii)..., (b) the scaling parameters are determined on the basis of only a few hundred glaciers with measurements at most."

No, that's not entirely true. The scaling exponent can be derived from theory, and two of the three scaling analyses in this paper use that theoretically derived value of gamma = 1.375 (Bahr et al, 1997; Arendt et al, 2006). I will agree that (to date) the scaling parameter c has been determined from data. But I'm not sure that a few hundred measurements can be called insufficient. Plenty of valid regressions have been done with less data. Bahr (1997, Water Resources Research) goes a step further and derives c as the mean of a probability distribution. It's the same data, but it's a different technique for which a few hundred data points is usually considered more than sufficient.

I can entertain the notion that there might be biases in glacier sizes that could impact the calculation of the mean value of c, but you would need to say this and then back it up with evidence. The appendix of Bahr (1997) shows a reasonable distribution of c that would not seem to support a size bias, but no specific tests were done.

Perhaps a more reasonable claim would be that c could vary from region to region. If so, there is insufficient data to establish this regional variability. That might be a real shortcoming, and it might be worth mentioning in this paper.

Pg. 4830, lines 14-15. The references on line 14 and 15 are good. They support your point that the scaling parameters vary. But why isn't the same potential variability discussed with respect to the numerical models? Surely there could be regional, temporal, or other variations in f, tau, etc.?

Pg. 4830. For a balanced presentation, please include another paragraph in this section that discusses the "conceptual aspects" of numerical models. The "*variety of possible combinations for the [model] parameters*" is one aspect. The necessary simplifications from full Stokes models is another aspect (albeit possibly less important, in my opinion). I am sure there are others aspects worth mentioning. Discussing the conceptual difficulties of scaling without discussing the conceptual difficulties of numerical models does not give your paper the appearance of a balanced presentation.

Pg. 4830. Continuing with the conceptual errors in section 5.2... Figure 8 is used in support of the arguments in section 5.2, but why is "measured" in quotes in the caption for Figure 8 and again on page 4825, line 17? If this is not real thickness data, then is this plot even a fair representation of the thickness-area relationship? Or do you mean to imply that GPR derived thickness is not an actual measurement, but is instead a complicated derivation from measurements? Please explain this.

I suspect the data in this plot is volume data (compiled by Meier, Bahr, Cogley, Grinsted, etc.) divided by glacier area. I'm not sure because the authors don't say it explicitly. But if so, then this is a plot of V/A versus A. That is an autocorrelation, and it will artificially increase noise (because A is in the denominator). Well, that's ironic!

By the way, scaling theory requires that the thickness be *measured* in the same place on every glacier. For example, it can be the average thickness (measured, not calculated), the average thickness on the centerline, the average thickness across the equilibrium line, the single value of thickness measured at the intersection of the equilibrium line and the centerline, etc. But whichever is chosen, it needs to be the same quantity in the same place on every glacier. If the values are not consistently measured in the same place on each glacier, then the scaling relationship is not guaranteed; at best, the scaling relationship would be very noisy, and in a worst case scenario, the scaling relationship would not be identifiable at all.

Although I am surprised at how good Figure 8 looks considering that it spans only one order of magnitude (see above), this plot would undoubtedly be less noisy if the mean thickness was *measured* as the same quantity on every glacier. As the authors note, the thickness has rarely been measured in the Himalaya-Karakoram and is interpolated and extrapolated from measurements that were not made in the same place (as the same quantity) on every glacier. Therefore it is probably a mistake to over-interpret the scatter in Figure 8. At the very least, the scatter in this plot needs a much deeper analysis.

For the above reasons, among others, total glacier area and total glacier volume are very nice choices for data (rather than thickness). By definition, volume and area will be measured as the same quantity on every glacier. As the authors correctly note (pg. 4825), volume measurements are difficult and rife with problems, but at least they are a consistent measurement, and therefore the physics guarantees that the volume will scale with area.

I would suggest eliminating thickness-area scaling as a conceptual shortcoming of volume-area scaling. In fact, it's quite the opposite. As noted above, volume-area scaling is a conceptual solution to some of the problems posed by area-thickness scaling.

Pg. 4833, line 11. Have you considered using the numerical models to estimate appropriate scaling parameters gamma and c for volume-area scaling in the Himalayas? Would the performance of volume-area scaling be improved if more appropriate scaling parameters were used for the Himalaya? Either way, this would be a very interesting result, and would be a more thorough and in my view essential exploration of the two different approaches (scaling versus numerical modeling).

In fact, until you have appropriate volume-area scaling parameters, can you really claim that the modeling approach is superior? Yes! I think you can make that claim based on the idea that the modeling can be tuned to the specific region without (unavailable) volume data. But *if* volume-area scaling will give results that are more in line with the numerical modeling when using the correctly tuned parameters for c and gamma, then I think your claims should be much more nuanced.

I realize that this constitutes a major change in the manuscript, but it would very much strengthen your results, and your comparisons would be much more definitive.

Pg. 4833, line 20. This line implies that the modeled results are based on ice dynamics while the scaling approaches are not. Consider rewording or removing this. The scaling approach is based on a derivation from the underlying physics/mechanics/dynamics.

Pg. 4833, line 22. Yes, this is probably one of your strongest points. Scaling doesn't play so well with arbitrary (and inconsistently placed) thickness measurements (as discussed above), so the ability to validate numerical models with arbitrary thickness data is a big advantage. Although I give you a hard time about your misinterpretation of the area-thickness scaling plot, it really plays into your hands. It shows that volume-area approaches are harder to validate; they aren't necessarily wrong, but without the ability to easily compare scaling results to existing thickness (or other) data, it is harder to tell if and when volume-area scaling is wrong.

Pg. 4853, Figure 8: I assume this is calculated from volume data that was collected globally rather than locally in the Himalaya. The global nature of the data should be mentioned because the rest of the paper is dealing with the Himalaya-Karakoram. At first glance, it is also very slightly confusing to have a box that says “V-A scaling relations” on a plot that shows a thickness-area relationship. The dashed lines on the plot are not the specified volume-area relationships but are instead V/A.

Summary

(1) Please rewrite section 5.2 to remove the inaccurate claims, preferably by dropping the discussion of area-thickness scaling altogether. For a more balanced presentation, also add to this section a discussion of “conceptual aspects” of numerical modeling.

(2) Use your numerical models (or other techniques) to tune the scaling parameters c and gamma to the Himalaya. Until this is done, the scaling techniques are not being given a fair assessment. So far, you've only shown that global values and Alaskan values for c and gamma don't work so well in the Himalaya.

(3) Otherwise, well done. I agree with your general conclusion that the numerical models offer many advantages over scaling. A more definitive and nuanced comparison will enhance your conclusion considerably.

David Bahr
October 4, 2013

References:

Bahr, D.B. 1997. Global Distributions of Glacier Properties: A Stochastic Scaling Paradigm. *Water Resources Research*, 33(7), 1669-1679.

Clauset, A., Shalizi, C. R., Newman, M. E. J., Power-law distributions in empirical data, *SIAM Review*, 51(4), 661703, doi:10.1137/070710111, 2009.

Lüthi, M., M. Funk, and A. Bauder, Comment on “Integrated monitoring of mountain glaciers as key indicators of global climate change: the European Alps”, *J. Glaciol*., 54(184), 199–200, 2008.

Raper, S. C., and Braithwaite, R. J. Glacier volume response time and its links to climate and topography based on a conceptual model of glacier hypsometry, *The Cryosphere*, 3, 183-194, 2009.