FY 2015 Report: Developing Remote Sensing Capabilities for Meter-Scale Sea Ice Properties

Chris Polashenski
USACE-CRREL Building 4070 Fort Wainwright, AK 99703
phone: (570) 956-6990   fax: (907) 361-5188   email: christopher.m.polashenski@usace.army.mil

LONG-TERM GOALS

The overarching goal of this work is to develop and validate remote sensing techniques to track sea ice physical properties of geophysical importance that occur below the pixel size of most global-coverage satellite assets, particularly melt ponds.

OBJECTIVES

We will collect a dataset of high resolution satellite imagery and develop and field-validate methods for detecting melt pond area fraction, floe size distribution, and ice surface roughness from this imagery at a number of sites in the Arctic. The primary objective, in years 1 and 2, is to demonstrate the capability for operationally monitoring these variables. In the 3rd and 4th years of the project, these measurements will be scaled up to basin scale estimates, using both interpolation between observation sites and improved spectral mixing techniques to classify the fractional mixture of surface types within low resolution remote sensing imagery pixels, such as MODIS.

APPROACH

Key Particpants:

Chris Polashenski (Research Geophysicist, ERDC-CRREL)
Elias Deeb (Research Scientist, ERDC-CRREL)
Karen Frey (Associate Professor, Clark University)
Don Perovich (Research Geophysicist, ERDC-CRREL)
Blaine Morris (Research Scientist, ERDC-CRREL)
Nicholas Wright (Graduate Student, Thayer School of Engineering, Dartmouth College)
Carolyn Stwertka (Graduate Student, Thayer School of Engineering, Dartmouth College)
Alexandra Arntsen (Graduate Student, Thayer School of Engineering, Dartmouth College)
Nelson Crone (Graduate Student, Clark University)
Justin Chen (High School Student, Hanover High School)

Key Tasks:
1. Task and acquire high resolution panchromatic and multispectral optical (e.g. Quickbird, Worldview, National Assets) and synthetic aperture radar (TerraSAR-X, COSMO-SkyMed, or RADARSAT-2) images of regionally representative ~10x10km test areas (Yr 1–4)

2. Collect in situ measurements of target variables for validation of data products derived from high resolution satellite assets, using surface based observations and aerial photography at three readily accessible sites. (Yr 1–2)
3. Develop and validate automated pixel classification algorithms to extract pond coverage, floe size distribution, and ice surface roughness from high resolution satellite assets, and use these algorithms to create data products at sample sites distributed around the Arctic. (Yr 1-2)
4. Further develop and validate algorithms to derive melt pond coverage and floe size distribution from lower resolution assets through techniques such as spectral mixing based on in situ spectral measurements. (Yr 2–4)
5. Rapidly disseminate all data products created, along with suitable metadata, and publish methodology developed. (Yr 1–4)

WORK COMPLETED

During 2015 we continued to capture imagery; refined, tested, and iterated methods to segment and classify the imagery; identified and addressed issues with our version-1 mass processing scheme, improved upon processesing schemes in preparation for mass-processing imagery, and transitioned our entire scripting library to a single common, open source library. We faced significant challenges in not receiving our funding increment for FY15. Work effort on the project was significantly below expected effort due to this lack of resources. Using carried funds and other efforts to ensure continuity, we focused on continuing to collect the dataset and testing/refining processing methods using sample images in preparation for being able to restart the project in earnest in 2016. The largest impact of the funding shortfall was in removing several key participants from the project necessary for mass processing and up-scaling efforts. Nevertheless, our accomplishments indicate success in key tasks 1,3, and 4. We also took further steps toward conducting in situ validation of our methods (task 2) by leveraging our participation in the Norwegian Polar Institute N-ICE campaign. We were able to gather an aerial photo and ground-based dataset for validation while in the field. We recognize the need to enhance our distribution of our data and methodology (task 5) but lacked manpower to address this through publications in FY15. We did, however, migrate our processing methodology entirely into an open source environment and package it for distribution. Below we detail our work efforts as organized around these five key tasks.

Task 1 - Imagery Acquisition
We have worked with a wide variety of partners to task and acquire optical and radar imagery tracking both fixed sites and drifting sites throughout the Arctic. Based on utility, our primary focus narrowed to multispectral optical imagery during FY14. In FY15, we have contined to collect imagery with commercial optical assets that includes sample sites across most of the Arctic basin and spans a broad range of sea ice concentrations and melt pond coverage, both across space and time. Our imagery library is now over 5,000 images, and holds the data needed to begin assessing the spatial and temporal variation of key variables at many of the sites we are to tracking.

We continued to collect scenes needed to track the temporal evolution at both fixed locations and drifting sites, using the scheme developed in FY14 to increase the number of cloud-free scenes acquired, and circumvent the Digital Globe cloudiness filter, which appeared to fail in differentiating clouds and sea ice. Multispectral, high resolution imagery from Quickbird and Worldview 2+3 has proven to be by far the most suitable for automated identification and classification of sea ice surface conditions and again we focused our acquisitions on these products. This strategy has continued to be successful and we can now track the temporal evolution of key sea ice surface characteristics at fixed sites well.

In addition to the fixed sites, throughout the summer 2015 season, eight (8) CRREL ice mass balance buoys were targeted on an approximately bi-weekly basis using commercial high-resolution satellite imagery to achieve a lagrangian view of the ice. The tasking produced approximately 350 satisfied requirements for commercial satellite collections from 5/1/2014 through 10/30/2014. We have not fully cataloged the imagery from this year's collections, but look forward to incorporating it into our new Python database next year.

Task 3 – Develop Imagery Processing Techniques

We have iterated and tested numerous algorithms to segment and classify optical imagery, using a combination of both off the shelf software such as ENVI, Imagine, ArcGIS, and IDL, and implementing our own algorithms in Matlab and Python. We identified an algorithm that produces high accuracy segmentation and classification during FY14. This year we have invested effort into refining that methodology and migrating the coding from a labrynth of expensive commercial software packages to an entirely open-source implementation in Python. The result will enable code sharing as well as more computationally efficient mass processing.

The method we've settled on uses a watershed-based segmentation scheme seeded at local minima in pixel intensity gradient. The scheme is broadly similar to ENVI's feature extraction methodology, but now implemented in an open source Python environment with many more available tuning parameters to improve seeding and segmentation characteristics. An otsu's method-determined threshold is used to isolate unponded ice from all other classes prior to watershed classification to save computational effort. The remaining area is then segmented. Watersheds are grown from the seeds in an iterative basis in a gradient image (sobel filter) until they reach zero or negative gradient boundaries. Tuning the watershed segmentation to ensure that feature boundaries appear at all physically important locations requires breaking down the image into segmentation objects smaller than the typical surface feature (i.e. melt pond) size. Recombining adjacent cells of like type later in the processing allows us to make features whole again for statistical analysis.

After segmentation, statistics of each object, including its color and brightness, but also shape, tortuosity, texture, variability, etc are calculated. The expanded set of characteristic variables has greatly helped us differentiate between spectrally similar areas in multispectral imagery, providing many more differentiable variables than color intensity available to a single pixel. Classification of thes multi-variable objects into categories is carried out with machine learning. We implement this using the random forest classifier of the SciKit-image package for Python. Statistical analysis of the forest trees grown from these methods shows that differentiability of ponds and open water approaches the level possible by user evaluation of imagery (~95%).

Hundreds of segments in images selected from across the entire range of seasonal ice conditions are manually identified in a Python-scripted GUI by the user as a training set and from these we create random forest classifier trees. The random forest classifier is then run on numerous images. The more similar the images are to the training set, naturally the better the results, but results have proven very good across the entire summer season using only a single training set. An example image being taken through this workflow as implemented in commercial software is presented in Figure 3, and as implemented in python in Figure 4, showing we've successfully transitioned the entire workflow from 5+ expensive software packages to an open source environment. Typical classification success shows both user and producer accuracy above the 94%. Accuracy in this case is determined by randomly selecting several hundred segments, identifying these visually by human eye, and comparing the human classification to the machine classification.

Once we settled on a segmentation and classification scheme we recognized that a large portion of the remaining error comes from the hard-to-classify edge pixels which span the boundary between surface types. To reduce the area of the image that is covered by edge pixels, incorporated pan-

sharpening into our workflow to further enhance our ability to resolve fine scale features. Because of the strong role textural characteristics play in classification, we found that methods which better preserve texture rather than color are best. All pan-sharpening schemes are, however highly computationally intensive, so we are exploring processing schemes which take the grayscale image resolution, but do not fully calculate a pan-sharpened image. In FY15 we transitioned this portion of the workflow to open source Python implementation.

Additional refinement of our technique is still underway, targeting the remaining 4-6% error. Dark areas of open water pixels immediately between white ice floes tend to be 'contaminated' with glare from the adjacent bright ice areas, resulting in mis-classification of open water as ponds. This blurring comes from atmospheric scattering of light from the highly reflective ice and from interpolation errors introduced by pan sharpening. We are currently exploring methods to incorporate the original panchromatic band data (which would have atmospheric, but not interpolation error) to handle this. Similarly, small dark segments within melt ponds are common, even in ponds that are not yet melted through to the ocean. These segments, typically under 5 pixels in size, are frequently misclassfied as open water. Though an insignificant error in overall area classification, these small pixels do impact statistics of pond and water area size/shape appreciably. We have just implemented a scheme that re-evaluates their type based on the surrounding environment, which appears to effectively remove the misclassification of such small pixels.

Scale to Production

With our methodology development at a stable version state, we've refined and scripted the workflow all onto a common Python environment to facilitate software sharing and to mass-processing the imagery. Processing time is improved but remains a major challenge. Pan sharpening, segmenting, and classifying a full 16 Gb scene (of which we have several thousand) currently occupies approximately 50 hrs on a single core. We have implemented a tiling scheme that breaks images into small tiles (256x256) for processing, greatly reducing RAM needs, and have identified several processor intensive steps which may be streamlined. At this time, however, we have not initiated multi-threading of a single image. We still expect that with additional effort toward efficiency we will be able to reduce processing time for an image to several minutes.

Finally, partial cloud cover, surface shadowing, and streaking in many images preclude accurate classification. Up to this point our work has only included the manually selected or masked clear sky images. We plan to develop automated detection for some of these imagery flaws during the coming year, but for now are simply manually selecting clear sky imagery.

Current workflow:
1. Collect images (generally in NITF format)
2. Manually identify clear sky imagery (Python database GUI)
3. Orthorectify images (Python GDAL)
4. Pan sharpen imagery (Python GDAL)
5. Tile images (Python)
6. Conduct Feature Extraction (Python Watershed SciKit-Image)
    6a. Create training dataset and build a random forest trees (Python SciKit-Image).
7. Execute Random Forest Classifier (Python SciKit-Image)
8. Merge like Features (Python GDAL)
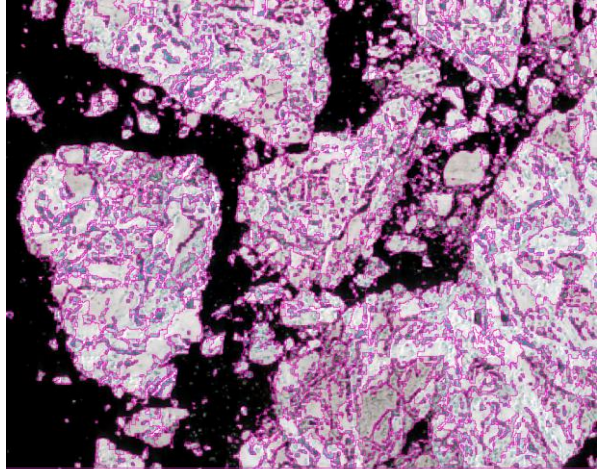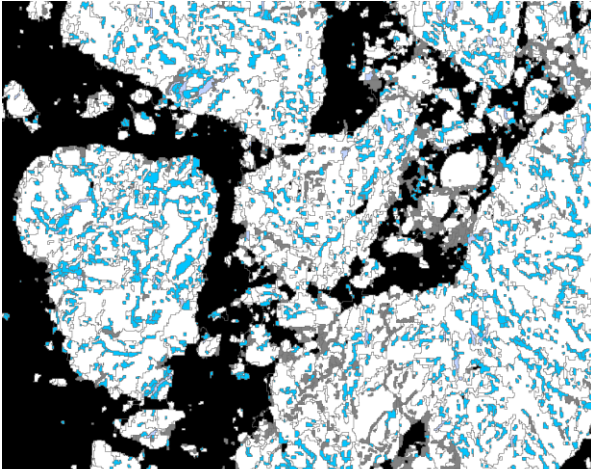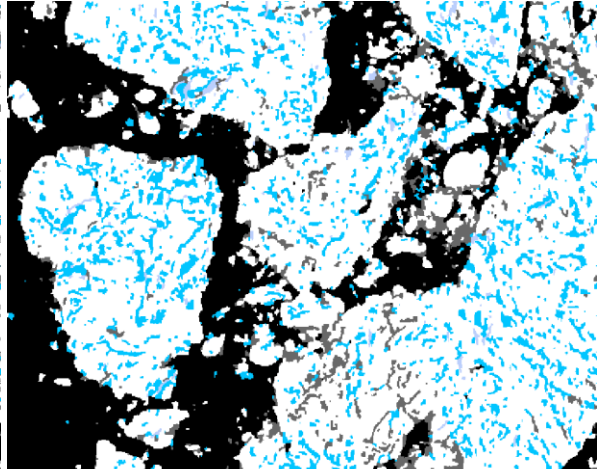9. Calculate statistics and archive result image (Python).
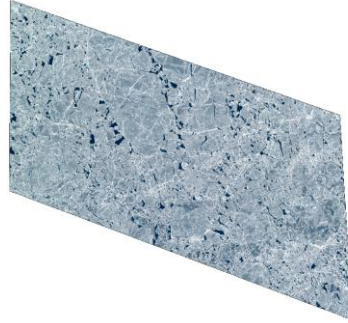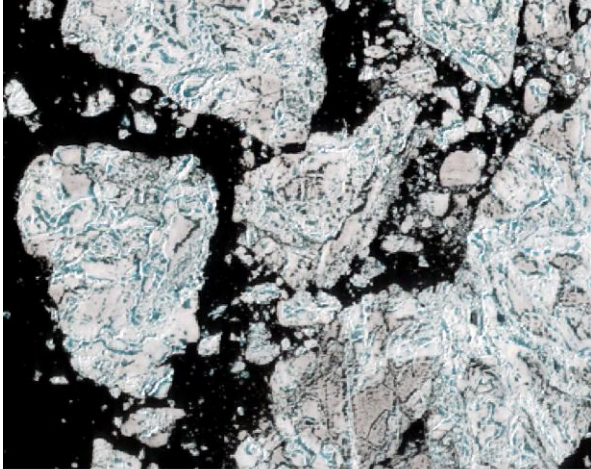
Figure 3. Sample sea ice image from July 10th at the Beaufort sea site (a.) and a zoomed in look at a subset of this image (b.), being segmented using the ENVI feature extraction module, (c.), having the segments classified using our Matlab implementation of a random forest classifier (d.), and having adjacent segments merged together to create a final image. In the final image, blue is melt ponds, purple is ridged ice, and gray is brash ice. Overall accuracy of this image classification was 94.5%, compared to a human classifier.

Figure 4. An image being carried through the processing steps in our open source Python Implementation. Top – original pan-sharpened WV2 image. Middle, combined Otsu's method (segmenting off the unponded, white ice as gray color) and watershed analysis (segmenting the remaining water and pond covered ice into super pixels for classification). Note that watershed analysis is conducted on 256x256 pixel sub images to reduce memory usage, hence the rectangular segment boundaries. Bottom – classification results using random forest technique. Ocean (dark blue), ponds (light blue), and ice (white). Accuracy relative to a human classifier is 96%, with most errors in discrimination between very dark melt ponds and melted-through melt ponds (classed as open water).
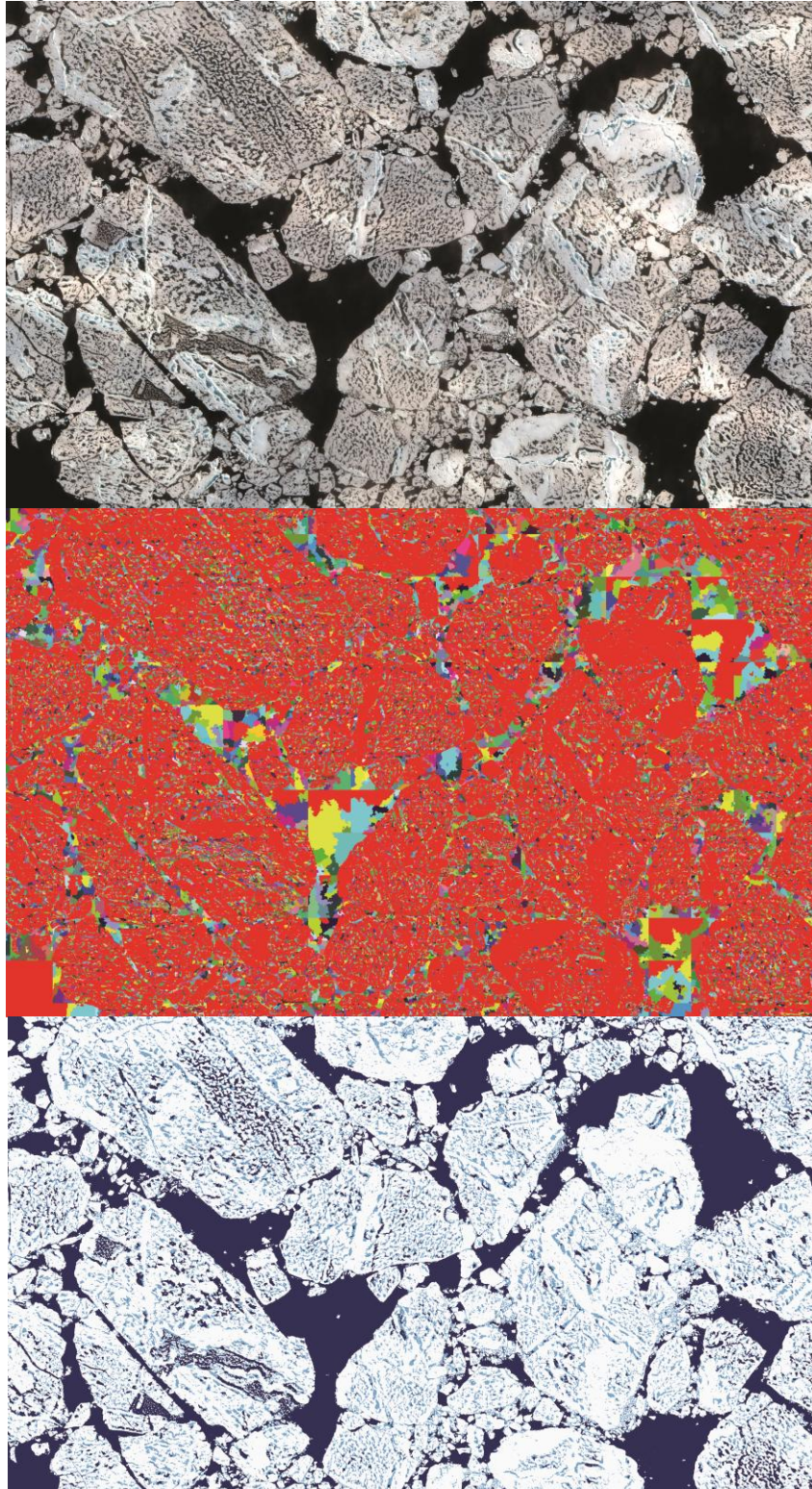
Task 2 – Collect in Situ Data for Field Validation

Though visual validation of our algorithms against human classifiers has been conducted extensively, we also proposed to conduct field validation where we collected higher resolution aerial imagery and ground based measurements of pond locations to compare to our derived pond coverage. We carried out another effort at this validation in FY15 as part of the Norwegian N-ICE experiment, collecting aerial imagery from a helicopter in collaboration with Dmitri Divine, which will be available for validation at higher resolution beneath several collections. This will help us to better address the contribution of edge pixels to overall image classification error.

Task 4 – Scale Up Using Lower Resolution Satellite Assets

We made no new efforts at up-scaling our results in FY15, largely due to limitations on effort.

Task 5 – Disseminate Data and Methodology

Three team members affiliated with this project attended the SAMS image processing workshop in Oban, Scotland July 4-8, 2015 to share our results to date and presented a poster at AGU Fall Meeting 2014. We have also submitted abstracts for the AGU fall meeting 2015. A paper in prep on our methodology was pushed forward somewhat. We have also shared our imagery collection with collaborators on the ONR projects and have re-written our processing scripts into a single open source language in a manner suitable for sharing as a toolkit. Dissemination and technology transfer will come into higher focus during early FY16 as funded time allows.

RESULTS

We have:
- Collected a library of over 5,000 high resolution satellite images at both fixed and lagrangian sites throughout the Arctic Basin.
- Conduct field validation studies capturing high resolution representations of ice beneath satellite images.
- Developed, tested, and validated a workflow for classifying surface types at meter scale in optical sea ice imagery.
- Shifted the processing scheme into a single language, open-source workflow.
- Shared our imagery library with several collaborators and begun writing publications on the methodology.

IMPACT/APPLICATIONS

The results of this study will include both a dataset of key meter-scale sea ice properties derived from our observation sites and the toolkit required to assess these properties in a uniform way from future imagery. This data and these techniques will enable synthesis activities seeking to explain the mechanisms and feedbacks governing ice loss in the Arctic. These synthesis activities are not hypothetical – PI Polashenski has secured NSF funding for a 3 year project to integrate the remote sensing with in situ buoy data to quantify solar partitioning, and the products of this work are highly integrated with several other ONR projects listed above. To ensure our methodology developed can be used beyond this project we plan to release our processing scheme as a standardized shell script downloadable from our website in the coming months.

RELATED PROJECTS

We are coordinating efforts and sharing data with a suite of four closely related ONR projects:

N0001413MP20105: Propagation of Shortwave Radiation Through A Spatially Complex Melting Ice Cover – Lead PI Donald Perovich, USACE-CRREL

N0001413MP20102: Evolution of Melt Pond Geometry on Arctic Sea Ice – Lead PI Ken Golden, University of Utah

N0001413MP20163: The Seasonal Evolution of Sea Ice Floe Size Distribution" – Lead PI Jacqueline A. Richter-Menge, USACE-CRREL

N0001414MP20126: Using discrete element modeling to improve resolved scale floe interaction modeling. Lead PI Arnold Song, USACE - CRREL

Imagery acquisitions and processing efforts are being coordinated closely with PI Richter-Menge's efforts to use similar imagery to track the seasonal evolution of floe size distribution and PI Song's efforts to conduct discrete element modeling. Imagery both from our remote sensing classifications and our planned UAV flights is being shared with PI Golden's efforts to track the evolution of melt pond geometry. Finally, the results of our efforts will produce a surface type classification dataset necessary for Perovich's work to assess the interaction of shortwave radiation with the summer ice cover.

An additional NSF-funded project (PI Polashenski) that will use the derived imagery products to track solar partitioning at the buoy sites began in FY15.

PUBLICATIONS

Stwertka, C., C. Polashenski, E. Deeb, A. Arnsten, D. Perovich, Seasonal evolution of melt pond characteristics derived from Worldview and Quickbird imagery, AGU Fall Meeting 2014.

Wright, N., C. Polashenski, J. Chen, A. Song, E. Deeb, B. Morriss, and D. Perovich, Constraining aggregate-scale solar energy partitioning in Arctic sea ice through synthesis of remote sensing and autonomous in-situ observations, AGU Fall Meeting 2015.

Istomina, L., Heygster, G., Huntemann, M., Marks, H., Melsheimer, C., Zege, E., ... & Katsev, I. (2015). Melt pond fraction and spectral sea ice albedo retrieval from MERIS data–Part 2: Case studies and trends of sea ice albedo and melt ponds in the Arctic for years 2002–2011. *The Cryosphere*, *9*(4), 1567-1578.

Istomina, L., Heygster, G., Huntemann, M., Schwarz, P., Birnbaum, G., Scharien, R., ... & Katsev, I. (2015). Melt pond fraction and spectral sea ice albedo retrieval from MERIS data-Part 1: Validation against in situ, aerial, and ship cruise data. *The Cryosphere*, *9*, 1551-1566.