Biogeosciences Discuss., 12, C4156–C4164, 2015
www.biogeosciences-discuss.net/12/C4156/2015/

Biogeosciences
Discussions
Open Access

Interactive comment on "Detecting methane ebullition on thermokarst lake ice using high resolution optical aerial imagery" *by* P. R. Lindgren et al.

D.R MUELLER (Referee)

derek_mueller@carleton.ca

Received and published: 10 August 2015

This paper explores the use of optical aerial imagery in estimating ebullition rates of gas from a thermokarst lake. This was done by examining imagery of a snow free ice surface – at a snapshot in time – and relating bubble presence to several types of seep and extrapolating to ebullition rates using empirical relationships. The method seems to work reasonably well and has potential. The paper is well written and is a good contribution to the literature. Scientific Significance, Scientific Quality, and Presentation Quality are all good to excellent. I recommend that it be published with minor revisions.

I have four general/main critiques that I develop immediately below. My other sugges-

tions are captured in the specific comments section. The vast majority of these are minor, some are suggestions, a few are questions. A limited number call for a more systematic rearrangement of the text or additional information.

Derek Mueller,

Carleton University

General comments:

I am vaguely aware that some of you wrote about detecting methane ebullition from lakes using SAR. I haven't read these papers but I see this method as being potentially complementary to optical imagery. You bring up this method in the Introduction and I kept looking for some kind of discussion of the relative merits of optical aerial image analysis and the SAR techniques examined by Engram, Walter and others (cited at the top of pg 7454). I think there is room to add this into the paper in the Discussion at least and this small touch would augment the value of this study.

Frankly, I kept wondering how you can detect bubbles in ice with optical imagery until I realized somewhere in the middle of the Introduction that you were looking at snow free imagery. Unless you want to know about open hotspots, this to me is a constraint (well relative to SAR, I imagine)... I think you should take care to make this clear early on in the manuscript and in the Abstract.

On pg 7452 and in several other places, you take care to mention that bubbling is highly episodic and seep bubbling rates are not constant over time. This makes me wonder about the suitability of this method in which you capture 2 to 4 days worth of ebullition and extrapolate this. Further, you check on this 'snapshot' with ground truth data taken after several weeks when the ice is safe to walk on. So a great many more bubbles ought to be present then. You acknowledge that the process is sporadic but don't really assess whether the method is able to capture the process sufficiently to allow for whole lake ebullition estimates. I would really like to see some further thought given to

this. At what time scales does the bubbling vary for each type of seep? What are the implications? Having said all that, all estimates of the ebullition rates seem to be fairly similar to each other (within the stated error/uncertainty of other estimates in any case). How did you arrive at these uncertainty values? How well did you do estimating the number of seeps/density of seeps for each seep type relative to the ground survey? How does this translate into how close your estimates for whole lake ebullition are? Are your aerial surveys close to the ground surveys by chance (over estimating certain types of seeps while underestimating others)? The results are in Table 1 but I think an analysis of this and the uncertainty with respect to the issues I raised above could be improved.

My last comment is an organizational one. You have 3 types of seeps initially, plus hotspots and then you add one (tiny) but don't treat it in the same way as the others. However, the hotspots should really be split into two classes based on the status of ice over top of them. I think you could improve your communication and be more organized if you thought about how and when to describe each seep type. More details are in my comments below.

Specific comments:

Title: Detection is one thing but I think you are doing that and more. Maybe revisit the title? I would add snow-free as a modifier of lake ice first off.

7451 ln 25 Can you quantify dominant? What percentage of methane is released via ebullition vs other means like diffusion? Also what percentage of methane in the bubbles in this lake (or typically)?

7452 ln 2-4 What are the implications of fast vs slow ice growth in this context? Does this alter bubble morphology and impact your bubble estimates?

7454 ln 9 Geological methane seepage could be confused with Hotspots. These are both methane sources, so that may or may not be an issue here. However, can you

comment on how upwelling from springs would be interpreted? Surely this would be confused as a Hotspot using this technique.

7454 ln 22 based on field-based (repetitive)

7454 ln 23 – please explain why you didn't conduct fieldwork immediately after the aerial survey. (later you explain, but it would be best to comment on this right away).

7455 – It would be nice to have a location map so we can see where the lake is relative to Fairbanks or other landmarks.

7456 ln 6 please give the number of GCPs and comment on how they were distributed across the study area.

7456 ln 9 – again, why wait so long?

7456 ln 15 – The tiny seep type really seems like an add on. Either mention this in the Introduction and follow through with it as with the other seep types or bring it up in the Discussion as a new type that was 'discovered'/characterized after this study was planned (this is what I assume happened). The Discussion is a great place to offer some insights and ponder the implications of this seep type. I don't see the middle of the Methods as an appropriate place to give background on this new seep type and start incorporating it into the manuscript.

7457 – ln 4 delete 'are'

7457 ln 16 elevations are given in m a.s.l. but you don't mention anywhere what the lake elevation is. It would be more appropriate to have the elevation in m above ground level anyway. Please consider providing this.

7457 ln17 and 1 : 17 000 for 2011 and 2012, respectively reads better.

7458 ln 2 a second order polynomial?

7459 the segmentation and classification settings/parameters are not provided. I

haven't worked with eCognition but in my experience there are a myriad of settings to tweak in doing both these procedures and any of these can affect the result. What were these settings? Were the same ones used in both years? How sensitive are your results to changes in these settings? How do we know you optimized this step? Please provide details on this.

7459 why did you segment and classify the land cover as well as the lake? Would it not be simpler to mask out the land first and study only the lake area?

7460 I assume that this step is, technically-speaking, an unsupervised classification but perhaps not? Can you please clarify? What are we to make of the 98% accuracy? This seems to reflect how well your customization of the eCognition routines worked on the image(s) you worked with. Since you didn't use an independent dataset to validate the classification routine, I don't see this as a true validation accuracy. That's ok, but please make this clear.

7460 as above, please explain a bit more on how you selected the threshold (what was the threshold ultimately?)

7460 ln17 It is fairly confusing to the casual reader that dark pixels are bubbles and light pixels are ice in the PC1 variable. I would suggest that you arrange for this to be the reverse, which is more intuitive. You could insert a step here 'the value of PC1 was multiplied by -1 [and offset by ##?] to reverse this band, making bubbles patches . . . ' I don't believe this will influence any subsequent results, but please check. Alternatively, you could just display the PC1 results in reverse and explain this in the figure caption.

7460 ln 22-3 We applied a post-hoc Tukey's Honest Significant Difference test . . . identify significantly distinct seep types. Also elsewhere you have 'Honestly' as well.

7461 ln 4 explain more on the MLC classification. What is the number of training samples? Were any samples taken for validation (at a different end of the lake)? What thresholding options and other parameters did you use?

7461 ln 8 integrating size as and additional feature – again, details are lacking. is this a step after the MLC? How did you do this?

7461 ln 12 seep types - add s

7461 ln 14 We studied - remove 'further'

7461 ln 21 Finally we evaluated. . .

7461 ln 27 – I would suggest that you not capitalize Complete Spatial Randomness and make this into an acronym. Spatial randomness is an easy concept to understand and by making it into an acronym, I was mistakenly thinking this was a procedure you applied. It slowed me down in this section. As well, you need to highlight in the part of your explanation that your null hypothesis is ultimately that the difference in spatial patterns are random, not that the spatial patterns themselves are random. This is important as your seeps predominate in certain parts of the lake.

7463 ln 5 – I would start this section a little more gently than getting into PCA results and correlations. How many seeps of each type were there? Did this change from year to year?

7463 ln 6 neg correlation – can you give more detail on this? It seems like there might be more to say. . .

7463 ln 7 – were assumptions for the ANOVA met? I see the ANOVA in supplementary material. Make it clear that hotspots are ice covered.

7463 ln 10 and 11 – writing significantly, $p < 0.05$ and 95% confidence interval is redundant. Chose one of these.

7463 ln11 to 14 – can we have more descriptive results on PC1? What is the range of bubble values in PC1? What is the range of lake ice (congelation and snow ice), snow, open water. . . lily pads, etc.

7463 ln 13 less bright ==> darker

7463 ln 18 given th spatial resolution

7463 ln 27 – it might be worth considering two subclasses of Hotspot for practical/logistical purposes throughout your paper – Ice covered Hotspots and open Hotspots (use your own terms, by all means). This distinction, once made clear early on, could then help you out here and elsewhere when you need to separate these two manifestations of hotspots. For instance your ANOVA is for ice-covered Hotspots only and seeking open hotspots is the sole purpose of the snow covered lake imagery, etc.

7465 ln 1-4 You made a big, late-breaking introduction of the fifth class of seep in the Methods and now we find that it cannot really be distinguished. I would suggest that you consider the place that Tiny-type seeps occupy in this paper (see my comment above). Depending on your decision here, you might consider adding Tiny type to a boxplot.

7465 ln 6-7 awkward sentence

7465 ln 10 – it would be nice to have a table for the users and producers accuracy

7465 ln 11 – mostly arising from

7465 ln 19 – aha, here we find out what took you so long to get out there ;-) Good reason, of course, but please write this in the Methods. I am sure we can all agree it is best not to have a delay between the imaging and the ground truth. Also, if seeps are sporadic, you are not integrating over much time (only 4 days). Longer would be better. How can this be mitigated – why not discuss this? I realize it could snow any time so you are probably wanting to get data right away after freeze up. . .. what about repeat surveys every 2nd day until the ice is safe to walk on? That way you can use the last possible date for study. One last question, when you did finally get onto the lake, was it snow-free?

7466 ln 3 could you say reductions instead of changes?

7466 ln 20 field-based = when exactly?

7469 ln 3 – can you just mention this was due to a different atm pressure, just to be clear?

7471 – here is a good place for a longer discussion on SAR vs optical for your application

7471 ln 22-27 – these are crucial points/limitations and should be also mentioned in your Abstract and Conclusion

7472 ln 1 allow to map – awkward/unclear

7472 ln 4 – bubble patch brightness, with some confusion.

7472 ln 5 – remove 'understandable'

7472 ln 10 – results also imply the potential, given the caveats raised above, to apply . . .

7473 – this section reads like a summary, can you bring your findings and their significance together and draw some more conclusions instead of recapping the Results?

7473 ln 12 – can you correct for, or account for, the difference in pressure? Or at least propose a way to do this in the Discussion? Seems like it might be possible.

7481 In the table, can you put two lines for the ground surveys – one for 2011 and another for 2012? Or if they only happened once, put the year.

7482 Is there a better way to lay out this table? It seems busy and isn't very intuitive. . . I don't have a suggestion right now, sorry.

7483 Nice to have the map of Alaska but it would be nice to have a zoomed in study site map of the lake and surrounding area. You can use this to display ground truth sites and other info that is missing from your current figures.

7484 I think lily pads should be brought up elsewhere in the manuscript – for example, how can these be distinguished from other classes in PC1 or otherwise?

7485 with p values < 0.05 as determined by Tukey's. . ..

7487 Fig 5a. Can you write out these empirical relationships? What about the p value of these relationships and the RMSE of your prediction? Also, I am having trouble understanding how the x values you have on the abscissa can produce the curves you drew in an exponential relationship!?

I wonder if it is possible to relate to ebullition rate as opposed to the area of the patches? I acknowledge it is a stretch but you might consider doing this perhaps in addition to the area (on a separate y axis?).

7487 Fig 5b. Please remove the perspective view of the lake. This adds nothing and makes it more confusing to understand. Can you explain why the N shore has changed? Is this infill?

7488 I think you should calculate your potential geolocation error and declare it here and were you look for spatial randomness in the text as well. Also, is it possible to look at whether the seep type changes year over year with your methods? (ie., not only the position. . .). Perhaps you can comment on that in the Discussion as this might be an interesting thing to do someday.

7490 Remove the legend from the graph. It is not required if you put this info in the caption. Explain the significance of where the black line crosses the shaded area, does that mean the process is random at that length scale? What does being on the other side of the curve mean?

Interactive comment on Biogeosciences Discuss., 12, 7449, 2015.