Network modeling of Arctic melt ponds

Meenakshi Barjatia, Tolga Tasdizena,b, Boya Songc, Christian Sampsonc, Kenneth M. Goldenc

a Electrical and Computer Engineering, University of Utah, USA
b Scientific Computing and Imaging Institute, University of Utah, USA
c Department of Mathematics, University of Utah, USA

ARTICLE INFO

Article history:
Received 27 February 2015
Received in revised form 14 August 2015
Accepted 28 November 2015
Available online 15 December 2015

Keywords:
Melt ponds
Horizontal conductivity
Mathematical morphology
Graph theory

ABSTRACT

The recent precipitous losses of summer Arctic sea ice have outpaced the projections of most climate models. A number of efforts to improve these models have focused in part on a more accurate accounting of sea ice albedo or reflectance. In late spring and summer, the albedo of the ice pack is determined primarily by melt ponds that form on the sea ice surface. The transition of pond configurations from isolated structures to interconnected networks is critical in allowing the lateral flow of melt water toward drainage features such as large brine channels, fractures, and seal holes, which can alter the albedo by removing the melt water. Moreover, highly connected ponds can influence the formation of fractures and leads during ice break-up. Here we develop algorithmic techniques for mapping photographic images of melt ponds onto discrete conductance networks which represent the geometry and connectedness of pond configurations. The effective conductivity of the networks is computed to approximate the ease of lateral flow. We implement an image processing algorithm with mathematical morphology operations to produce a conductance matrix representation of the melt ponds. Basic clustering and edge elimination, using undirected graphs, are then used to map the melt pond connections and reduce the conductance matrix to include only direct connections. The results for images taken during different times of the year are visually inspected and the number of mislabels is used to evaluate performance.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Sea ice is a critical component of Earth’s climate system and a sensitive indicator of climate change. The dramatic losses of summer Arctic sea ice observed in the past few decades have a substantial impact on Earth’s climate system, yet most global climate models have significantly underestimated the rate of decline (Boé et al., 2009; Serreze et al., 2007; Stroeve et al., 2007). One of the fundamental challenges of climate science is to develop more rigorous representations of sea ice in climate models and to incorporate important small scale processes and structures into these large scale models. For example, during the melt season the Arctic sea ice cover becomes a complex evolving mosaic of ice, melt ponds on the sea ice surface, and open water. While white snow and ice reflect most incident sunlight, melt ponds and the ocean absorb most of it. The overall reflectance or albedo of sea ice floes – the ratio of reflected to incident sunlight – is determined by the evolution of melt pond coverage and geometry (Perovich et al., 2002; Polashenski et al., 2012; Scott and Feltham, 2010). As melting increases, the albedo is lowered, which increases solar absorption, leading to more melting, and so on. This key mechanism is called ice–albedo feedback (Curry et al., 1995), and has played a significant role in the decline of the summer Arctic ice pack (Perovich et al., 2008; Pistone et al., 2014). Sea ice albedo is a significant source of uncertainty in climate projections and one of the most important parameters in climate modeling (Flocco et al., 2010; Pedersen et al., 2009; Polashenski et al., 2012; Scott and Feltham, 2010).

While melt ponds form a key component of the Arctic marine environment, comprehensive observations or theories of their formation, coverage, and evolution remain relatively sparse. Available observations of melt ponds show that their areal coverage is highly variable. This is particularly true for first year ice early in the melt season, with rates of change as high as 35% per day (Polashenski et al., 2012; Scharien and Yackel, 2005).

Such variability, as well as the influence of many competing factors controlling melt pond and ice floe evolution, makes the incorporation of realistic treatments of albedo into climate models quite challenging (Polashenski et al., 2012). Small and medium scale models of melt ponds which include some of these mechanisms have been developed (Flocco and Feltham, 2007; Scott and Feltham, 2010; Skyllingstad et al., 2009), and melt pond parameterizations are being incorporated into global climate models (Flocco et al., 2010; Flocco et al., 2012; Hunke and Lipscomb, 2010; Hunke et al., 2013; Pedersen et al., 2009). As melting progresses during the season, the evolution of melt ponds from small isolated structures into large interconnected networks is responsible for a number of processes that help control the rate at which the ice pack melts. It is believed (Hohenegger et al., 2012) that this evolution of connectedness is an example of a percolation transition.
(Christensen and Moloney, 2005; Stauffer and Aharony, 1992). Such a transition occurs when one phase in the microstructure of a composite material, for example, becomes connected on macroscopic scales as a controlling parameter exceeds a critical value called the percolation threshold (Broadbent and Hammersley, 1957; Christensen and Moloney, 2005; Stauffer and Aharony, 1992).

In the case of melt ponds the controlling parameter which gives rise to critical behavior is thought to be the fraction of the area of the sea ice surface covered by melt ponds.

An important example of critical behavior related to percolation theory as applied to sea ice, and important for melt pond drainage, comes from the study of fluid flow through the porous microstructure of sea ice. Specifically, the brine microstructure displays a percolation threshold at a critical brine volume fraction of around 5% in columnar sea ice (Golden et al., 1998; Golden et al., 2007; Pringle et al., 2009), which corresponds to a critical temperature \( T_c \approx -5 \) °C for a typical bulk salinity of 5 ppt. Below this threshold the brine phase of the sea ice consists primarily of isolated, disconnected pockets. It is only above the threshold where the brine phase becomes connected over large scales. This threshold acts as an on-off switch for fluid flow through sea ice, and is known as the rule of fives. It leads to critical behavior of fluid flow, where sea ice is effectively impermeable to fluid transport for brine volume fractions below 5% and increasingly permeable for volume fractions above 5%.

In addition to identifying the critical behavior of fluid transport in sea ice, the percolation theory of fluid and electrical transport through lattices (Christensen and Moloney, 2005; Stauffer and Aharony, 1992) was used to produce models of the fluid permeability of sea ice as a function of brine volume fraction (Golden et al., 2007). In this work X-ray computed tomography images of the brine microstructure of sea ice were analyzed and mapped onto random graphs of nodes and edges, in order to establish the percolative behavior of the system (Golden et al., 2007; Pringle et al., 2009), and the rule of fives in particular.

Other types of network models have also been used to describe both fluid and electrical transport in the brine phase of sea ice. For example, in the random pipe model, the diameters of random pipes, which represent brine channels in the ice, are chosen from lognormal probability distributions that describe the cross-sectional areas of the brine inclusions in sea ice and then assigned to the edges in a square lattice (Zhu et al., 2006). The fluid permeability of the model is then computed by using a random resistor network representation of the system and employing a fast multigrid method to find its effective conductivity which can then be related to the permeability. This same approach can also be used to directly model the electrical conductivity of the ice, an important parameter in remote sensing of sea ice thickness, fluid transport properties, and microstructural transitions (Addison, 1969; Buckley et al., 1986; Fujino and Suzuki, 1963; Ingham et al., 2008; Reid et al., 2006; Thyssen et al., 1974). Network models have been used extensively in analyzing the transport properties of composite materials (Milton, 2002; Torquato, 2002).

It has been suggested that percolative behavior occurs for melt ponds on the sea ice surface. As they cover more of the surface, disconnected, isolated ponds begin to evolve into large connected structures with complex boundaries, presumably achieving large scale connectivity above a critical area fraction (Hohenegger et al., 2012).

Increased connectivity of melt ponds promotes further melting through increased heat transport, contributes to the break-up of ice floes, and allows increased horizontal transport of meltwater toward drainage avenues such as large vertical brine channels, cracks, leads, and seal holes (Polashenski et al., 2012; Scharien and Yackel, 2005). Other melt pond models including both vertical and horizontal transport of melt water, such as a type of cellular automata, have been developed elsewhere, as in Scott and Feltham (2010).

In this work we begin to develop techniques for network modeling of melt ponds, their connectivity, and horizontal flow characteristics. Some of the groundwork for this type of modeling was laid in Hohenegger et al. (2012). Images of melting Arctic sea ice collected during two Arctic expeditions—the 2005 Healy-Oden Trans Arctic Expedition (HOTRAX) (Perovich et al., 2009) and the 1998 Surface Heat Budget of the Arctic Ocean (SHEBA) expedition (Perovich et al., 2002)—were analyzed for area–perimeter data on thousands of individual melt ponds. Algorithmic methods of distinguishing melt ponds from the ocean in leads between the sea ice floes were developed. This data was used to discover that pond fractal dimension transitions from 1 to 2 around a critical length scale of 100 m² in area (Hohenegger et al., 2012). Pond complexity was found to increase rapidly through the transition as smaller ponds coalesce to form large connected regions, reaching a maximum for ponds larger than about 1000 m² whose boundaries resemble space filling curves.

In earlier work on melt ponds and sea ice albedo, image processing has been used to measure the area fractions of melt ponds and leads from aerial and satellite images. In Perovich et al. (2002) these area fractions from June to October, using SHEBA images taken in 1998 (Perovich et al., 2002), show how the area fraction of melt ponds increases as summer progresses, and starts decreasing again at the end of summer as new ice forms. A probability distribution for the size of melt ponds is also derived from the data, which depends on the progress of the melt season.

In the work reported here, the connectivity of these melt pond networks is determined using aerial images of Arctic sea ice from the SHEBA and HOTRAX databases. We develop an algorithmic method of mapping a configuration of melt ponds onto a graph of nodes and edges. These melt pond configurations may be disconnected individual components, or partially or completely connected across an image. The edges are assigned values which indicate the width of “bottlenecks” separating larger pools of melt water, which are identified with the nodes of the graph.

The horizontal flow of water between melt ponds depends on the narrowest bottlenecks between them and the width of these bottlenecks is inversely proportional to the fluid conductance between them.

Mathematical morphology based image processing techniques (Gonzalez and Woods, 2008) are used with a clustering algorithm and graph theory to find a conductance graph associated with each melt pond configuration studied. Further work will explore the relationship of these graphs and associated conductance networks with the actual flow of fluid in the pond network, and the effect on sea ice albedo.

2. Method

The images of melt ponds from the SHEBA and HOTRAX expeditions are in color. The intensity and color of each pixel in the image are encoded using the intensities of the Red, Green and Blue colors that make up each pixel. The image is represented as a matrix of pixels, with each pixel being a vector of three variables — red, green and blue color values. These are called, respectively, the red, green and blue channels of the image.

These images are converted to gray-scale to reduce each pixel to only one intensity and lessen the number of computations required. The gray scale image is derived using the red channel as we see the largest difference between ice and water there.

A simple thresholding operation, as described in the Appendix A, is sufficient to segment the melt pond water from ice and produce a binary image. Otsu’s method (Gonzalez and Woods, 2008) is used to determine this threshold individually for each image, which is then segmented based on this threshold. Fig. 1 shows a histogram of the intensity levels of a gray-scale aerial image with Otsu’s threshold. After having segmented water from ice, it is also possible to use the blue color intensity in the images to distinguish between the ocean water leads and melt pond water. However, in this paper, we have selected images that do not contain any ocean water leads.
The images used are cropped from those in the SHEBA and HOTRAX databases, which have dimensions around 865 × 770 pixels. The size of the images does not affect the algorithm as long as the resolution remains the same. Only the processing time varies with image size. In this paper we calculate the sizes of ponds and bottlenecks only in terms of pixels. The number of pixels could be converted to a physical scale by knowing the helicopter altitude, camera characteristics, and so on, or if there were an object of known size, such as a ship, in the image. This information can be different for each image used, thus our focus here on pixel size. However, in Fig. 1 the horizontal scale of the image is about 80 m.

2.1. Preprocessing the image

The binary image produced by Otsu’s method can have small pieces of ice floating in the melt ponds, melt ponds that are too small to provide much information, and other small artifacts due to noise. These can clutter up the final connectivity graph with unnecessary data. Basic mathematical morphology operations involving erosion and dilation as described in Gonzalez and Woods (2008) and Appendix A are used to clean up the image.

A predetermined mask or structuring element of fixed size is centered at each pixel of the image and only those pixels, at which the structuring element fits inside the original image, are set to one. So, if a $3 \times 3$ structuring element is used, it will remove the outermost layer of pixels from the foreground, a $5 \times 5$ structuring element would remove two layers and so on. Morphological dilation is a complementary process where all those pixels, at which the intersection between the structuring element and the image is non-zero, are set as one. Dilation by a $3 \times 3$ structuring element would cause the foreground to grow another layer of pixels. Opening involves erosion followed by dilation with the same structuring element and is used to remove smaller structures from the foreground like protrusions and narrow connections. Closing on the other hand is dilation followed by erosion and it fills in small gaps in the foreground. Geodesic opening or closing involves finding the intersection of the result of opening or closing with the original image to preserve the shape of the image. The image is first cleaned up using geodesic opening of melt ponds to remove inconsequential melt ponds and geodesic closing to remove floating ice. Circular masks, as shown in Fig. 14, are used for these processes to maintain the curvy shapes of ponds. The mask size can be adjusted as desired. Here a $3 \times 3$ mask, like the first image in Fig. 14, is used. Note that care should be taken to ensure that the mask size is at least smaller than the narrowest bottleneck in the image, otherwise this connection will be lost.

2.2. Isolating melt ponds

The previous step results in large interconnected melt pond networks. The next step is to find individual melt ponds. First, connected components described in Appendix A are used to find all the separate unconnected melt pond networks and label each uniquely as $X_i$ where $1 \leq i \leq N_{\text{max}}$. Here, $N_{\text{max}}$ is the number of unconnected melt pond networks in the image.

Each of these networks is then eroded progressively with a $3 \times 3$ circular mask. Every erosion iteration scrapes away the outermost layer of pixels from the melt pond network image and the connections between the melt ponds gets narrower and some connections may break. In other words, at each erosion, some ponds might break away from the main network. These can be identified from an increase in the number of unconnected regions in the image, which are found using connected components. The $j^{\text{th}}$ region that breaks away from the network $X_i$ is labeled as $X_{ij}$. This can be an individual melt pond or a smaller network of melt ponds. The connection strength of the separated region $X_{ij}$ to its parent network $X_i$ is proportional to the number of erosion iterations after which it breaks away. Each of these smaller melt pond networks $X_{ij}$ is further eroded in a similar manner. The aim is to continue this until all the networks have broken down into their individual components, i.e., to separate out all the individual melt 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 until all the connections between melt ponds are broken. This was done for a sample image in each image set in Table 1. 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

![Image](https://example.com/image1.png)

**Fig. 1.** An aerial grayscale image of melt ponds from HOTRAX is shown on the left, with horizontal scale of about 80 m. A histogram of red channel intensities of the image is shown on the right. The bi-modal distribution is evident and Otsu’s threshold, marked on the histogram, can be used to separate melt ponds and ice.

![Image](https://example.com/image2.png)

**Fig. 2.** The red lines indicate constricted regions that should be marked as a boundary between two different melt ponds. The green line indicates a region that is simply a part of one large melt pond but might be treated as a divider between two different melt ponds because it is slightly constricted. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
individual ponds. For example, if the widest bottleneck is 60 pixels across, 30 erosions are required. So the repeated erosions 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 ponds and not networks of smaller melt ponds.
Consider the image in Fig. 2. Here, the red lines indicate bottleneck regions and should be eroded away eventually as they are connections between melt ponds. The green line indicates 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, we consider a so-called constriction ratio, \( CR \), defined as

\[
CR = \frac{\text{pond area}}{\text{bottleneck size}}.
\]

It was empirically found that a constriction ratio of \( CR_{\text{min}} = 20 \) worked well with the images used in this paper. During any erosion step, if a network under consideration has a constriction ratio \( CR \) such that \( CR < CR_{\text{min}} \), then this is probably a melt pond and should not be broken down any further. Another example of this is shown in Fig. 3. The image on the left shows the melt pond network that is obtained without using the constriction ratio. It can be seen that the ponds that are labeled 8 and 9 in this image are just a part of the long channel that connects ponds 7 and 10. The image on the right uses the constriction ratio and correctly labels the ponds.

2.3. Connections between melt ponds

The last part of the problem is finding the conductances between the individual melt ponds. As already described in the previous section, this is done while the interconnected melt ponds are being separated into smaller melt pond networks. To re-iterate, each erosion with a \( 3 \times 3 \) mask removes the outermost pixel layer. Thus two layers of pixels, one from each side of the bottleneck, are removed. If a region separates from the main network at the \( k^{th} \) erosion iteration, then the bottleneck joining this region to the network is \( 2 \times i \) pixels wide.

Until this point, the method has concentrated on grouping the melt ponds that are connected to each other and finding the sizes of bottlenecks between a melt pond and the network to which it belongs. The next step in the algorithm is to find out exactly which melt ponds are connected to each other and represent them using undirected graphs.

The problem also requires that we find only direct connections between ponds. To understand what this means, let each individual melt pond be a node in an undirected graph. Consider the 3rd image in Fig. 4 which labels each pond as a node. The node 56 is connected to all the nodes from 47 to 55 in this image. However, it is immediately or directly connected only to node 52. We can characterize the connection between nodes 56 and 48 by instead using the connection between modes 56 and 52 followed by connection between nodes 52 and 48.

At each erosion iteration, the interconnected melt pond network splits into a number of smaller ponds in the same erosion step. We have to find out which ponds are directly connected to each other as described above. Two simple methods of doing this would involve the following operations:

(i) morphological dilation (Gonzalez and Woods, 2008),
(ii) a simple clustering approach (Gonzalez and Woods, 2008) followed by a graph theory method (Van Steen, 2010).

In the first method, at each iteration, the eroded image is subtracted from the original image to get only the bottlenecks that were eroded away. This resulting image is then dilated and a simple overlapping operation (using the logical OR function) is performed to check which ponds form a direct connection with each other. This is illustrated in Fig. 5. A major problem with this approach is that sometimes the dilation is not sufficient to cause an overlap with the expected ponds and this leads to incorrect or missing connections.

In the second method, the center of each melt pond pixel-cluster is located using the mean of the cluster with Euclidean distances. One
may try to use k-means clustering on the initial image to separate the ponds, but as this only uses Euclidean distances between pixels and needs a fixed estimate of the number of clusters at the output, it will assign more than one cluster center to larger ponds and may ignore the smaller ponds. The geodesic distances between these cluster centers are calculated. The distance between unconnected ponds is set to infinity because the strength of the connection between two ponds decreases with increasing distance and an infinite distance corresponds to the absence of any connection between ponds. These distances are then used along with the conductance strengths calculated below to construct a graph of the melt pond network.

The nodes of the graph are the cluster centers found above, and all the nodes belonging to connected melt ponds are connected to each other with graph edges. Note that the conductance strength here only refers to the width of the channel connecting different ponds and gives a basis for relative comparison of ease of flow of fluid between these channels. Let the conductance strength between nodes \( i \) and \( j \) be denoted by \( c_{ij} \) and the geodesic distance between them be \( d_{ij} \). Each edge between two nodes \( i \) and \( j \) is assigned a weight \( w_{ij} \) given by,

\[
    w_{ij} = \frac{c_{ij}}{d_{ij}} \tag{1}
\]

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. Between each pair of connected nodes, the direct path and all paths involving only one intermediate connection are considered. For any node, there are \((n-1)\) possible paths to another node, or \((n-2)\) indirect
paths with one intermediate node and one direct path. The weight of the $k$th indirect path connecting two nodes is calculated as,

$$W_{ij}^{(k)} = \left( \frac{C_{ij}}{\sum_k C_{ij}} + \frac{C_{ik}}{\sum_k C_{ik}} \right). \forall k \neq i, j. \quad (2)$$

Here, $C_{ij}$ is the weight of the edge from node $i$ to node $k$. The weight of the edge which directly connects nodes $i$ and $j$ is $W_{ij} = \frac{C_{ij}}{\sum_k C_{ij}}$. Only the path corresponding to the maximum weight between two nodes is retained and all the edges corresponding to other paths are removed. This favors paths which are either very short or have large conductances. At each step, one pair of nodes in the graph is considered. For the next pair, the previously updated connection graph is used so that the edges that no longer exist are not reconsidered. 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 step is performed 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.

The second graph method performs much better for mapping connections than the dilation method. Fig. 6 shows the results obtained using the two different approaches. Consider nodes 5 and 6 at the bottom right corner in the first figure. The connection between the two nodes is not detected because dilation of the connection shown in Fig. 5 is not sufficient to overlap with ponds 5 and 6. Thus, pond 6 is shown connected directly to pond 1. This issue is solved in the second figure by using the clustering and graph method.

Fig. 8. Melt ponds in June from SHEBA early in the melt season. Here there are no complete connections that go across the entire image, so that the conductivity factors are zero.
2.4. Conductivity factor calculations

To calculate the horizontal fluid “conductivity”, first two battery nodes are added to the left and right of the image. 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 the battery nodes as well as the connectedness and local conductances of the graph representing the circuit. The left battery node is connected to all the ponds touching the left edge of the image with a conductance value of 1 for each connection. The right battery node is similarly connected. The purpose of the battery nodes is to simulate the computation of the effective or equivalent conductivity of a conductor network, which must be subjected to a potential difference, most easily visualized by connecting a battery. The effective conductivity of the network, between these battery nodes, is then measured. The conductivity of very large networks can be calculated approximately by considering smaller sections and then replacing these subsections with their equivalent conductivities in a hierarchical fashion similar to renormalization group techniques (Goldenfeld, 1992). The conductivity of each section could be calculated to create a new, simpler graph model.

To calculate the conductivity between battery nodes, and thus the effective conductivity of the graph with given bond conductivities, let $c_{ij}$ be the conductivity of the edge between nodes $i$ and $j$, and consider the formulation of the problem of finding the effective conductivity of a graph as found in Golden (1991). Here, each $c_{ij}$ is the normalized edge weight,

$$c_{ij} = \frac{w_{ij}}{\max_{ij}(w_{ij})}, \quad \forall i, j.$$  

Let $M$ be the total number of nodes in the graph, including the two battery nodes. We define the $M \times M$ matrix $A$ such that

$$A_{ij} = -c_{ij}, \quad i, j = 1...M, i \neq j,$$  

$$A_{ii} = \sum_{\forall j, j \neq i} c_{ij}, \quad i = 1...M.$$  

The matrix $A'$ is the $(M - 1) \times (M - 1)$ array obtained by removing the first row and column of $A$, which corresponds to the left battery node. Removing the last row and column of matrix $A'$, corresponding to right battery node, gives the $(M - 2)(M - 2)$ matrix $A''$. The conductivity factor of the image represented by matrix $A$, between the battery nodes, is given by Golden (1991) and Mason and Zimmermann (1960)

$$\sigma(A) = \frac{\det(A')}{\det(A'')}.$$  

It should be noted that the conductivity factor obtained is then related to the fluid permeability of the network, but not equal to the effective conductivity of the network, due to the length scale involved. As noted in the Introduction, further work will explore the relationship of this computed network conductivity to the horizontal fluid flow properties of melt pond configurations. Our goal here is to establish a viable method of transforming images of arrays of melt ponds and map them onto random, labeled graphs. The connectivity and local conductance characteristics of these graphs provide idealized, mathematical models of melt pond connectivity and effective, horizontal flow properties.
A brief summary of the method discussed in this section is presented in the form of a flowchart in Fig. 7.

3. Results

The above method is used to generate conductance graphs for different sets of images as described in Table 1. MATLAB is used to implement the method summarized above for each of these images.

This method was found to be most useful for images obtained in mid-summer, i.e. July, as the melt ponds are large and interconnected. The average time taken for different sets of images was calculated and is shown in Table 2. The SHEBA images taken in July were processed the quickest, because the images consist of larger and fewer melt ponds. Consequently, the operations involving connected components and the calculation of geodesic distances do not occupy the processor for too long. When these times are compared to the August melt pond images from SHEBA, which have many more melt ponds per image, the computations take much longer. Only about 10% of the computation time is spent in the calculation of geodesic distances and using graph methods to eliminate all but the direct connections between melt ponds. A major part of the computation time is spent in iteratively eroding the image, finding all the connected components and updating the bottleneck widths at each iteration. This can be sped up by using parallel processing for different connected components. Another step

![Fig. 10. Continued from Fig. 9, melt ponds in July from SHEBA. Here the white nodes are battery nodes. Conductivity factor values are calculated across these battery nodes and give an indication of the horizontal fluid conductance from the left to the right edge of the image. The white lines represent direct connections to the battery nodes while the blue lines are simply connections between ponds. For the conductivity to be nonzero there must be at least one full connection from left to right. Ponds which do not connect to battery nodes or which do not connect across the image are unlabeled to reduce clutter in the image. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).]
in reducing the time latency would be to ignore all ponds that have no other connections. 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. The conductivity factors shown in Figs. 9, 10, 11, and 12 respectively. Fig. 13 shows the conductance only missed when pond size is much larger than channel size. We most 1 in 10 of connections in our images are missed and typically are connection is very small or appears broken in the image. We manually count the number of mislabels and missed connections in each image. A mislabel occurs when a large channel is labeled as a pond or a large pond with a complex morphology may be labeled as 2 or more connected ponds. We find that less than 10% of labels are obvious mislabels. In terms of calculating the conductance, whether or not a large channel is considered a pond, or one large pond is considered two connected ponds, is less important. It is the connectedness that matters most for this calculation. Missed connections occur when the connection is very small or appears broken in the image. We find that at most 1 in 10 of connections in our images are missed and typically are only missed when pond size is much larger than channel size.

The processed images from July, August and June are shown in Figs. 9, 10, 11, and 12 respectively. Fig. 13 shows the conductance graph obtained for the 3rd image in Fig. 11. The conductivity factors for these figures are shown in Tables 4, 5 and 6. Note that in the images shown in Figs. 9, 10, 11 and 12, 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 (these melt ponds do not contribute to the horizontal conductivity calculations). The images shown in Fig. 8 do not have any complete connections that go across the image from left to right. For this reason, unlike the abovementioned figures, the images are shown without removing the melt pond labels which are unconnected to the battery nodes. The conductivity factor values for all these images are zero.

4. Conclusions

Melt ponds play a critical role in determining the albedo of the sea ice pack. Understanding their role in climate processes and incorporating their impact into climate models is fundamental challenges in climate science. In particular, quantifying key characteristics of melt pond geometry and connectivity are critical to quantifying and modeling melt pond growth, decay, and evolution. We have developed here a method of extracting the essential connectivity and scale characteristics of complex melt pond configurations and representing them in a discrete model. We have used image processing techniques in order to map melt ponds onto graphs whose edges represent horizontal flow pathways through the configuration. By computing the effective conductivity of these graphs, we obtain an idealized way of estimating the ease of horizontal flow of meltwater, which is important in melt pond evolution.

After visual inspection, it can be concluded that the algorithm we have developed does a very good job of identifying individual melt ponds, labeling their connections and creating the conductance matrix. More work can be done to improve its speed and remove the few mislabeling errors. The edge elimination method used assigns weights to the edges between nodes (melt pond centers) based on geodesic distance and widths of the connections. The function assigning weights to the edges can be modified and the weights of the nodes (areas of melt ponds) can also be used.
The work done here can be used to aid in understanding both the horizontal water and heat flow between melt ponds. These are important parameters to consider when modeling melt pond evolution and drainage which are the major controlling factors of ice albedo during the melt season. The conductivity factors calculated can be used to represent effective behavior of the ice–pond composite, and in turn this can be used to develop simpler models of the complex processes which govern melt pond evolution, ultimately to include them in climate models.
Acknowledgments

We gratefully acknowledge support from the Arctic and Global Prediction Program at the Office of Naval Research (ONR) through Grant N00014-13-10291. We are also grateful for the support from the Division of Mathematical Sciences and the Division of Polar Programs at the U.S. National Science Foundation (NSF) through Grants ARC-0934721, DMS-0940249, and DMS-1413454. We would like to thank the NSF Math Climate Research Network (MCRN) as well for their support of this work, and Don Perovich for providing the melt pond images.

Appendix A

Here we explain in more detail some of the concepts and techniques of image processing and analysis used in the body of the paper.

(i) **Thresholding:** Consider an image in which the value of each pixel is denoted by \( f(x, y) \) where \( x \) and \( y \) are the 2-dimensional coordinates of the pixel. If the image is bimodal, i.e., most pixel values fall in two major groups, then this image can be converted to a binary image by performing the following threshold operation:

\[
g(x, y) = \begin{cases} 
0 & \text{if } f(x, y) > T \\
1 & \text{if } f(x, y) \leq T
\end{cases}
\]

In the resulting image \( g(x, y) \), all unity-valued pixels are considered to be the foreground and the zero-valued pixels are the background.

(ii) **Mathematical morphology:** This uses set theory and is commonly applied in image processing solutions as it applies well to the analysis of geometric shapes and structures. For example, in a binary image, the set of all zero-valued pixels can be considered to represent the background, and the unity-valued pixels, the foreground. When applied in image processing, mathematical morphology usually employs structuring elements or masks which are used to perform various operations on the images of interest. Some examples of structuring elements are shown in Fig. 14.

![Structuring elements of different sizes](image_url)
Table 1  List of images considered.

<table>
<thead>
<tr>
<th>Set</th>
<th>Month</th>
<th>Number of images</th>
<th>Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>June</td>
<td>5</td>
<td>SHEBA</td>
</tr>
<tr>
<td>2</td>
<td>July</td>
<td>10</td>
<td>SHEBA</td>
</tr>
<tr>
<td>3</td>
<td>August</td>
<td>10</td>
<td>HOTRAX</td>
</tr>
</tbody>
</table>

Table 2  Average time to process each image.

<table>
<thead>
<tr>
<th>Set</th>
<th>Month</th>
<th>Database</th>
<th>Number of iterations</th>
<th>Average time (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>June</td>
<td>SHEBA</td>
<td>31.66</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>July</td>
<td>SHEBA</td>
<td>9.06</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>August</td>
<td>HOTRAX</td>
<td>18.04</td>
<td></td>
</tr>
</tbody>
</table>

Translation: A 2-D image can be represented by a 2-D integer space \( \mathbb{Z}^2 \). Each pixel in the image has a value and is associated with a fixed location \( z = (x, y) \). Translation of a set \( B \), by \( z = (z_1, z_2) \) is given by,

\[
(B)_z = \{c | c = b + z, \forall b \in B \}.
\]

Translation shifts every point in \( B \), \( (x, y) \), to \( (x + z_1, y + z_2) \) to result in \( (B)_z \).

Reflection: Reflection of a set \( B \) is given by \( \Bar{B} = \{ c | c = -b, \forall b \in B \} \).

Reflection is the mirror image of \( B \) such that every point in \( B \), \( (x, y) \), is reflected to \( (-x, -y) \), resulting in \( \Bar{B} \). If \( B \) is symmetric then \( B = \Bar{B} \).

Erosion: The erosion of \( A \) by \( B \) is defined as \( A \ominus B = \{ z | (B)_z \cap A \} \) where \( A, B \in \mathbb{Z}^2 \).

Erosion of \( A \) by \( B \) results in a set of points \( z \) so that all the elements of \( B \) translated by \( z \) fit completely inside \( A \). Erosion usually results in the removal of the outermost layers of the foreground.

Dilation: The dilation of \( A \) by \( B \) is defined as \( A \oplus B = \{ z | (B)_z \cap A \} \) where \( A, B \in \mathbb{Z}^2 \).

Dilation of \( A \) by \( B \) results in a set of points \( z \) so that at least one element of \( B \) translated by \( z \) overlaps with \( A \). Dilation results in the addition of layers to the foreground in an image.

Geodesic opening: Opening of \( A \) by \( B \) is erosion of \( A \) by \( B \) followed by dilation of the result by \( B \),

\[
A \circ B = (A \ominus B) \oplus B = A \ominus B
\]

This results in smoothing of the image by the removal of small protrusions and breakage of narrow connections.

Geodesic closing: Closing of \( A \) by \( B \) is dilation of \( A \) by \( B \) followed by erosion of the result by \( B \),

\[
A \cdot B = (A \oplus B) \ominus B.
\]

This results in smoothing of the image by filling in of small gaps and fusing of narrowly separated components.

Geodesic distance: The geodesic distance between two points in a binary image is the distance length of the path between the two points in pixels, such that the entire path lies in the same set as the two points. For example, in the foreground (unity-valued pixels), the distance between two pixels is measured along paths in which all pixels are unity.

Connected components: This is a technique for finding all the elements in a binary image that are connected to each other. Let \( A \) be the image in which we are trying to find connected components and \( B \) be a \( 3 \times 3 \) structuring element. \( X_k \) is an image with the same size as \( A \), but containing one unity-valued pixel at the same location as the component of interest in \( A \). This is called the seed. To find all the pixels in \( A \) connected to the seed, the following operation is performed recursively until \( X_k = X_{k-1} \).

\[
X_k = (X_{k-1} \oplus B) \cap A \text{ with } k = 1, 2, \ldots
\]

Table 3  List of parameters used for different image sets.

<table>
<thead>
<tr>
<th>Set</th>
<th>Constriction ratio</th>
<th>Number of erosion steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 4  Conductivities for image set 1.

<table>
<thead>
<tr>
<th>Image1</th>
<th>Image2</th>
<th>Image3</th>
<th>Image4</th>
<th>Image5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5  Conductivities for image set 2.

<table>
<thead>
<tr>
<th>Image1</th>
<th>Image2</th>
<th>Image3</th>
<th>Image4</th>
<th>Image5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6  Conductivities for image set 3.

<table>
<thead>
<tr>
<th>Image1</th>
<th>Image2</th>
<th>Image3</th>
<th>Image4</th>
<th>Image5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0542</td>
<td>0.1353</td>
<td>0.1216</td>
<td>0.0563</td>
</tr>
</tbody>
</table>

References


