Reviewer #2: Review of "Network Modeling of Arcitc Melt Ponds"  
  
General comment:  
  
The paper present an algorithmic technique to map photographic images of melt ponds onto a discrete conductance network with the intention of allowing in future studies the ease of lateral flow in summer melt pond covered sea ice. The algorithm consists of 4 steps: i) preprocessing to eliminate small ponds and small elements of floating ice; ii) identification of individual melt ponds; iii) in parallel finding connections between melt ponds; iv) and small node deletion.  
  
The paper focuses on the model description and applies the algorithm to a small subsample of remote sensing images at the expense of systematically analysing extensive datasets of summer melt ponds images. While this choice can be justified in principle it calls for an extremely clear presentation of the algorithm itself. As it stands I do not think that the authors achieve a high enough standard in presenting their algorithm. I believe that in its current state the paper does not allow the reader to reproduce the method without a substantial amount of additional bibliographical research. I recommend at the very list a clearer presentation possibly including a detailed schematic of the algorithm in Appendix or as a figure. I would also urge the authors to make their algorithm available to the scientific community, maybe as supplementary material.  
  
Nevertheless I consider the approach presented here worth publication and I recommend the paper to be accepted with major revisions.  
  
In addition to this main comment I list below a series of suggestions that I think need to be addressed for the paper to be accepted for publication.  
  
Below I describe in more details the places that I think could be improved.  
  
Abstract:  
  
- In the version I received there are 2 abstracts and they differ. Correct

The spacing in the abstract is different from the pdf (latex takes care of the spaces) and the 2nd from last sentence is a little different. I will fix this.

- You mention seal holes. Is this an important contribution to water drainage ?  
- You state in the last sentence that the number of mislabels is used to evaluate performance while I find this is not done in the paper.

For this, we can include another table with the percentage or number of missed connections and number of mislabeled connections and classify them by months (June, July, Aug). Is it possible to get Boya to help out with this?   
- You mention graph theory. I am not an expert but it seems to me that you could be more specific here.  
Maybe we can frame the sentence as “Using basic clustering and edge elimination, the conductance matrix is reduced….”

Alternately, we can say “Basic clustering and edge elimination using undirected graphs, …”

Introduction:  
  
- L54: cite also Hunke et al, 2013 ("Sea ice, Albedo, Melt ponds, Ridging, Modeling, Arctic"), Flocco et al, 2012 (10.1029/2012JC008195).  
- L68-L76: Is this really needed ?  
- L80-L146: I am not sure if all this is directly relevant for the present paper and therefore needed ? I understand that it helps set the context of percolation phenomena but here you look at horizontal processes. As such I would more interested to have a more in depth discussion of scaling aspects of your work. For example along the lines of your reference [15]. One question that will arise later in the paper is how the conductivity values that you determine scale with your sample size. I think this kind of discussion would be more useful than a general discussion on percolation phenomena.  
- L159: you mention algorithms that helped you in [15] distinguish between melt ponds and open water in leads between floes. I am not sure that you use the same methods in the current paper. I think you should. If you do not explain why.

Can we explain that here we can distinguish between them the same way that we can distinguish between ice and melt ponds – using the intensity of red color? I know that we are not doing this in the current images because we have selected pictures that do not have ocean water in them. I thought that ocean water will be a sink for all the melt pond water

Explained that possible with blue intensity but we do not select images that contain ocean water.

- L161-L171: I think a more in depth discussion on scaling of melt ponds geometrical characteristics would be welcome maybe after this paragraph.  
- Is the daily cycle relevant for your findings. For example if you measure pond characteristics at different times of the day would you expect different thermodynamic forcing and hence different conductivity maps…  
- L200: should you discuss melt freeze up, melt lid formation a bit more here ?  
- L205-210: I think you should provide some reference illustrating how transverse flow of water is important, maybe Scott et al, 2009 ("Modelling the evolution of Arctic melt ponds"). How do the timescales of drainage vs horizontal flow compare for example ?  
- L214: do these technique have a name ? Can you give some examples of applications in other fields ?

I can elaborate a little more on the techniques like geodesic opening, closing, dilation, erosion, etc and give examples of how they change an image.  
- L219-L221: Scott et al, 2009 tackled some of that.  
  
Method:  
  
- L241: Otsu's method, again here I am not sure if you do not mix up melt ponds and open water between floes sometimes. This could in turn affect your values of conductivities.

I don’t know how to tackle this. We remove any smaller pieces of floating ice using image processing, but large pieces of floating ice are indistinguishable in the images from solid, fixed ice ice. I’m assuming the reviewer is speaking about this and not pond vs ocean.

Have explained that Otsu’s threshold can be used again on the blue intensities to distinguish between water in melt ponds and ocean water.

- L247-255: not clear here if images are a the same resolution in HOTRAX and SHEBA and even for different days. You should provide the pixel to m conversion value or at least discuss this.  
We use SHEBA images for the June and July photographs and HOTRAX for August photographs. They definitely have different resolutions. I remember that Dr Golden mentioned that there was a ship which we could use to determine the pixel to physical size scaling. Is this still an option? Otherwise, can we say that we have just used the maximum bottleneck sizes and pond sizes based in terms of pixel values in the graphs, but if height from which the photographs are taken and the actual size of a feature are known, we can easily get the pixel to physical size values?

I explained as follows:

I this paper, we have calculated all sizes, for example the largest bottle-neck in the melt pond network, in terms of the number of pixels in an image. If the physical size of an object in the image is known, for example a ship, then the physical size to pixel ratio can be calculated and instead of using number of pixels as a size measure, the physical dimensions of melt pond network can be used.

Section 2.1:  
  
- L264-289: I find your description of erosion + dilation not very clear for someone not familiar with the method. You could offer a schematic with the shape of the mask. For example I do not understand what a circular 3x3 mask looks like ? Also discuss sensitivity of your results to size of mask. For example how are the results for conductivity impacted when your ignore narrow bottlenecks ? Should you ?  
 I can elaborate more on erosion and dilation and provide figures to show what these techniques do. Also provide a figure of the mask. Explain how if the mask is larger than the narrowest bottleneck, the bottleneck will be eroded away in the first step and not registered. Ignoring narrow bottlenecks will reduce the horizontal conductivity calculated, leading to inaccurate results.

Section 2.2:  
  
- L303-L304: reverse sentence order.

Will change sentence order.

- L306: maybe say how you obtain connected components (what method you use).

Elaborate on the connected components technique.

Include in appendix

- L323: should it read 'minimum' ?

No. The sentence is “This step is repeated until a pre-defined maximum bottleneck size is reached.” I will explain how we assume that after this step, the remaining ponds are considered to be individual ponds and not interconnected ponds, because we have already eroded the largest connection between ponds.

Depending on the season in which the photographs are taken and the resolution of the photographs we can find the expected largest bottlenecks in the melt pond network empirically by performing the above erosion steps repeatedly till all the connections between melt ponds are broken. This was done for a sample image in each image set in table \ref{table:list}. Knowing that a $3\times 3$ circular mask erodes two layers of pixels from the bottleneck - one from each side, we can calculate the number of erosion iterations that are needed to break the network into individual ponds. For example, if the widest bottleneck is 60 pixels across, 30 erosions are required. So the repeated erosion defined above are performed until this maximum bottleneck

size is reached.

After this maximum bottleneck size is reached in

the above connected components process, it is assumed that all the remaining melt ponds are individual melt-ponds and not networks of smaller melt ponds.

- L325-334: I find this paragraph very unclear and qualitative. The reader cannot reproduce your method exactly as you do not provide values and justification of the choices of this ratio.  
- On figure 2 you should specify the different ratios used.  
I will explain this more clearly. I will also mention the ratios used and justification for using them.

I have changed a lot in this section to better explain this concept – included new figure.

Consider the image in figure \ref{bottleneck.ratio}. Here, red lines show regions that are bottlenecks and should be eroded away eventually as they are connections between melt ponds. Green line shows a region that is slightly constricted, but cannot be considered as a bottleneck as it is large relative to the pond surrounding it. Simply performing erosions as described above would eventually break all of these connections. To prevent this, contriction ratio, $CR$, is defined as

\begin{equation\*}

CR = \frac{pond \; area}{bottleneck \; size}

\end{equation\*}

It was empirically found that a constriction ratio of $CR\_{min} = 20$ worked well with the images used in this paper. During any erosion step, if for a network under consideration, $CR \lt CR\_{min}$, then this is probably a melt pond and should not be broken down any further.

Section 2.3:  
  
- This is by far the most difficult part to reproduce. You must provide some kind of clearer summary of the algorithm.

Again, will have to elaborate on the method used in this section.

Have re-written large parts of this section

- L349-357: unclear. Figure 3 is not clear either and lacks explanations in the legend. Why is Fig  3b the negative of Fig 3a for example ? What are the brown lines in Fig 3b ? Define direct connections between ponds ? Define intermediate connections ? Vague.

I can explain how Fig 3b is just a binary version of 3a with the geodesic distances shown in a third color(brown) for clarity – I will explain these in the figures also. Will explain differences between direct connections and intermediate connections using the same figures.

The first figure on top-left is the input image used. The second figure on top-right shows geodesic distances between melt pond nodes, this figure is a binary version of the first figure - blue is ice, green is water and maroon shows the smallest geodesic paths between nodes. The third figure on bottom shows the final connections obtained after edge elimination.

- L358-359: add references for these 2 methods.

I can add text book references for dilation and clustering. I can also add some references for the graph theory method and explain that we have used a different equation to calculate the weights assigned to edges.

Added reference for dilation, have to add for graph theory

- L365: should this be eroded ?

No. “dilated” is correct – I can make the explanation clearer.

- L365-L373: this seems ad-hoc and is unclear. Explain more clearly with better schematics maybe.

Will explain better.

- L374- … Line numbers missing here!  but this method description is very poorly explained. At times description makes no logical sense: "the distance between unconnected ponds is considered to be an arbitrarily large number, which is larger that the maximum distance between two ponds". How can the distance between 2 ponds be larger than the maximum distance between 2 ponds ?! You must provide schematics and a much more detailed description of the "simple clustering approach" and the "graph theory" !

Again I will make this more understandable – we use both the distances between the ponds and the width of the connections to assign weights to eliminate the edges connecting the ponds. The distance here is the geodesic distance between them, i.e., distance between the ponds through water, not ice. If the ponds are unconnected, we just assign a really large number to the distance between them (the actual path does not exist as they are unconnected – so they are an infinite distance apart through water).

The distance between unconnected ponds

is set as infinity because the strength of connection between two ponds decreases with increasing distance and an infinite distance corresponds to absence of any connection between ponds.

I have already mentioned the following about the clustering approach - he center of each melt pond

pixel-cluster is located using the mean of the cluster

with Euclidean distances

I have given details about how the nodes and edges in the graph are determined and eliminated.

- You use sigma for conductance strength is analogy to work in material science I suspect. Please provide some references of similar work in other fields here.

I can, maybe, cite something from electrical engineering, where the conductance of a wire is directly proportional to the cross-sectional area and inversely proportional to the length of the wire?

The above equation is analogous to conductance in an electrical circuit, which is directly proportional to conductivity of the wire and inversely proportional to the length of the wire.

- Fig 4 mentions dilation in the legend nowhere to be seen on the figure.

I need to add an additional dilation step to the figure.

- L385: improve description of noodle deletion. Why ratio of 20 chosen, discuss sensitivity of results to that parameter.

I can explain the node deletion better. I can explain the ratio and how the result varies with its choice. I need to think a little more about why 20 was chosen.

The final step of the algorithm is for node deletion,

where the algorithm searches for very small nodes that

lie between two or more much larger nodes, and

eliminates these small nodes based on a predetermined

ratio. This is because, if a really small melt pond lies between two much larger melt ponds, it is probably just a part of the channel connecting the two large melt ponds and should not be labeled as an individual melt pond. For the results presented in this paper, this ratio is empirically set to $20$

- L390: replace latter by second.

Change as suggested.

- L398: dilation is not shown. It is not clear to me how pond 1 and 6 can become connected.  
I will explain this better – it will also be clearer once I change Fig 4 to have the additional dilation step.

Section 2.4:  
  
- L405: you introduce battery nodes in the framework of sea ice which sounds a bit odd. Please say in analogy to…

I’m not sure – similar to an electrical circuit?

Not sure about this:

This is analogous to an electrical circuit, where the conductivity between two points can be calculated and the flow of current through the circuit depends on the potential drop across battery nodes.

- Lines numbers missing between L420 and L425. Replace "Let the M" by "Let M".

Change as suggested.

Changed as suggested

- L424: explain equation (6) and provide reference.  
- L425-L430: very unclear. What are you trying to express ?  
  
Results:  
  
- L455-L457: unclear. If you look at smaller areas you get faster results. I do not understand.

I think I should remove this line – even if you get faster results for smaller images, you will need more iterations to cover the entire summed image.

Removed this line.

- L460: reduced resolution of the conductance values ? Why ? Unclear.

I should remove this line also – I wrote it and it is unclear to me too.

Removed this line.

- L462: using "graph methods". What are they ? Unclear ? It sounds like you use a package from Matlab and that you ignore its scientific content.

I don’t know if I need to elaborate on the methods used (assigning weights to edges and then edge elimination) every time I mention them. Maybe reference the section in which I explain this – “ Using graph methods as described in section …” ?

- L473-L476: unclear.

I can explain it better, maybe – it seems clear to me.

However, this

choice would be application specific, as even the

isolated ponds may be used to study the evolution

of networks with time, because they might, at some point further in time join larger interconnected networks.

- L477-L480: I do not see what kind of ground truth you are talking about here.

Explain the ground truths are images in which all the ponds and connections are already marked and verified and these would be used to evaluate our method, if they existed.

In image processing, ground truth refers to data from images that have already been processed and is known to be correct. Ground truths are often used to evaluate the performance of an algorithm as they provide a desired solution to the problem under consideration.

- Reorder figures so that Fig 10 (June) is first.

Change as suggested.

- It looks to me that the figures are at a different scale and I expect this to affect the conductivity value. Am I correct ? You cannot compare conductivity at different scales. The scales are important and must be clarified throughout the text.

This is an extremely relevant point, but I do not know how to address this.

- L486-L492: I do not understand.

This is just a description of what the image shows. I will explain it better.

For this reason, unlike the above mentioned figures, the images are shown without

removing the melt pond labels which are unconnected to the battery nodes.

- In figures 6 to 10 I do not understand why only a subsample of melt ponds is highlighted ?

I’m assuming that the reviewer is talking about some of the melt ponds not being labeled as nodes – this was done because they are not connected to the battery nodes and hence do not contribute to horizontal conductivity. I will include this explanation.

Note that in the images shown in figures \ref{July1}, \ref{July2}, \ref{Aug\_hot1} and \ref{Aug\_hot2}, the melt-ponds that are not part of the network which connects the battery nodes, have not been labeled to prevent excess clutter in the figures.

Conclusions:  
  
- The conclusion is far too weak. What are the implications of this work ? Why do we care…?