Response to reviewer # 1, D. R. MUELLER

P. R. Lindgren et al.

Comment on “Detecting methane ebullition on thermokarst lake ice using high resolution optical aerial imagery”

We are very grateful to the reviewer for giving the time to read the manuscript in detail and for providing very valuable feedback. We have revised our manuscript based on the reviewer’s comments. Below is our response to all the comments.

<u>General comments</u>

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.

This is a very good point. Thank you for the suggestion. We have now included the relative merits of optical aerial image analysis and the SAR techniques in section 5 ‘Benefits and challenges of aerial image analysis for ebullition seep mapping’. We have highlighted the limitations of SAR particularly on spatial resolution to be able to detect ebullition bubbles on small lakes and along thermokarst shoreline where ebullition is very active. High-resolution optical data can provide more realistic distribution of bubbles on the entire lake under snow free conditions and can also identify open hole Hotspots on snow covered images.

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.

We have now specified the use of snow free imagery to detect bubbles in ice with optical imagery in the Abstract and the Introduction.

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.

We agree that since many more bubbles are present in lake ice by one to three weeks following freeze-up [time of ground surveys] compared to two to four days following freeze-up [time of aerial photos], there is the potential for the aerial survey method to underestimate the total number of seeps on the lake. That said, other factors come into play, such as hydrostatic pressure affects on bubbling rate during the 2-4 day aerial survey time frame, which can lead to some seep misclassification in the aerial photos. These factors are evaluated in Section 4.1, 4.2, 4.3 and 4.5 of the revised manuscript.

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. At what time scales does the bubbling vary for each type of seep? What are the implications?

Walter Anthony et al. (2010; Limnology and Oceanography Methods) provided analysis of the temporal heterogeneity of ebullition for each seep class. Large fluctuations in the daily bubbling rate are seen in the long-term bubble trap data sets for each seep class. Usually these fluctuations are related to hydrostatic pressure changes. The implications of temporally fluctuating fluxes for the seep classes are discussed in Section 4.1 of the revised manuscript.

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.

Standard error of mean (SEM) of each flux type estimated by Walter Anthony and Anthony (2013; JGR Biogeosciences) is the measure of uncertainty. We have stated this in the revised text. We have discussed the comparison of seep density for each seep type and total flux derived from aerial surveys relative to ground survey in section 4.3.

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.

We agree that introducing the Tiny-seep class in the Methods section is confusing to the readers since very little research has been done on this seep class. Therefore, we removed Tiny-type seep from the early part of the paper and only mentioned it in the results and discussion.

To improve clarity, we have now mentioned, where appropriate, whether Hotspots were ice-free (open-hole Hotspots) or ice-covered at the time of our surveys; however based on the following two observations we did not subdivide the Hotspot class: (1) due to temporal variability in bubbling within the Hotspot class, it is possible for an ice free Hotspot to be covered by a thin layer of ice on one day an not another, and other hotspots to follow the opposite pattern depending on the recent history of their bubbling. (2) We observe ice-free cavities in spring-harvested blocks of ice from Hotspots that were originally identified as open and closed on early winter surveys.

Specific comments:

1. **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.**

   We agree and appreciate your feedback. We have changed the title to "Detection and Spatio-Temporal Analysis of Methane Ebullition on Thermokarst Lake Ice Using High Resolution Optical Aerial Imagery". A small portion of our study also shows the potential application of snow-covered images to identify open-hole Hotspots. Therefore, we have decided not to add 'snow-free' in the title.

2. **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)?**

   Greene et al. (2014; Biogeosciences Discuss.) showed that ebullition comprised 83% of total annual emissions from Goldstream Lake. The concentration of methane in Goldstream Lake's bubbles was 82-89% (Greene et al. 2014; Biogeosciences Discuss.). We have added this information in section 2 of the revised manuscript.

3. **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?**

   It does not have major implications in our study because we look at the very first ~ 5 to 7 cm of ice in the winter. Subsequently, winter ice growth and ebullition rate determine bubble morphology but it does not affect the estimates of our bubble surveys, as both the field data and the imagery are from very early in the season.

4. **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.

Geologic methane seepage and "ecological" Hotspots have in common the trait of ice-free holes in winter lake ice. They differ distinctly in associated fluxes (i.e. geologic seeps are several orders of magnitude higher flux than Hotspots) and spatial distribution. Hotspots are typically found along thermokarst margins of yedoma-type lakes and within a yedoma permafrost region they occur in nearly all lakes. In contrast, geologic ebullition seeps are not limited to a specific permafrost type, but can occur in any type of water body. In regions where they occur, typically those underlain by hydrocarbon reservoirs (coal, petroleum, sedimentary basins), geologic seeps nonethless are found in only a small fraction of the water bodies. Geologic seeps are far more rare than ecological Hotspots. Also, because they have higher fluxes, the open-water holes in lake ice are usually much larger than those of ecological Hotspots. These distinguishing factors should be taken into account in analysis of open holes in remote sensing images of lake ice.

It should also be noted that ebullition associated with groundwater springs typically has a very low CH_4% (<5% CH_4) and the geologic ebullition seeps described by Walter Anthony et al. (2012; Nature Geoscience) did not include seeps associated with groundwater springs.

5. **7454 ln 22 based on field-based (repetitive)**

   We have modified the sentence ‘based on with ebullition collected data collected on the lake….’.

   Important note: This paragraph in P 7457 now serves as a summary of our methods in the Methods section. We made this change to make it easy for the readers to get an overview of our multi-steps method before getting into details of each step.

6. **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).**

   We have now provided reason why we didn’t conduct fieldwork immediately in this section.

   Important note: This paragraph in P 7457 now serves as a summary of our methods in the Methods section. We made this change to make it easy for the readers to get an overview of our multi-steps method before getting into details of each step.

7. **7455 – It would be nice to have a location map so we can see where the lake is**

relative to Fairbanks or other landmarks.

The lake is located in Fairbanks. We have corrected that in the text. We have provided latitude and longitude location of the lake in section 2 and Figure 1.

8. **7456 ln 6 please give the number of GCPs and comment on how they were distributed across the study area.**

   We have now added the details related to GCPs.

9. **7456 ln 9 – again, why wait so long?**

   We have explained why it took long for us to do our field survey.

10. **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.**

    Yes, the presence of Tiny-type seeps and their role in altering brightness values of other seeps in the patch was discovered later in the study. We have removed Tiny-type seep from the Introduction and Method. They are only discussed in the Results and Discussion section (section 4.1).

11. **17457 – ln 4 delete 'are'**

    'are' is deleted.

12. **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.**

    We have added the lake elevation.

13. **7457 ln17 and 1: 17 000 for 2011 and 2012, respectively reads better.**

    We have made the change.

14. **7458 ln 2 a second order polynomial?**

    Thank you for noticing this. Yes, it is a second order polynomial. We corrected the sentence.

15. **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.**

We appreciate your comments. These are some crucial technical questions. Yes, there are number of settings in eCognition that can affect the result (please see reply #17 below). The procedures we used are same in both years except the threshold values applied in some of these ruleset settings are different. This is just because image quality varies depending on the prevailing conditions during image acquisitions. Therefore, threshold used in one image band to identify some object may not be the same between years. We have explained the object hierarchy and general rule applied to map bubble patches in both years in the paper (section 3.3.2 and supplementary section in the new version of manuscript). Please see some examples on how we set rules in reply #17.

This paper is focused on the analysis of ebullition dynamics using information derived from optical images. Therefore, we have only provided a general overview of the mapping technique. We realized that having technical details as well as the analysis section in one paper was going to be way too much information for readers. Many readers may find this complicated. A technical paper on methane ebullition bubble mapping using Object Based Image Analysis (OBIA), which provides all the details, is in preparation for peer-reviewed publication. We have cited the paper in prep. in this section of the manuscript to guide the readers if they are interested to learn more about the technical details of the methods.

Please note that we also decided to move some portions of methods in section 3.3 in the revised manuscript to the supplementary section, as our method section was long and difficult to follow for other readers. We have now provided details on Principal Component Analysis (PCA), object-based segmentation and classification, and classification of bubble patches as supplemental text. This has helped to keep the method description simple and straightforward in the main paper. To improve the clarity, we also sub-divided sections of methods in the revised manuscript. Please see the replies, #24 and #25, below.

16. **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?**

In the OBIA technique (or in eCognition) segmentation is always the first step. So, even to mask out the land first, the image has to be segmented and then classified to separate land from lake to study only the lake (which is the first stage in our classification). Also, we used winter images for classification; hence it is less straightforward to have a land-lake (water) classification as first step. One of the options could have been manually mask out the land and only work on the lake part. But one of our goals was to develop a semi-automatic ruleset in eCognition to create transferable and efficient steps to study multiple lakes. And this current method is capable of separating land from frozen lakes with few manual inputs. The manual input would be to assign a correct threshold on the setting if needed.

17. **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.**

    Yes, technically speaking traditional supervised classification requires training dataset and we do not use that, so in that sense this could be an unsupervised classification but still it is not a completely unsupervised classification. We make rules by defining thresholds on variables such as image bands, object's relation to neighbors etc. to classify image objects. For example, we set a threshold on PC 2 brightness component to separate lily pads from lake ice. We use a threshold on PC 1 component to separate dark and white ice on the lake. We use morphological filter to identify bubble edges for which we use PC 1 component. We also define rules based on object's neighbor's relationship. These are just a few examples. There are many different rules and the classification is narrowed down to the lowest level to identify bubble patches. Therefore, it is correct that the accuracy reflects how well our customization of the eCognition routines worked on the image.

    Due to thin lake ice condition on image acquisition days, it was not feasible to collect ground truth data of ebullition bubble patches for the purpose of accuracy assessment. We manually classified bubble patches on the image to use them as validation sites. We did not check our classification accuracy in terms of object's geometry or boundary delineation. We only performed quantitative site-specific accuracy assessment using error matrix that only checks the agreement of object classes between manually classified reference sample and object-based classification results. We have provided more details on accuracy assessment in the revised manuscript as a supplemental material to keep the paper concise.

18. **7460 as above, please explain a bit more on how you selected the threshold (what was the threshold ultimately?)**

    The threshold was a DN value on image bands, an input for contrast split algorithm, that separated open-hole Hotspots from the snow-covered lake. We have now provided this information. We have provided reference for contrast and split algorithm in eCognition for more information.

19. **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.**

    Thank you for your suggestion. We agree that the way PC 1 variable is related to the bubble brightness in the paper is not very intuitive. We have now reversed the PC 1 values (i.e. dark pixels/low DN values in PC 1 component is less bright bubbles and bright pixels/high DN values is brighter bubbles). Yes, this does not influence any subsequent results. We have explained inverted PC 1 in section 3.4.1 and more detail is provided in the supplemental text.

20. **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.**

    We changed the sentence to "We applied a post-hoc Tukey's Honestly Significant Difference (HSD) test, if the null hypothesis was rejected, to identify significantly distinct seeps."

21. **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?**

    We have added more details in section 3.4.2 regarding the number samples for training MLC and validating the MLC results. Most of the samples are concentrated along the thermokarst margin of the lake. But we do have other randomly distributed samples from all other parts of the lake where bubble patches are visible.

    The MLC classification only uses spectral characteristics derived from all visible bands (RGB) and PC 1 component of these bands of the training samples we provided. We have stated this in section 3.4.2.

22. **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?**

    Yes, there is another step after the MLC. The MLC categorized bubble patches solely based on pixel spectral characteristics i.e. only included brightness parameter derived from the training samples. MLC was a pixel-based classification; therefore a segment of bubble patch had more than one class of seep assigned depending on the variation in brightness of the pixels within the patch. Since the size of bubble patches is also an additional important indicator of seep class and methane flux (Walter Anthony et al., 2010; Limnology and Oceanography Methods), in the next step we further investigated the size of each seep class in a bubble patch identified by MLC. Based on the size information and combination of seep classes in a patch, we then re-assigned bubble patches to a more accurate methane flux. We have added more details in the supplementary section of the revised manuscript (Supplementary Text S3 and Table S1 in the revised manuscript).

23. **7461 ln 12 seep types - add s**

    'ʻs' is added.

24. **7461 ln 14 We studied - remove 'further'**

    We removed 'further'. This is a separate subsection now in the revised manuscript (Section 3.4.3).

25. **7461 ln 21 Finally we evaluated. . .**

    We removed 'Finally'. This is a separate subsection now in the revised manuscript (Section 3.4.4).

26. **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.**

    Complete spatial randomness is in lower cases and removed the acronym.

    We agree that the null hypothesis is confusing in the manuscript. We have revised the text to highlight that our null hypothesis is ultimately the difference in spatial patterns are random.

27. **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?**

    We have now described PC 1 brightness characteristics of each seep type at the beginning of section 4.1. We have also shown number of samples (number of seeps we used to compare field-based flux with aerial survey) and PC 1 brightness mean±standard deviation of each seep type in Figure 3 in the revised manuscript. Yes, the number of seeps of each type used for comparison changed. This is just because 2011 had less number of bubble patches mapped compared to 2012.

28. **7463 ln 6 neg correlation – can you give more detail on this? It seems like there might be more to say. . .**

    We have rearranged the sentences to make it clear.

29. **7463 ln 7 – were assumptions for the ANOVA met? I see the ANOVA in supplementary material. Make it clear that hotspots are ice covered.**

    We have now made it clear that Hotspots are ice covered.

30. **7463 ln 10 and 11 – writing significantly, $p < 0.05$ and 95% confidence interval is redundant. Chose one of these.**

    Thank you for noticing this. We are only using $p < 0.05$ in the revised manuscript.

31. **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.**

    More detail regarding the PC 1 values of different image objects is discussed in the technical paper (in prep.). The PC band value range of lake, snow, open water etc. varies between years. Please see reply #15 above. In the revised manuscript, we have reported the results on PC 1 values (mean and standard deviation) of each bubble seep type in section 4.1 and Figure 3.

32. **7463 ln 13 less bright ==> darker**

    We have replaced ‘less bright’ with ‘darker’.

33. **7463 ln 18 given th spatial resolution**

    We corrected this line.

34. **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.**

Thank you for your suggestion. We have made it clear what we are looking at on the images whether ice-covered or open-hole Hotspots in section 3. We also made it clear early in the Methods section 3.1 that ice-covered ebullition seeps (A, B, C and ice-covered Hotspots) are referred to as ebullition bubble patches and we are using snow covered lake imagery to map open-hole Hotspot locations. Please see our reply in general comments above regarding ice-covered vs. open-hole Hotspots.

35. **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.**

    We have removed the discussion of Tiny-type seep from the Methods and only mentioned it in the Discussion. Please see our replies related to Tiny-seep above in general comments.

36. **7465 ln 6-7 awkward sentence**

    We have changed the sentence structure to make it clear.

37. **7465 ln 10 – it would be nice to have a table for the users and producers accuracy**

    Accuracy table is now provided as a supplementary material.

38. **7465 ln 11 – mostly arising from**

    We replaced 'rising' with 'arising'.

39. **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?**

    We have now provided the reason for a gap between image acquisition and fieldwork early in the Methods. We agree that repeated survey is an option to mitigate this issue but may not be very cost efficient. Also, it is very rare that snow free condition will prevail until the ice is safe to walk on for ground surveys. We have stated this issue in section 5.

40. 7466 ln 3 could you say reductions instead of changes?

We added ‘dynamics’ after ‘ebullition’ in the current sentence: ‘It is well established that ebullition dynamics are inversely related to changes in barometric pressure (Mattson and Likens, 1990; Fechner-Levy and Hemond, 1996; Scandella et al., 2011).’ This is more accurate. A rise in pressure suppresses ebullition, while a drop in pressure enhances ebullition. So, pressure changes in both directions have differing affects on ebullition

41. 7466 ln 20 field-based = when exactly?

The results are from the observations over multiple. We have now made it clear.

42. 7469 ln 3 – can you just mention this was due to a different atm pressure, just to be clear?

We have now mentioned it.

43. 7471 – here is a good place for a longer discussion on SAR vs optical for your application.

We have now added the discussion on SAR vs. optical in section 5. Please see our reply in general comments.

44. 7471 ln 22-27 – these are crucial points/limitations and should be also mentioned in your Abstract and Conclusion

We have mentioned these limitations in the Abstract and Conclusion.

45. 7472 ln 1 allow to map – awkward/unclear

We corrected this line.

46. 7472 ln 4 – bubble patch brightness, with some confusion.

We replaced brightness with ‘PC 1 brightness values’.

47. 7472 ln 5 – remove ’understandable’

We removed ‘understandable’.

48. 7472 ln 10 – results also imply the potential, given the caveats raised above, to apply…

We corrected the line.

49. **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?**

    Thank you for your suggestion. We have shortened the Conclusion and only highlighted the major findings and their implications.

50. **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.**

    We haven't yet looked into this correction in detail. But we agree by using multi-temporal imagery (several years of data) this could be possible.

51. **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.**

    The ground flux estimate is based on multiple year data. We have stated that in the manuscript in section 4.2.

52. **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.**

    We have removed the table since the results are already explained in the text. We felt that the information is redundant and the table is confusing. Removing the table also helps to make the manuscript more concise

53. **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.**

    Please see reply # 7. We have displayed ground truth sites (surveyed polygons) on the figure. We have explained important characteristics of the surrounding area in the Study Area as well that we think can help readers to get an idea about the surrounding.

54. **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?**

    We used PC 2 band to distinguish lily pads from lake ice since vegetation had distinct characteristics (very low PC 2 values compared to lake ice). We briefly mentioned that we also used PC 2 bands to classify image objects in the paper (now in supplementary section of the revised manuscript). Please see reply # 15 and # 17 above. The details regarding techniques on identification of bubble patches are discussed in the technical paper (in prep.).

55. 7485 with p values < 0.05 as determined by Tukey's. . ..

'by Tukey's HSD test' is now added to the sentence.

56. 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!?

Thank you for the suggestion. We have now added regression equations, p-value and standard errors in the figure. We understand the confusion on the x-axis values and exponential relationship. We had the distances from the 1949 eastern lake margin to illustrate high methane bubble concentration along the thermokarst margin. We have replaced the graph with distances from the modern lake thermokarst margin (2012 thermokarst margin). The exponential relationship is between the distance from 2012 margin and the bubble patch percent coverage. We have marked where 1949 and 1978 thermokarst margins existed in relation to the modern margin in the graph.

57. 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?).

We appreciate your suggestion but for now we would like to keep the relationship as it is since the observations are made during the two very short windows of time 2–4 days after freeze-up. Relating distance from the thermokarst margin to bubble patch area is the most direct and accurate representation of our observations.

58. 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?

Thank you for your feedback. We have now removed the perspective view of the lake.

We suspect that the lake either partially drained or that the water level decreased sometime after 1949, though we have insufficient information to explain the cause for the observed change in lake shape along the north shore between 1949 and 1978.

59. 7488 I think you should calculate your potential geolocation error and declare it here and where 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.

We appreciate your comment. We understand that geolocation error is important. Since our geo-location error was not that significant (on average < 20 cm) in our analysis, we didn't mention it in the paper. Now we have declared it section 4.5 and Figure 6.

It is challenging to look at whether the seep type changes with only two-year dataset. However, with many years of data it could be possible to check how bubble patches change (by comparing brightness and size) to analyze possible seep type changes. We have highlighted this possibility in section 4.5.

60. **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?**

    The legend has been removed. As suggested we have provided the meaning and shape of the curves and other details in the caption.