Reply to review of

Grinsted, A.: An estimate of global glacier volume, The Cryosphere Discuss., 6, 3647-3666, doi:10.5194/tcd-6-3647-2012, 2012.

Aslak Grinsted says:

> *I answer the review comments point by point in indented italics below.*

Anonymous Referee #1

This paper is timely – new estimates of total GIC volume are very important to the ongoing efforts to predict sea level rise. The paper is easy to read and the analysis is described clearly. The results are unique and significantly different from previous studies of the same topic, and for this reason, the paper deserves careful attention. In many places, additional citations would be helpful. In particular, this manuscript's scaling exponents depart significantly from previous work; therefore, a thorough list of previous volume-area regressions would be helpful in the introduction. A summary table of regression values from previous studies (exponents and constants) would also help place this paper's work in context with others. The primary justification for the new exponents is an improved regression technique.

> *I have added a table with published scaling laws. I have also included a new figure which hopefully helps to explain both the methods and why I get different exponents to previous studies.*

I heartily endorse an improved regression of the data, but the choice and justification may need some additional thought. There is a great deal of literature describing appropriate ways to find power-law scaling exponents (e.g., Clauset et al, 2009, referenced below), and the author's selection is not discussed in that context. Again, citations would be helpful.

> *In order to justify the particular methods I use, I have now included a test to investigate the performance of different calibration methods when tested against data from a virtual world (I use the Huss and Farinotti 2012 data as the truth in this virtual world). This gives much stronger support for the regression technique that I finally employ. This has also allowed me to improve the regression technique further (by not having k as a free parameter). I believe this test provides much stronger support for the approach that I use.*
>
> *One of the main arguments of the paper is that when you calibrate an empirical relationship, then you must consider the purpose for which you plan to apply it. You should optimize the skill with respect to performance measures that are considered most important for the task at hand. This is why I optimize it with respect to absolute volume deviation. My goal is not to*

have the smallest percentage error on any particular volume estimate, but to design a relationship that results in the smallest error in the total volume.

Clauset et al. is an excellent paper, but while some of the considerations are useful, then it is only directly applicable when the power law you are fitting is describing a distribution. In my case I have volume on the y-axis rather than a frequency or a probability density. The dangers of log-log regression that Clauset et al. list in their appendix A do not apply in this case. Clauset do not describe cross validation (but endorses it). This is the approach I use in my model selection. Another important difference between Clauset and my study is that he is talking about real power laws, whereas I am talking about empirical fits. Although I use the term "scaling law", then I don't actually think of it as a law, but rather as a useful approximation which works well up to certain point. -Probably because it is pretty close to a law as derived by Bahr.

The conclusions of this paper depend on the quality of the volume and area data for the 210 glaciers and 34 ice caps shown in Figure 2. That list is short enough to warrant publication of the data as a table, and would help others to both evaluate the quality of the data and to attempt their own regressions. I would consider publication of this data essential to the publication.

Graham Cogley has collected this data and I believe that he is still adding to it. So, for that reason I would much prefer that potential new users went to Graham Cogley for the data rather than using the particular set of data available to me at this time. However, I will add the contents of this file as supplementary material ….

I am concerned that there is such a large discrepancy between this paper's very high estimate of the percentage total volume of glaciers greater than 100 km^2 when compared to Bahr and Radic (2012, The Cryosphere). Normally, I wouldn't be too concerned with such a discrepancy, except that in this case the author is already proposing significantly different scaling parameters.

While my ~85% is greater than your calculation below based on Bahr and Radic (2012), then it is very compatible with the 86% estimated by Huss and Farinotti (2012). Also if you read off from Bahr and Radic's final graph then you get 70-75% which is not that far away from the estimate we propose.

The difference this seems to simply be a feature of the area distribution in the RGI, rather than a consequence of my scaling laws. Indeed applying other scaling laws with larger exponents would place an even greater fraction of the total volume in the large ice bodies.

The Bahr and Radic analysis can be applied to the scaling parameters derived in this paper, so the discrepancy may indicate a calculation bias in this paper that is over-weighting large glaciers (relative to small glaciers) by assigning them too much volume and/or by assigning small glaciers too little volume. This would be consistent, for example, with assigning too large of a value to the scaling constant.

I do not follow how you are able to conclude which study has a bias based on the disagreement alone. I do not dispute that my scaling fits might have a bias, but I do not think this affects the discussion I have here. My scaling laws have smaller exponents than most other published estimates. This would act to place a smaller volume fraction in the largest glaciers. So this cannot explain the difference.

However, the issue of RGI containing glacier complexes for some regions could lead to a positive bias in the largest ice volumes. I now estimate the bias in the total ice volume due to this glacier complex problem to be in the order of 15% of the total volume. This seems too small to explain the difference though. I believe that Huss & Farinotti's approach is able to give much better volume estimates for glacier complexes, than what I would get from glacier scaling, and they also get ~85%.

Regarding over-weighting large glaciers: I have now added a simple weighting to the least.abs.dev misfit function to reduce the effect of a sampling bias in the volume database. (As was suggested as a refinement in the old version of the manuscript). This has lead to larger exponents, but they are still below published estimates. The global volume estimate is unchanged.

Figure 3 does seem to show a significantly higher value for this constant when compared to previous regressions in other studies. The estimate of total SLE depends on the relative volumes assigned to the few large glaciers versus the many more numerous small glaciers. I have no way of evaluating this discrepancy quantitatively in the context of this work, but it is a concern. At the very least, I'd recommend noting (and possibly discussing) the discrepancy somewhere in this manuscript.

When I use a smaller exponent then I would have to compensate with a greater constant to maintain the fit. When this is taken into account then my scaling relationships are actually below most other published studies. (see the new figure 2).

Detailed comments:

(2) Page 3648, Line 18: This line mentions that the volume-area power law is "empirical but physically reasonable." I would instead say that the power law is "both empirical and physically derived."

Thank you. Revised to "physically justified".

(2) Page 3650, Line 26: The symbol n is often used for Glen's flow law. The Greek symbol gamma is frequently used for the volume-area scaling exponent. For clarity and consistency with previous

publications, gamma would be preferable. To avoid my own confusion in this review, I'll use the traditional n for the flow law exponent and gamma for the volume-area scaling exponent.

I have changed to gamma.

(3) Page 3651, Line 3: This line mentions that the theoretical power law exponent of gamma = 1.375 is derived from perfectly plastic ice. Not so. The scaling exponent is derived for arbitrary choices of Glen's flow law exponent n (see equation 7 in Bahr et al, 1997). The specific exponents of gamma = 1.375 and gamma = 1.25 are derived using n = 3.

Thank you. This misrepresentation has been fixed. However I note that Bahr et al. (1997) assume a 'square-root' shaped ice cap profile when getting 1.25 which is essentially boils down to assuming perfect plasticity.

(4) Page 3651, Line 4: This line mentions that volume-area scaling has been derived in Bahr et al (1997) for linear glaciers and circular ice caps only. In fact, there is no restriction on the geometry.

Thank you. The misrepresentation has been fixed. Bahr's scaling considerations are very general. The particular exponents of 1.375 and 1.25 are consistent with particular idealized geometries and perfect plasticity, although this is not actually assumed in the closure choices.

The quantities used in Bahr et al (1997) are "characteristic values" of the length, width, etc. These characteristic values summarize in a single number the complex geometries of real-world glaciers. This is exactly the same as the characteristic values of time, length, thickness and velocity that are used to summarize the complex shapes of glaciers in Johannesson et al (1989, J. Glac.) when evaluating the response time of glaciers. This is also the same way that non-dimensionalizations of other problems in physics and engineering handle complex geometries.

(5) Page 3651, Lines 8-11: The author implies that there may be inadequacies with the theoretically-derived volume-area scaling relationship because it assumes a fixed scaling constant k. Actually, neither Bahr et al (1997, JGR) nor Bahr (1997, WRR, vol 33) make that assumption. In fact Bahr (1997, WRR) shows that k has a distribution of possible values (with a well-defined mean) – see Figure 6a. The author should cite more explicitly those publications which perhaps erroneously make the assumption that k is a constant, and then note that the derivation from the physics does not make this assumption.

I do not mean to imply that there is anything wrong with the theoretical foundations. I now clarify that Bahr 1997 considers a distribution.

I also do not think that it is wrong using a constant k in the scaling law, as long as you realize this is an approximation that will not work well for every glacier.

(6) Page 3651, lines 12-14: Yes, it is good that this is stated explicitly. Scaling laws are best applied to large sets of glaciers and not to individual glaciers. This is not widely acknowledged but should be, and I appreciate seeing that here.

(7) Page 3651, line 16: It would be best to cite specific references that use regressions to derive gamma and k. In fact, this manuscript derives radically different values for the scaling exponent, so a review of previous regressions would be most helpful. It's also worth noting that not all papers use a regression to derive a value for k. For example, consider the way Bahr (1997, WRR) treats the scaling constant. In that paper, the constant is shown (from data) to have a distribution of possible values with a well-defined mean (with what looks like a roughly normal distribution); that derivation assumes that the scaling exponent gamma is fixed. Other papers (e.g., Slangen and van de Wal, 2011; Radic et al, 2008) have also looked at the effect of variations and errors in k to assess potential problems associated with deriving its value from a regression.

I have added a new table which shows a list of earlier published relationships. Several of the papers I list in this table 1 use log-log regression to estimate the parameters (although sometimes regressing against average thickness rather than volume). Bahr 1997 is an exception.

I have updated the misfit function to put less weight on large glaciers (to partially account for sampling bias in the calibration data set). This has rendered the exponents less 'radical' (but still smaller than many published studies). The impact this has had on total volume and volume distribution is minimal however.

(8) Page 3651, lines 22-23: I agree that errors in the largest glacier volumes are more important when calculating the sea level equivalent (SLE). However, the errors in the larger glaciers are mitigated somewhat by the fact that there are very, very many small glaciers. Therefore, a scaling error at small glacier sizes is multiplied by tens of thousands when calculating the SLE. This tends to level the playing field somewhat between the large and (numerous) small glaciers. When calculating the SLE of glaciers and ice caps, it's not quite as biased toward large glaciers as the author implies. Bahr and Radic (2012) discuss the role of many small glaciers.

The small glaciers are clearly also very important, but they receive an undue share of the weight when you do log-log regression. Consider the 50% of the volume held by small area glaciers, and the 50% of the area in the large glaciers. These two groups should receive equal weight. However, there are hundreds to thousands of times as many in the bottom half compared to the top half. (The exact number depends on whether I look at the volume database, or the volumes I estimated for the RGI or WGI/GLIMS).

That said… The least abs. dev. estimator probably gave too much weight to the largest glaciers/ice caps, because there appears to be a sampling bias in the volume database. There seems to have been a preference to measuring the largest glacier units. I have now introduced a weight in the misfit function to reduce the effect of this sampling bias (as alluded to in the previous version).

(9) Page 4, lines 6-17: The author makes a very good point that least-squares regressions can be fragile. It's always good to check the physics, and this manuscript performs a more sophisticated regression where "volume misfit" is minimized. The author's goal here is a good one. The proposed regression technique is resistant to outliers in the data, and a citation of this important and valuable feature of the regression should be added. However, techniques of "least absolute deviation" have other shortcomings that can cause unstable regression slopes when there are small errors in the glacier area (the errors do not need to be large). In other words, the "least absolute deviation" technique could potentially derive non-unique and unstable solutions to the volume-area exponent. In particular, the regression slope (exponent) could change dramatically with small changes in glacier area data. Under normal circumstances I would not be too concerned with this possibility, but this manuscript does derive radically different scaling exponents when compared to most previous studies. As part of a possible explanation, the potential for this instability (and non-unique solution) needs to be examined very carefully in this manuscript. Frankly, I doubt that there is any instability, but it would be much better to explicitly rule it out. Other possible power-law regression techniques are outlined in many references. For example, see Clauset, A., Shalizi, C. R., and Newman, M. E. J.: Power-law distributions in empirical data, SIAM Review, 51, 661–703, doi:10.1137/070710111, 2009. This reference uses the similar maximum likelihood estimator, and they discuss the slope instability problem in great detail. Their approach ad conclusions could be very useful in this paper.

I have added two references about least absolute estimators. I also evaluate the performance of the various regression techniques in two ways: I do cross validation, and I now also test the skill in on a surrogate dataset where I know the truth (I use the Huss and Farinotti 2012 set og areas vs volume). If there is any instability then the model will not have good cross validation statistics and will be weeded out by this procedure.

(10) Page 3652, line 17: The Akaike Information Criteria is not explained and needs a reference.

I have changed this model selection criterion to a simple mean squared prediction error of log V against withheld data. I do not need a penalty for high-order models (as in AIC) because I am doing out of sample validation. Whatever model does best in an out of sample validation should be considered the best regardless of the number of parameters. This simplification of the model selection criterion has only minimal impact on the results, and was found to be more robust in tests (less chance of over-fitting).

Note this change in model selection criteria does not influence the total volume much. The new simpler model selection criteria allows for more complex statistical models even if they only give slight improvement in the cross validations.

(11) Page 3653, line 26: The correct units are 3-2*gamma, not 3 – 2k.

Thank you. Typo fixed.

(12) Page 3654: line 16-17: The author says the total volume is concentrated in the largest ice masses. By his wording here (and elsewhere in the paper), he is implying that the largest of the very largest GIC contain the bulk of the mass. But Bahr and Radic (2012, The Cryosphere) show that about 50% to 60% of the total mass is in all of the remaining glaciers (depending where the cutoff is assigned). This is acknowledged later in the text. So this statement on lines 16-17 is not quite wrong, but it's also not quite precise. A little more precision is warranted in this statement.

I have revised the sentences. At this point in the paper I do not think that the specific numbers are very important. I simply want to highlight that a very large fraction of the total volume is concentrated in a few very large ice bodies. If you do log-log fitting then you will have extremely many points in the small size classes, and the fit will be constrained primarily by them.

Later in the paper (in the final figure) I quantify the volume in the largest glaciers and I get ~85% in the largest glaciers using RGIv2 (size threshold =100 km2). This is higher than Bahr and Radic 2012, but very compatible with Huss and Farinotti 2012 who gets ~86%.

(13) Page 3654, line 18: The author acknowledges that the Randolph inventory contains ice masses that combine multiple glaciers in a single outline. But then he says this will lead to a 20% error and does not actually separate the ice masses. This 20% is a guestimate (evaluated from only one Devon Ice Cap example) and should not be treated as +/- error because it is a systematic overestimate. Not separating the contiguous ice masses into separate glaciers will always lead to an overestimate of volume. However, the amount of the overestimate can vary considerably. For example if the ice outline is separated into two glaciers that contain 10% and 90% of the original area, then the error is only 9% (when gamma=1.375). If separated into two glaciers of equal size (50% and 50% of the total area), then the error is 33%. All possible errors between 0% and 100% are possible. The choice of 20% needs additional justification; or perhaps a better alternative to drop this arbitrary error value. The author can then note that that "not subdividing ice masses" will always give an upper bound on the SLE. This would be consistent with the author's argument that the total GIC SLE should be lower.

I have included a new more solid estimate of the standard error and the potential for bias. The bias arising from not sub-dividing glacier complex, can be huge (+80%). However, on the global total I estimate the RMSE to be 20% and bias to be +15%. (I suppose I was lucky with my earlier guestimate.)

There may however also be another negative bias as illustrated by the Devon ice cap example. The ice cap volumes will be too low if they generally have been subdivided into many small units (as was the case for Devon).

(14) Page 3656, line 2: This paper finds the total volume of all the glaciers in the inventories, but does not find the total volume of GIC on the Earth. It's a small semantic difference, but an important one. Many glaciers that are smaller than 1 to 2 km^2 are not in the inventories (depending on the region).

This has been clarified in the conclusions.

(15) Page 3656, lines 15-17: Using Figure 3 and gamma = 1.13 for glaciers and gamma = 1.19 for ice caps, the author estimates that 85% of the total volume is in glaciers larger than 100km^2. Bahr and Radic (2012, The Cryosphere) give a technique for determining the total volume in all glaciers larger than a specified size. If we use gamma = 1.13 in Bahr and Radic's 2012 analysis, then we find that only 58% of the mass should be in glaciers larger than 100 km^2. If we use gamma = 1.19 in Bahr and Radic's 2012 analysis, then we find that only 44% of the mass should be in glaciers larger than 100 km^2. Reality is a mix of glaciers and ice caps, so the correct value should be between roughly 44% and 58%. That's a huge departure from this manuscript's result of 85%. The departure is difficult to explain, but the difference does not depend on the scaling constant; the constant k is not part of the derivation in Bahr and Radic (2012). Also, Bahr and Radic's 2012 analysis does not rely on the underlying physics and only takes scaling laws as "fact." In other words, the scaling laws can be either empirical or theoretical, and the volume-area scaling exponent can be assigned any value; the result is not dependent on using the theoretical value.

This disagreement (<58% versus 85%) is possibly indicative of a flaw in the manuscript's analysis. One possibility is that the scaling constant k in this paper is too large for glaciers. For a given and fixed scaling exponent (e.g., gamma = 1.13) a large value of k would tend to exaggerate the volume of all the largest glaciers relative to all the smallest glaciers. Interestingly, one of the drawbacks of a "least absolute deviation" regression (as used in this manuscript) is that the solution is not guaranteed to be unique. Therefore, it is unlikely but theoretically possible that the "least absolute deviation" regression could give the incorrect value of k (even while giving the correct slope gamma). Another possibility is that the contiguous ice masses which have not been subdivided are significantly biasing the total volume of large glaciers.

My estimate of 85% is supported by Huss and Farinotti, and is not sensitive to excluding regions with glacier complexes. I do not think this can be explained by the scaling relationships I apply, but the source of the disagreement must be in the area distributions in RGI compared to the power law area distributions in Bahr and Radic (2012). Further homogenization of RGI would certainly help here.

When I read off Bahr and Radic's final graph, then I get 70-75% of the volume in the glaciers larger than 100 km^2. I really do not think that is much of a disagreement considering the assumptions that go into Bahr and Radic's number. E.g. A single scaling exponent, a specified maximum area, and not a perfect fit to very large glaciers. Additionally my study may have a positive bias. Bahr and Radic's approach is a great way to estimate the order of magnitude *error you make from excluding very small glaciers. - But I am less sure that you can use it to get very precise estimates of volume fractions below a threshold. Can it really distinguish between 70% or 85%?*

(k affects all glaciers equally. A larger value of k would have to be compensated by a smaller value of gamma (otherwise it the fit is lost). A smaller value of gamma would place a smaller fraction in the large glaciers. So k cannot explain why I get a large fraction.)

(16) Page 3661, line 4: “Ice sheets” should be changed to “ice caps.”

Fixed.

(17) Page 3665, Figure 2: Please include another table that shows the specific glaciers, volume, area, etc. that were used to construct this figure and the regression. It will help readers evaluate the results.

This will be included as supplementary material.

(18) Page 3665, Figure 2, Line 6: The figure caption notes that the theoretical exponent is too high compared to nature. That may be true, but the author should instead say that “Theoretical scaling law exponents are generally higher than what is derived from the regressions in this analysis.”

Sentence rephrased.