

Continuing Project

Monroe Community College Geospatial research experience with University of Maine – CAFS

Funding provided by *Skills Training in Advanced Research & Technology (START) Supplemental Funding Request for ATE at Monroe Community College (Award #1955256) with IUCRC Phase 3 at University of Maine - Center for Advanced Forestry Systems (CAFS).*

PI Name: Jonathon Little State University of New York, Monroe CC
Dr. Aaron Weiskittel, University of Maine and CAFS Director

Presenter: Jonathon Little



Justification

- MCC has a NSF Advanced Technological Education grant: (Award #1955256) *Meeting Workforce Needs for Skilled Geospatial Technicians through Virtual Geospatial Information Science Technology (GIST) Education*
- *Program consists of an online stackable A.A.S. in GIST, 24 credit GIST Certificate, and a 9 credit GIST micro-credential for professionals (Python, web mapping, data)*
- *Program offers alumni and peer support, virtual desktop, virtual internships (for credit and paid), and more.*
- *Up until the START grant, the program was lacking research opportunities.*



Objectives

- 1) Partner with the University of Maine's CAFS faculty/staff
- 2) Provide two 2-year Community College students with a paid 8-week summer internship
 - Collect and process field data for remote sensing applications, particularly tree species composition
 - Present work
- 3) Provide 2 students with a credit-based virtual internship
- 4) Present work



(LEFT): Casmir & Rissa recording data. (CENTER): Wayne calculating azimuth with a compass. (RIGHT): Casmir using a BAF prism to identify variable point samples.



Methods

- MCC faculty provide guidance/support
- Students work with CAFS partners in Maine.
 - Barbara Wheatland Geospatial Lab (Tony Guay, Dr. Kasey Legaard, Dave Sandilands, & Dr. Dan Hayes)
 - Penobscot Experimental Forest (Graduate students, Stephanie Wilsey and Rissa)
 - Town of Orono (student Dave Ludwig)
 - University of Maine at Fort Kent (Dr. Ned Rubert-Nason)
 - Dr. Pari Rahimzadeh
 - Schoodic Institute (Dr. Peter Nelson)

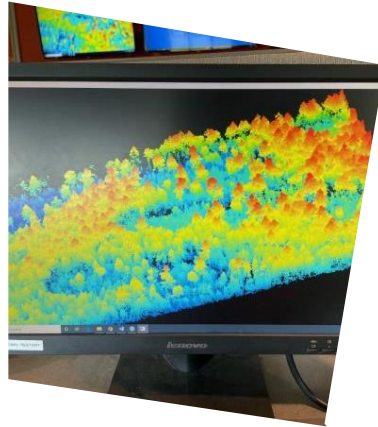


Kasey, Casmir, and Bryon working on Round 2 of training samples for cloud masking.



Major Findings

- Successful experience thanks to Dr. Aaron Wiskettiel and CAFS partners!



Tony giving a presentation on remote sensing and data collection/processing.



Deliverables

- Student Story Map (Week 1-4, 5-8)
- Student Presentations and Poster

Evaluation of Sentinel-2 imagery cloud & shadow masking by a machine-learning algorithm and Fmask post-processing

Casmir Brown
August 4, 2022

Background

The ability to remove clouds from satellite imagery that will be analyzed/processed is crucial in the reduction of inaccuracies and systematic errors when producing maps. Current standard of Fmask developed in 2012 has high accuracy, but not enough for the detailed work done in forests of Maine.

Methodology

Goal is to train the XGBoost ML algorithm to detect cloud and shadow pixels. Two images are needed: one cloud-free (control) and one heavily clouded (variable). Satellite imagery was acquired from Sentinel-2, processed in QGIS.

Round 1 training points on variable images are selected via a principal components analysis to define 300 clusters of similar pixels, from which samples are drawn at random. Points are user-classified in attribute table as clear, cloud, shadow, or uncertain; points are digitally organized and aided by implementation of Python code for QGIS toolbar.

Control, variable images, and classified training points are fed into XGBoost algorithm, which produces intermediate cloud/shadow masks and calculates areas of low confidence from which Round 2 training points on variable images are manually selected. Round 2 points are fed into ML algorithm again in a repeating cycle to fine-tune ML output of cloud/shadow mask.

ML cloud/shadow mask is then post-processed with Fmask's spatial processing to match shadows to a cloud based off the 5007 Bernli and azimuth angles, and the geometric relationship between a cloud and its shadow.

Research Question

If a ML algorithm can be trained to detect and mask cloud/shadow from Sentinel-2 imagery, how will the trained ML algorithm's ability to detect and mask cloud/shadow pixels compare to current standard, Fmask? Is it possible to create a better tool? be available to the public?

Figure 3. Comparing output for blocks of training for ML algorithm: the varied state of forest.

Figure 4. Filtering pixels with trained standard ML algorithm: only cloud/shadow and unclassified pixels in cloud mask.

Figure 5. 2 clouds of training ML algorithm results: high shadow detection and better, more detailed coverage.

Figure 6. Post-processing results from Round 2 ML output: detection of trees, creates less accurate result.

Figure 7. ML Round 1 & 2 Comparison - Cloud (True vs Predicted).

Figure 8. ML Round 1 & 2 Comparison - Shadow (True vs Predicted).

Figure 9. Fmask Post-processed ML output vs. Round 2 ML T19TDM 09/06/21.

Future Work

The form accuracy assessment & quantitative comparison on ML output

- Compare to Fmask accuracy (between 94 and 96.4% dependent on terrain and source of imagery)
- Evaluation of additional band-on imagery
- Further refinement of ML
- More training points for shadow
- Tests for cloud/shadow detection on full-striping imagery
- Software development
 - Expand service to predict all clouds on imagery within Maine
 - QGIS Plugin
- Continued work with an MCC student intern for a GISQGIS Student Capstone Project, Spring 2023

References & Data Sources

Images acquired from Sentinel-2. All maps are projected in WGS 1984 UTM Zone 18N.

- Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sensing of Environment, 118, 83-94. <https://doi.org/10.1016/j.rse.2011.11.014>
- Zhu, Z., Wang, S., & Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm: cloud shadow, and snow detection for Landsat 8-7, 8, and Sentinel 2 images. Remote Sensing of Environment, 150, 268-277. <https://doi.org/10.1016/j.rse.2014.12.014>

Acknowledgements

University of Maine – Center for Research on Sustainable Forests, School of Forest Resources
Dr. James Inglehart, Assistant Research Professor
Barbara Whelan, GIS Geospatial Lab
Dr. David Howe, Director
Andrew Gony, Remote Sensing Specialist
David MacIsaac, Aerial Survey Pilot & Remote Sensing Technician
University of Maine – Center for Advanced Forestry Systems
Dr. Anne Moskaluk, Director
Maine Community College – Department of Chemistry & Geosciences
Jonathan Little, GIS Professor & Principal Investigator
Wayne Rowland, Computer Science & GIS Professor & Senior Practitioner

The experience is thanks to NSF Training in Advanced Research & Technology (IATRT) Supplemental Funding Award for GIS in Maine Community College (Award #201820) with IATRT Phase 2 at University of Maine – Center for Advanced Forestry Systems (CAFS). Available for educational use only. Created 2022.

Results

Fmask VS. Round 2 ML Output – Clouds (Figure 2)

- Fmask cloud coverage is greater, but less detailed (due to buffering of mask)
- XGBoost ML output underestimates cloud coverage

Shadow Output (Figures 3 to 5)

- Fmask overestimates areas of shadow in areas of visible clear pixels
- Underestimation of shadow in areas of visible shadow pixels
- Round 2 ML output refines shadow coverage, reduces overestimation from Round 1
- Post-processing ML Fmask output is less accurate than the Round 2 ML output

ML Round Comparison (Figures 6 & 7)

- Cloud matching nearly identical, increased this cloud detection with Round 2
- Shadow refinement reduces errors, fewer clear pixels labeled as shadow

Post-processing ML Fmask output VS. Round 2 ML Output (Figure 8)

- PP ML predicts clouds with same high accuracy as Round 2 ML output, but adds a buffer
 - Beneficial in areas of steep, thin clouds
- Round 2 predicts shadow much more accurately than PP ML output.
 - PP ML still relies on cloud-shadow matching. Highly selective algorithm needs to be fine-tuned to allow for more variation in potential best-match imagery.

Results

Fmask VS. Round 2 ML Output – Clouds (Figure 2)

- Fmask cloud coverage is greater, but less detailed (due to buffering of mask)
- XGBoost ML output underestimates cloud coverage

Shadow Output (Figures 3 to 5)

- Fmask overestimates areas of shadow in areas of visible clear pixels
- Underestimation of shadow in areas of visible shadow pixels
- Round 2 ML output refines shadow coverage, reduces overestimation from Round 1
- Post-processing ML Fmask output is less accurate than the Round 2 ML output

ML Round Comparison (Figures 6 & 7)

- Cloud matching nearly identical, increased this cloud detection with Round 2
- Shadow refinement reduces errors, fewer clear pixels labeled as shadow

Post-processing ML Fmask output VS. Round 2 ML Output (Figure 8)

- PP ML predicts clouds with same high accuracy as Round 2 ML output, but adds a buffer
 - Beneficial in areas of steep, thin clouds
- Round 2 predicts shadow much more accurately than PP ML output.
 - PP ML still relies on cloud-shadow matching. Highly selective algorithm needs to be fine-tuned to allow for more variation in potential best-match imagery.

Results

Fmask VS. Round 2 ML Output – Clouds (Figure 2)

- Fmask cloud coverage is greater, but less detailed (due to buffering of mask)
- XGBoost ML output underestimates cloud coverage

Shadow Output (Figures 3 to 5)

- Fmask overestimates areas of shadow in areas of visible clear pixels
- Underestimation of shadow in areas of visible shadow pixels
- Round 2 ML output refines shadow coverage, reduces overestimation from Round 1
- Post-processing ML Fmask output is less accurate than the Round 2 ML output

ML Round Comparison (Figures 6 & 7)

- Cloud matching nearly identical, increased this cloud detection with Round 2
- Shadow refinement reduces errors, fewer clear pixels labeled as shadow

Post-processing ML Fmask output VS. Round 2 ML Output (Figure 8)

- PP ML predicts clouds with same high accuracy as Round 2 ML output, but adds a buffer
 - Beneficial in areas of steep, thin clouds
- Round 2 predicts shadow much more accurately than PP ML output.
 - PP ML still relies on cloud-shadow matching. Highly selective algorithm needs to be fine-tuned to allow for more variation in potential best-match imagery.

Center for Advanced Forestry Systems 2022 IAB Meeting

Company Benefits

- Expand institutional knowledge re: student summer research experiences
- Share geospatial educational opportunities at MCC

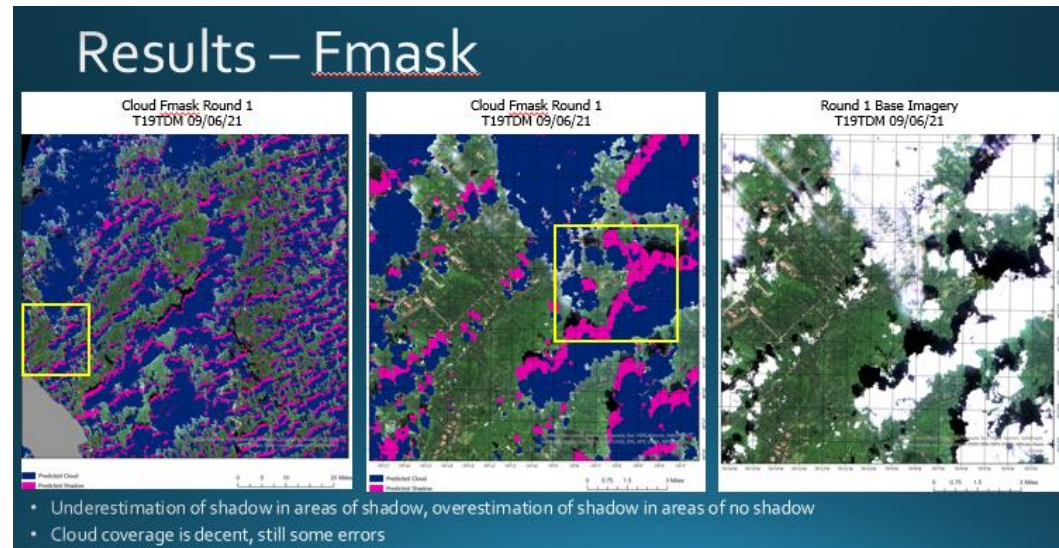


The screenshot shows the Esri website header with navigation links: Products, Industries, Support & Services, Stories, and About. A search icon and a user profile icon are visible in the top right. The main content area features a blurred background image of a classroom. Overlaid on this is a white box with the text 'USER STORY' and the title 'How to Modernize GIS Education with a Microcredential to Ensure Workplace Success'. Below the title is a short paragraph: 'Monroe Community College (MCC) in Rochester, New York, has shaped its coursework with labor market intelligence to ensure that its campus of more than 12,000 students is meeting emerging workforce needs.' A small blue horizontal line is positioned below the paragraph.



Recommendations

- We are always looking for internship hosts for our Geospatial Capstone course offered Feb-May each spring. This is the last course in our 5 course GIS Certificate program and in our 8 GIS AAS. If interested, let me know.
- Consider partnering with CAFS on a NSF REU (Research Experience for Undergraduates)



Summary

- 1) Partnered with the University of Maine's CAFS faculty/staff and provided two 2-year Community College students with a paid 8-week summer internship
- 2) Thank you Dr. Aaron Weiskittel and everyone we worked with this summer with CAFS!
- 3) Looking ahead
 - Two virtual internships planned with Dr. Kasey Iegaard and Barbara Wheatland Geospatial Lab spring of 2023.
 - Present at CAFS IAB meeting spring of 2023
 - Consider partnering with CAFS on NSF REU

