GIST Advisory Board Meeting 2022
Oct 7, 2022

MCC's Geospatial Information Science Technology (GIST) Program

Virtual GIST at MCC

https://atecentral.net/msites/MCC_GIST

Monroe Community College
STATE UNIVERSITY OF NEW YORK

The *Meeting Workforce Needs for Skilled Geospatial Technicians through Virtual Geospatial Information Science Technology Education* project was funded through the U.S. National Science Foundation (NSF) Office of Advanced Technological Education under Grants Award # 1955256 to Monroe Community College. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Welcome and Introductions

MCC's GIST Team

Wayne Howard: GIST Professor + senior collaborator
Jon Little: Geography/GIST Professor + Principal Investigator
Heather Pierce: Geography/GIST Professor + co-PI
Catherine DuBreck: MCC GIST alumni + co-PI

-name
-your role
-your favorite *new* GIST app/software/method that you've used or plan to learn, &/or map

MCC's GIST Advisory Board

Dan Allen: MRB Group
Rui Li: Professor Geography/GIS at SUNY Albany
Vince DiNoto: GeoTech Center PI + GIST Professor
Tabassum Insaf: Research Director-Environmental and Occupational Epidemiology at New York State Dept of Health
Buffy Quinn: GIST/UAS Professor + NSF ATE PI Onondaga CC

Overview of NSF ATE Grant

Meeting Workforce Needs for Skilled Geospatial Technicians through Virtual GIST Education $467,639 *(June 2020-May 2023)*

Project developed one of the nation's first Associate in Applied Science degree programs in GIST that is fully accessible both on campus and online to prepare students for the GIST workforce.

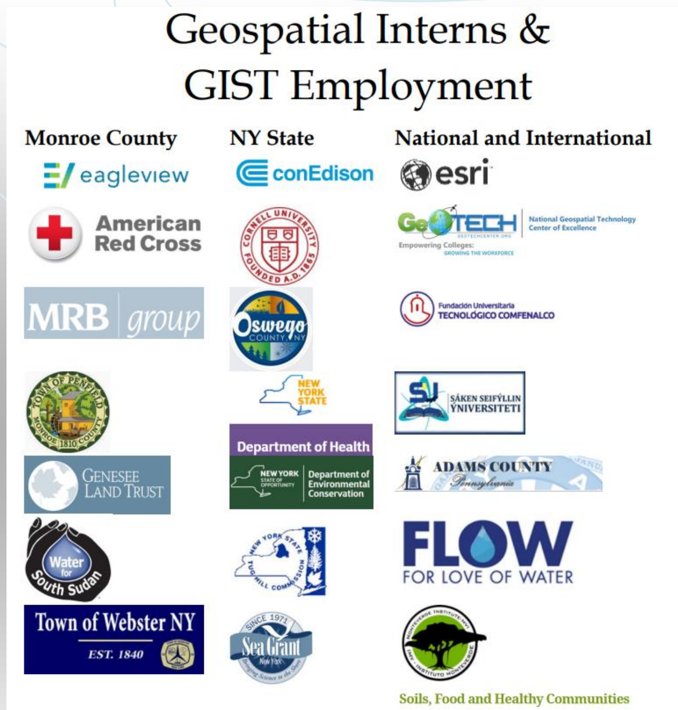
Current program:
- A.A.S. in GIST
- 24 Credit GIST Certificate
- micro-credential
- A.S. Geography concentration in GIST

Grant web site: https://atecentral.net/msites/MCC_GIST

Grant Information

Meeting Workforce Needs for Skilled Geospatial Technicians through Virtual GIST Education $467,639 *(June 2020-May 2023)*

Grant: https://atecentral.net/msites/MCC_GIST



MCC Geography/GIST home page:
https://www.monroecc.edu/depts/geography/

A.A.S. Degree

<u>Objective 1:</u> Build A.A.S. Degree in GIST by adding new courses and updating existing courses, including online.

- Virtual Desktop available
- Available online

Table 1

FALL Year 1	**Cr**	**SPRING Year 1**	**Cr**
Introduction to GIST	3	***Web Mapping***	3
Cartography	3	Spatial Analysis	3
English	3	Art/Foreign Language	3
Introduction to Remote Sensing	3	Physical Geography Lab	1
Math	3	Physical Geography	3
		Physical/Health Education	2
FALL Year 2	**Cr**	**SPRING Year 2**	**Cr**
GIS Data Acquisition and Management	3	***Introduction to Programming for GIS***	3
Statistics	3	Capstone Course in Geospatial Technology	2
Elective	3	American History	3
Human Geography	3	Program Elective	3
Elective	3	Elective	3
		Elective	3

Monroe Community College
10 hrs ·

Did you know MCC has the Geospatial Information Science and Technology (GIST) program for just that?! In the first ever GIST AMA (Ask Me Anything) Session, recent MCC grad and current MCC employee Catherine DuBreck spent some time answering student questions on getting more involved in the field. Check the audio recording out here! https://bit.ly/3iXLB3L

9 credit Micro-credential for GIST Professionals

Table 1

FALL Year 1	Cr	SPRING Year 1	Cr
Introduction to GIST	3	***Web Mapping***	3
Cartography	3	Spatial Analysis	3
English	3	Art/Foreign Language	3
Introduction to Remote Sensing	3	Physical Geography Lab	1
Math	3	Physical Geography	3
		Physical/Health Education	2
FALL Year 2	**Cr**	**SPRING Year 2**	**Cr**
GIS Data Acquisition and Management	3	***Introduction to Programming for GIS***	3
Statistics	3	Capstone Course in Geospatial Technology	2
Elective	3	American History	3
Human Geography	3	Program Elective	3
Elective	3	Elective	3
		Elective	3

Micro-credential

Micro - credential

Micro-credential

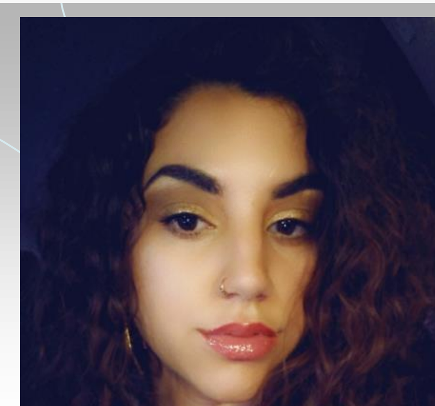
Alumni Support

Objective 2: Provide “Open” GIST lab with Virtual Student Mentors + Alumni Mentors.

New in person computer lab.

Ask Me Anything sessions 4/year (audio)

Alumni Mentor Events including Ask Me Anything (AMA) Audio only
2020 8/26 Agenda || 9/30 AMA || 12/7 AMA
2021 03/31 AMA || 04/28 AMA || 8/28 Agenda || Sept 30 AMA || Dec 3 AMA
2022 March 3 AMA || May 4 AMA || Sept 26 video



Internships

Objective 3: Provide virtual GIST internships

Geospatial Interns & GIST Employment

Monroe County

- eagleview
- American Red Cross
- MRB group
- Town of Penfield, Monroe County 1810
- Genesee Land Trust
- Water for South Sudan
- Town of Webster NY, EST. 1840

NY State

- conEdison
- Cornell University, Founded A.D. 1865
- Oswego County, NY
- New York State Department of Health
- New York State of Opportunity, Department of Environmental Conservation
- New York State Tug Hill Commission
- Sea Grant New York, Since 1971

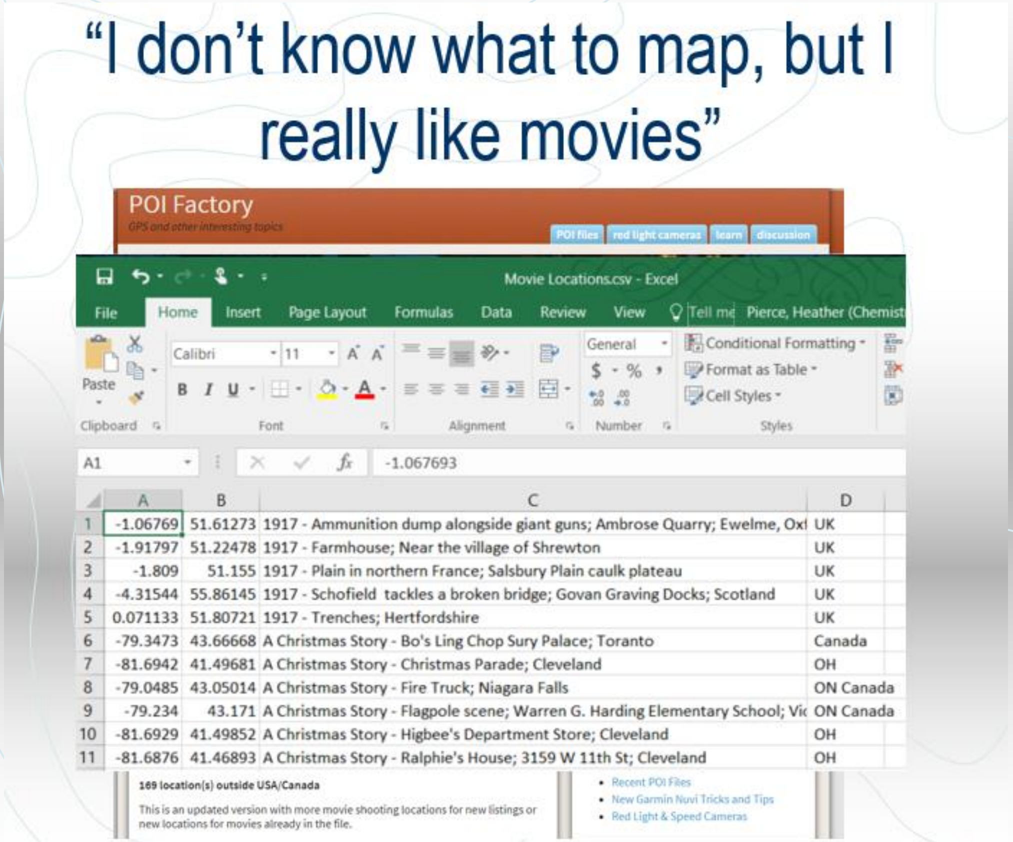
National and International

- esri
- GeoTECH, National Geospatial Technology Center of Excellence, Empowering Colleges: Growing the Workforce
- Fundación Universitaria Tecnológico Comfenalco
- Sáken Seifýllin Ýniversiteti
- Adams County Pennsylvania
- FLOW, For Love of Water
- Instituto Monteverde
- Soils, Food and Healthy Communities

Librarians

Objectives 4: Deliver outreach to public librarians + virtual/in person support from public librarians to introductory students

- Library professional development FL 2020, 2021, & 2022
- GIST Data support FL 2020, 2021, & 2022



Outreach

Objectives 5: Recruitment of GIST students + Professionals

- Upward Bound Summer camps
- Virtual GIS Day (GeoTech Center Collaboration)
- Every semester mapathon!

GeoTECH GEOTECHCENTER.ORG | National Geospatial Technology Center of Excellence
Empowering Colleges: GROWING THE WORKFORCE

Status of New courses:

GEG 236: Geospatial Data Acquisition and Management

Timeline: Offered FL21 and FL22

Course Description: Learn important geospatial data management skills that are in high demand!

Topics Covered:

- Data models, data formats and data management
- Best practices for data collection and processing
- Database management systems and schema
- Advanced geodatabase design
- Topology
- Enterprise geodatabase design
- Using QGIS in a multiuser, postGIS environment
- Introduction to Python automation

Pre-requisites: Introduction to Remote Sensing (GEG 133) or permission of the instructor.

Software: ArcGIS Pro, QGIS, PostgreSQL/PostGIS

Status of New courses:

GEG 236: Geospatial Data Acquisition and Management

Download a course flyer: https://atecentral.net/msites/MCC_GIST

- Adding some open source and third-party drone data processing software
- Added an ALE on Creating a simple Geodatabase
- Adding an ALE to help ease into SQL – based on student feedback

Comments/Questions?

Status of New courses:

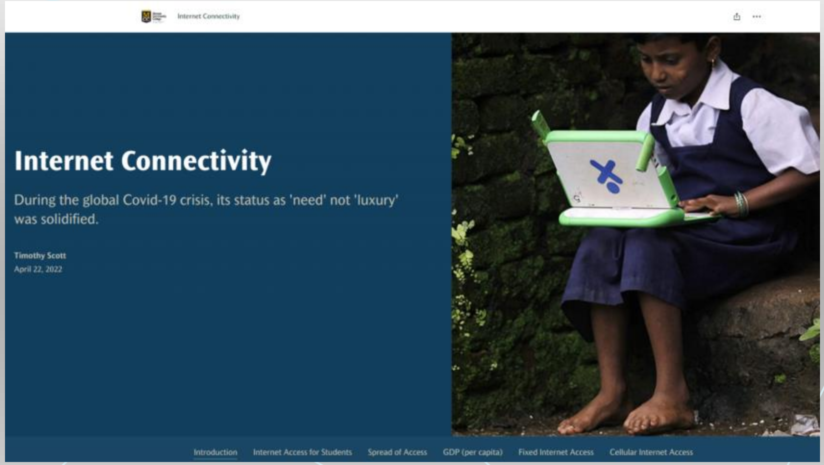
GEG 237: Web Mapping

Offered Online Spring 2022 and plan to offer every spring online

Course Description: Students will learn about the usefulness and application of Web GIS tools such as ArcGIS Online, Google Maps, Volunteered Geographic Information, and Map services (Mapbox). Students will become adept at storing and accessing spatial data in the cloud, practice developing Story Maps to communicate spatial data, and learn how web mapping is key to mobile GIS applications such as field data collection.

Specific Topics from last spring

- What is Web Mapping (1 week)
- Story Maps (4 weeks) w/ a little ArcGIS Hub
- Esri Dashboards and Dashboard Operations (2 weeks)
- ArcGIS Field Maps (2 weeks)
- Mapbox: Create map/data visualization (1 week)
- MapBox: Sheet Mapper Live with basic scripting (2 weeks)
- Volunteered GMapping: Humanitarian Mapping (1 week)
- Story Map Project (3 weeks) - Sample Story Maps



Pre-requisites: Digital Earth (Intro to GIS), or permission of the instructor

Status of New courses:

GEG 237: Web Mapping

Course flyer:
https://atecentral.net/downloads/12223/GEG237_Geospatial_Flyer_SP22-Final.pdf

Changes for Spring

-Plan to combine 2 labs into one.
-Add one new lab on ArcGIS Hub?

Comments/Questions?

Status of New courses:

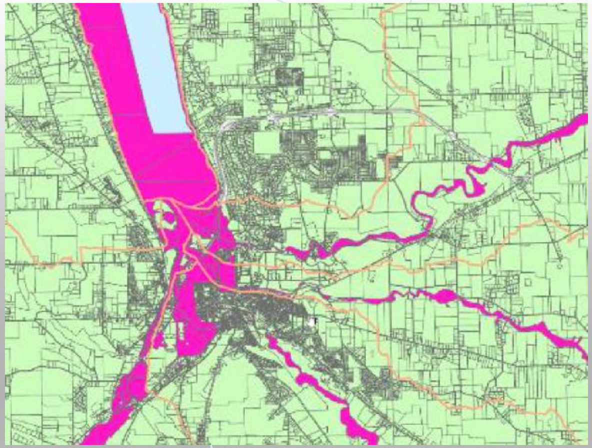
GEG 238: Introduction to Geospatial Programming

Timeline: Offered SP22. Will be offered SP23.

Course Description: Learn to automate geoprocessing tools and to modify and create scripts in Python.

Topics Covered:

- Introduction to Python and geoprocessing tools.
- Model Builder and programming fundamentals.
- Geoprocessing and object-oriented programming.
- GIS inventory using the data access module.
- Debugging and error handling.
- Data access and creation with geodatabases.
- Working with geometry and map layout.
- Jupyter notebooks.



Pre-requisites: GEG 130, GEG 133, and GEG 230 or GEG 236 all with a grade of C or higher or permission of the instructor.

Software: ArcGIS Pro, Jupyter Notebooks

Status of New courses:

GEG 238: Introduction to Geospatial Programming

Considering adding a module on Rasters – but need to remove something. i.e., Working with Map Layout

Comments/Questions?

Status of Course Updates + Online Conversion:

Virtual desktop (60 concurrent users) Introduction to GIS (a.k.a Digital Earth)

- ArcGIS Pro transition completed
- Online conversion completed

Cartography

- QGIS / Online conversion completed

Remote Sensing

- ArcGIS Pro/online conversion completed

Spatial Analysis

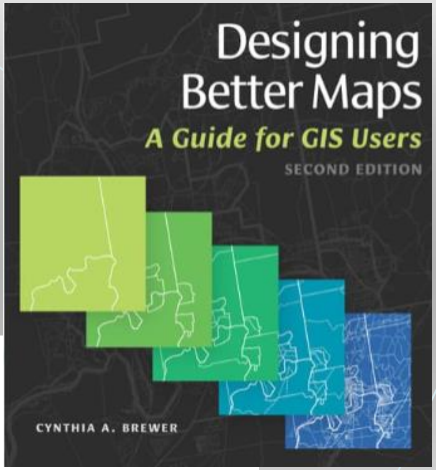
- ArcGIS Pro/online conversion completed

Capstone in Geospatial Technology

- ArcGIS Pro conversion completed
- Online conversion completed
- Many course revisions – SP21 & SP22
- Enhancing module on Ethics in GIS

Status of other courses:

GEG 131: Cartography
Offered: Fall, online
Course Description: This course introduces fundamental cartographic concepts. Upon completion of this course, successful students will be able to employ design principles to create effective visual representations of geographic data (e.g. maps) in different formats (e.g. hardcopy, digital, web). Specific topics include map element design, typography, label design, selecting appropriate color schemes, and symbology. (SUNY-Art)



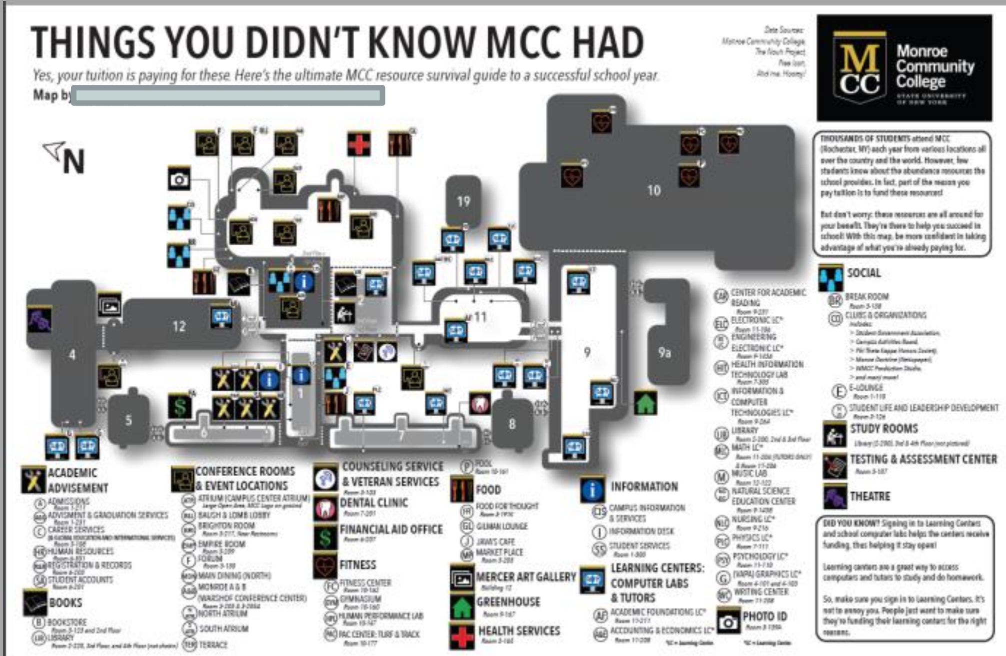
Pre-requisites: None, but should have basic computer skills

Software: QGIS

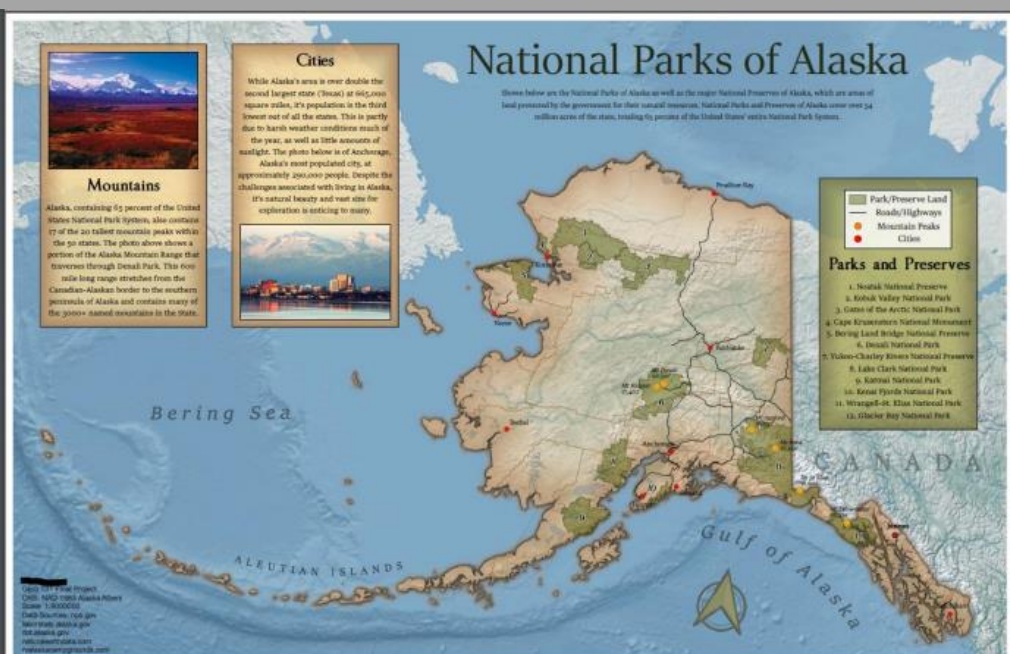
Course Learning Outcomes

By the end of the semester, you will be able to:

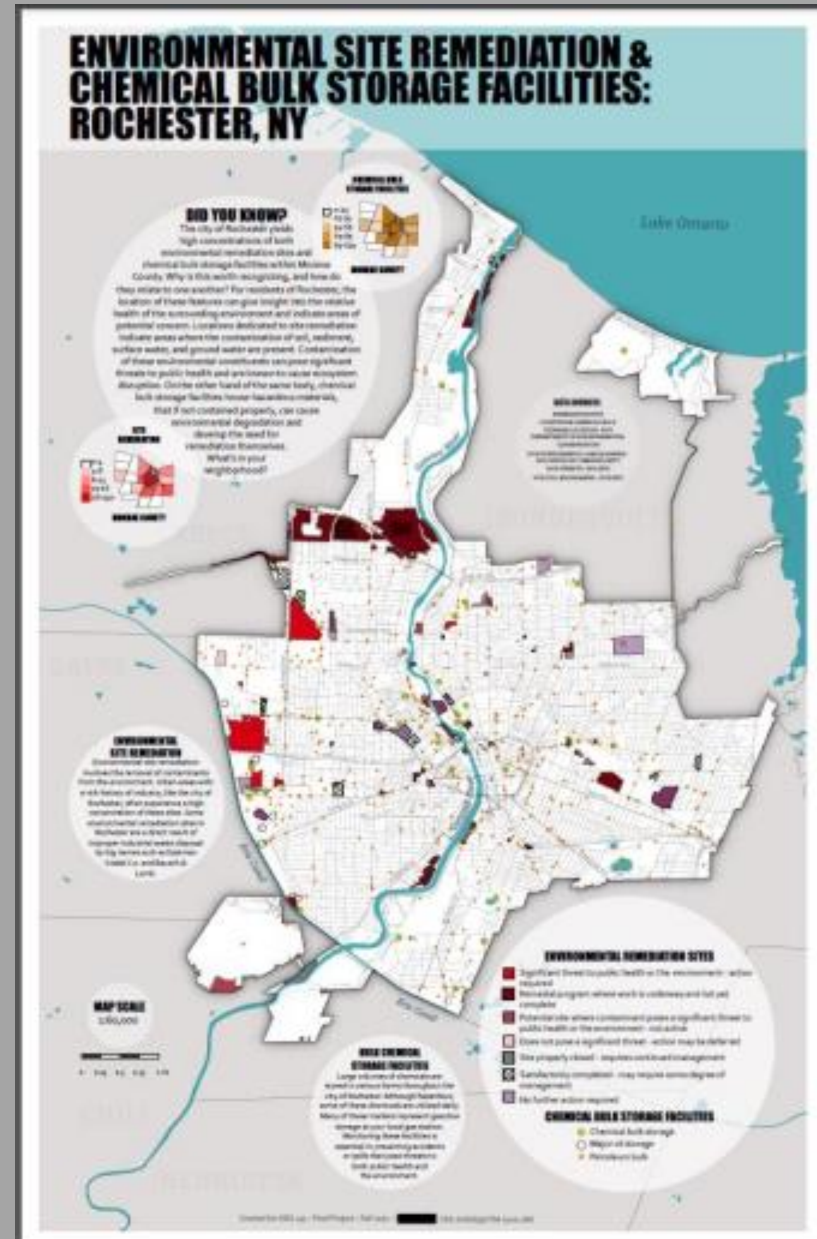
1. Categorize maps according to type.
2. Identify the essential components of a map.
3. Select an appropriate geographic referencing system (i.e. datum, projection, coordinate system) for a given purpose.
4. Determine the appropriate thematic map for a given purpose.
5. Produce maps that effectively communicate quantitative and qualitative geographic data.
6. Design professional quality maps that use cartographic principles.
7. Critique maps for appropriate use of cartographic design principles.



THINGS YOU DIDN'T KNOW MCC HAD
Yes, your tuition is paying for these. Here's the ultimate MCC resource survival guide to a successful school year.
Map by
Monroe Community College
N
19
10
12
4
11
9
9a
5
8
ACADEMIC ADVISEMENT
CONFERENCE ROOMS & EVENT LOCATIONS
COUNSELING SERVICE & VETERAN SERVICES
DENTAL CLINIC
FINANCIAL AID OFFICE
FITNESS
FOOD
INFORMATION
LEARNING CENTERS: COMPUTER LABS & TUTORS
MERCER ART GALLERY
GREENHOUSE
HEALTH SERVICES
BOOKS
PHOTO ID
SOCIAL
STUDY ROOMS
TESTING & ASSESSMENT CENTER
THEATRE



National Parks of Alaska
Mountains
Cities
Park/Preserve Land
Roads/Highways
Mountain Peaks
Cities
Parks and Preserves
Bering Sea
ALEUTIAN ISLANDS
Gulf of Alaska
CANADA
[redacted]



ENVIRONMENTAL SITE REMEDIATION & CHEMICAL BULK STORAGE FACILITIES: ROCHESTER, NY
DID YOU KNOW?
Lake Ontario
ENVIRONMENTAL REMEDIATION SITES
CHEMICAL BULK STORAGE FACILITIES
MAP SCALE
[redacted]

Status of other courses:

GEG 230: Spatial Analysis & GIS
Offered: Spring, online
Pre-requisites: GEG 130 (Intro GIS) w/ C or higher

Software: ArcGIS Pro

Description: This course introduces students to geospatial analyses that are used for problem-solving and decision-making. Students will learn how to perform these analyses using Geographic Information Systems (GIS) software. As a guiding framework, this course shows how GIS is used to answer fundamental questions in geography that are related to problem-solving and decision-making in a wide range of careers and academic disciplines. Upon completion of this course, students will be able to make informed decisions when choosing GIS-based approaches for conducting geospatial analyses.

Lecture Content	**Lab Skills/Tools covered**
Intro Activities	
Topic 1: Review of GIS Basics	Review of GIS Basics
Topic 2: Point Data trends, part 1	Collect Events, Central Feature, Mean Center, Median Center, Generate Near Table, and Calculate Field
Topic 3: Point Data trends, part 2	Project, Directional Distribution (Standard Deviation Ellipse), IDW, Kriging, and Kernel Density
Topic 4: Networks & Modelling Interaction	Network Analyst – Creating a Network Dataset, Route, Closest Facilities, Location-Allocation
Topic 5: Raster Basics Exam Review	Project Raster, Int tool, Raster Calculator, Extract by Mask, Extract by Attributes
Exam 1 on Topics 1-4	
Topic 6: Working with Lidar data	LiDAR basics, LAS to Raster, Hillshade, Linear Directional Mean
Topic 7: DEMs & Modeling Overland Paths	Mosaic to new Raster, Environment Settings, Feature to Raster, Slope, Cost Distance Analysis tools
Topic 8: Hydrology Tools	Hydrology tools, Reclassify (raster)
Topic 9: Site Suitability Analysis	Zonal Statistics as Table, Join Field tool, Erase, Select by Location
Topic 10: Mapping Inequality Exam Review	TBA

We're Not In Kansas Anymore: Tracking Tornado Alley and Its Shift Across America

[redacted]

GEG 230 – Spring 2022

Research Question

The change in season and the subsequent temperature and moisture levels across the central US allow Tornado Alley to shift its position over the course of each year, potentially endangering millions of people. Where exactly does the average center of Tornado Alley fall? How far has its center shifted since 1980?

Methodology

US tornado track data was acquired from National Oceanic and Atmospheric Administration (NOAA). The Fujita scale was used to classify tornados prior to 2007 (Table 2) it was updated to the Enhanced Fujita scale after studies showed that more catastrophic damage could be achieved with lesser wind speeds, deducing that some windspeeds prior to 2007 could have been overestimated. Because of this difference, I chose to compare wind speed values of both scales and chose F/EF2 as a starting point for creating the Kernel Density maps (Figures 1-3), since that rating had the most similar initial break value (F=113MPH, EF=110MPH) while also being more destructive tornados. Attribute queries divided data into three 13-year periods and made into separate layers, which were then each processed with the Kernel Density tool. The output cell size was 0.17.

Utilizing the same data layers that created the KD maps, each layer was individually processed with the Mean Center tool to find the average location of all tornados for that time period (Figure 4). The measure tool was used to find the total distance that the center of all tornados traveled over time.

Tornado data from 2007 – 2020 was utilized for the Directional Distribution maps, as this is the most current data. These maps show the general area that the tornados occupy, although do not perfectly encompass every tornado within its radius. For the first map (Figure 5), the Directional Distributions of tornados were calculated for the time frames Feb-April and May-July, when it is peak tornado season. This illustrates the shift of the range from southeast to the northeast. The second map (Figure 6) illustrates the average range of tornados by magnitude over the course of a year. The third map (Figure 7) overlays tornado paths over Directional Distribution, to help provide clarity as to why the EF-4 tornado DD is skewed more than the other magnitudes.

Magnitude	2007-2020	1993-2006	1980-1992
F0/EF0	5,098	10,112	5,583
F1/EF1	3,185	4,399	4,744
F2/EF2	1,442	1,274	1,714
F3/EF3	321	384	462
F4/EF4	80	81	100
F5/EF5	9	5	8

Table 1. Totals of tornados between three different time periods, broken down by magnitude

US Tornados, 1980-2020

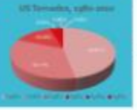


Chart 1. Breakdown of tornado frequency by magnitude. EF0 and EF1 are most common, but cause the least amounts of damage and have the lowest wind speeds.

Background

Each year, more tornados than anywhere else in the world form in the center of the contiguous United States. This is the perfect spot where cool, dry air masses from over the Rockies can crash into much warmer, moister air masses from the Gulf of Mexico. The relatively flat expanses of land provide ample space for tornados to tear across the landscape for miles, giving the area the nickname "Tornado Alley".

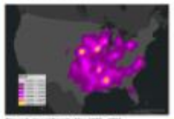


Figure 1.

Figure 2.

Figure 3.

Mean Center of All Tornados, 1980-2020

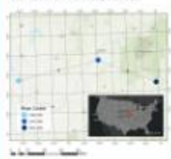


Figure 4.

Results

Kernel Density analysis revealed the different pattern of paths for each time period. Between 1980 – 1992, 3 main hotspots were located: Nebraska, Iowa, and Mississippi. 1993 – 2006 showed one main hotspot in Arkansas, which fell somewhere in the center of the previous hotspots. By 2007 – 2020, two hotspots in Mississippi and Alabama are the most prominent. Something to note is that while all maps have the same starting value, the maximum values increase with every subsequent time period. Figure 3 reveals that EF2+ tornados occurring the South/Southeast have been happening at a greater frequency than ever before. During this same time period, the hotspots from Figure 1 are barely visible. Figure 1 also reveals a more widespread pattern, but this could be attributed to the F-scale used at the time which may have overestimated windspeeds. I chose a rough equivalent breakpoint of EF2, but for future study it would be better to convert each F-scale speed to its EF-scale equivalent to get a more accurate KD map.

While KD analysis help define the tornado hotspot regions, the Mean Center tool helped define its exact average center. Comparing the MCs for each time period shows that all three were located within the southern part of the state of Missouri but viewing the data in large scale highlights the differences: a consistent noticeable shift eastward and the slightest shift southwards. The Measure tool revealed an 82.36-mile change in the average center of all tornados occurring between 1980 and 2020.

Figure 5 illustrates the general shift of tornado distribution annually during peak tornado season. In the earlier months of February – April, distribution is much further southeast with an east-to-west spread. In the months of May – July, the distribution shifts to the northwest, with a less elliptical spread. The Feb-April DD falls almost squarely within the average DD for the entire year, while the May-July DD falls further outside the average DD, which is most likely due to occasional tornados occurring near the West coast or Pacific Northwest. The Mean Centers of tornados by magnitude in Figure 5 highlight the same general area of southern Missouri that the MC of All tornados does, but the MC of EF5 tornados seem to have moved slightly further south into Arkansas, and EF2,EF3 tornados have MCs to the east, which is consistent with the previous analysis findings.

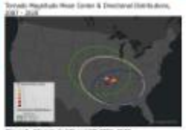


Figure 5. Magnitude MC and DD 2007-2020

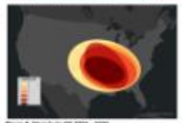


Figure 6.

Figure 7.

Fujita Scale		Enhanced Fujita Scale	
F-0	40-72 mph winds	EF-0	65-85 mph winds
F-1	73-112 mph	EF-1	86-110 mph
F-2	113-157 mph	EF-2	111-135 mph
F-3	158-206 mph	EF-3	136-165 mph
F-4	207-260 mph	EF-4	166-200 mph
F-5	261-318 mph	EF-5	>200 mph

Table 2. F and EF Scale Windspeeds

Discussion

The biggest anomaly that resulted from the analysis was the difference in DD for EF4 tornados compared to others, seen in Figure 6. The DD is highly elliptical and has a more North-South alignment than the other DDs. One possible reason for this is the Appalachian mountain range, which provides less-than-perfect conditions for a tornado to thrive. Tornados fare much better in the Great Plains, where there is more open space for storms to develop and have no topography to inhibit its growth or movement. To investigate this further, Figure 7 illustrates the paths of EF3 and EF4 tornados overlayed on the DDs. Several EF4 tornados occurred much further north than usual in North Dakota and Minnesota. EF3 tornados seemed to form in the same places as EF4, but had a wider range extending further west, south, and even east of the Appalachians.

Many of the tracks on the East coast did not originate from the Midwest, so there may be differences in the driving forces behind their creation. Tornados are known to form in the wake of hurricane paths, which are a common occurrence on the coast, however, any storm generated has little room to grow between the coast and the Appalachians, which could limit the severity of storm ratings to a maximum of EF3. Nonetheless, the wide range of EF3 storms allows for a more even and spherical DD than EF4 storms. Also to note, is that while tornados form on the East coast, their range does not extend far beyond Virginia during the entire course of the year. This also limits the sphericality of the distributions.

References & Data Sources

All maps use the NAD 1983 US Contiguous Albers Equal Area Conic Projection. Figures 1-3 & 5-7 are seen at a 1:18,750,000 scale.

- NOAA's NWS Storm Prediction Center: https://www.spc.noaa.gov/gis/svrgis/
- US Census Bureau for country and state borders. https://www.census.gov/data.html

Acknowledgements

Thank you to Professor Pierce for her guidance and advice on this project and throughout the semester!

#WDWBDE: Walt Disney World Best Day Ever

[redacted]

GEG 230 – Spatial Analysis & GIS Spring 2021

Background

As an avid Disney Park-goer my whole life, I carry much love and passion for Walt Disney World (WDW). I wanted to use the spatial analysis tools I learned in this course to create the "Perfect Day at Disney" or the most time-efficient way to spend a day in Magic Kingdom. Planning your day ahead of time is key when vacationing in the "Most Magical Place on Earth", especially for a first-time guest. Not only can the park feel overwhelming, but also frustrating when you need to take wait times into account.

My goal for this project was to create a network dataset of Magic Kingdom (MK) and then run a network analysis to create an itinerary of what order to stop at each attraction. The route takes attraction duration and wait time(s) into account, as well as suggests what the three closest restaurants are at various intervals throughout the day.

I chose Magic Kingdom for my analysis for a few reasons. For one, it's the first WDW park to open on October 1, 1971, making many of its attractions "must-sees" for any first-time guest. MK is also the most visited theme park in the world, 2019, alone, counting 20+ million guests. Lastly, Magic Kingdom is the park that embodies the magic atmosphere that is Walt Disney World. It completely transports its guests to a whole new world where they can forget the outside world have fun!

Introduction

How does one best maximize their time whilst on vacation? Better yet, how does one best maximize their time while vacationing at a place like Walt Disney World in Orlando, Florida?

Figure 1: The digitized pathways of Magic Kingdom park.

Methodology

In ArcGIS Pro, I created a geodatabase and then a feature dataset. Within the dataset, I created a lines feature class so that I could digitize the walking paths of Magic Kingdom (using OpenStreetMap as my basemap). I used Tobler's value for a flat trail (.012 mins/meter) and multiplied that by the shape length of the polylines. I created my network dataset using the feature dataset and then adjusted the properties accordingly before building my dataset. I then created a Magic Kingdom attractions point feature class and a Magic Kingdom restaurants point feature class within the geodatabase. To find my route, I ran a network analysis with the attractions layer imported as my stops. Once I had my route and stop sequence, I selected three points throughout the "day" (at the start, middle, and end) to choose when to eat and exported them to their own layer so that I could import them as incidents when I ran the closest facility network analysis (the restaurants layer was imported as facilities).

Map 1: "Where to Start and Where to Go From There"

Map 2: "Where (and When) to Eat?"

Results

As you can see, *Map 1* displays a step-by-step guide of what order one should do each attraction currently offered at Magic Kingdom, with your starting and ending point(s) being at the front of the park. While the time is not exact, as we are assuming the walking speed is 12 mins/km and the attraction is a walk-on (i.e., no wait time), the time calculated is ≈6 hours to cover the entirety of Magic Kingdom and a total walking distance of 3,982 yards. *Map 2* displays the 3 closest eating options for 3 different points along your day— at the beginning of your day, about midway through your day, or right after your last attraction of the day!

Figure 2: Magic Kingdom attractions attribute table

Figure 3: Magic Kingdom restaurants attribute table

Discussion

Although I pride myself in knowing Magic Kingdom like the back of my own hand, I learned new things while working on this project! For one, the exact middle attraction of the park, is literally positioned in the center of the park (either Mickey's PhilharMagic or Prince Charming Regal Carrousel, depending on which side of the park you start on). I believe these maps could be used as a resource for vacation planning, but one must remember that wait times and a slower walking pace were not considered. Although I did calculate them, they weren't accurate enough for me to feel like they could be properly used (even as a resource). Future work would be possibly expanding this network dataset to include more of WDW (other parks, resorts, Disney Springs, etc.)"

Figure 4: Back entrance to Pecos Bill Tall Tale Inn & Café

References & Data Sources

The data required for this analysis was a feature class/shapefile of Magic Kingdom walking paths, Magic Kingdom attractions, and Magic Kingdom Restaurants.

All data used was created myself within ArcGIS Pro, using OpenStreetMap, disneyworld.disney.go.com, and my own knowledge as reference(s).

Acknowledgements

I'm so grateful to Professor Pierce and the whole geography department at Monroe Community College. All opinions, findings, conclusions, or recommendations expressed in this presentation are solely my own and do not necessarily reflect the views of Professor Pierce or Monroe Community College, nor their endorsements.

GEG 239 Unpaid virtual internships (for credit)

Spring 2022 Hosts:

Benemérita Universidad Autónoma de San Luis Potosí (Mexico)
American Red Cross – Rochester, NY
City of Traverse City, FLOW – Michigan
FLOW - Michigan
Freshwater Future – Michigan
Genesee Land Trust – Rochester, NY
GeoTech Center – Louisville, KY
MCC Library
NYS Department of Health
NYS Seagrant

Comments/Questions?

Paid internships Summer 2021-2023

2021:

Water for South Sudan – Rochester, NY

NYS Department of Health – Albany, NY

2022:

Genesee Land Trust – Rochester, NY

NYS Department of Health – Albany, NY

University of Maine – Advanced Research & Technology (START) – Center for Advanced Forestry Systems (CAFS) project

Comments/Questions?

MCC Student work U Maine

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

Casmir Brown
August 4, 2022

Background

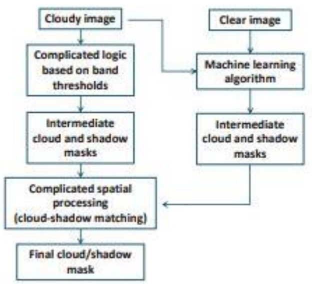
The ability to remove clouds from satellite imagery that will be analyzed/processed is crucial in the reduction inaccuracies and systematic errors when producing maps.
Current standard of Fmask developed in 2012 has high accuracy, but not enough for the detailed work done in forests of Maine.

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 solar zenith and azimuth angles, and the geometric relationship between a cloud and its shadow.

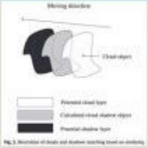


Figure 1. *Cloud shape (white) is projected to ground (grey) and slid forward and back along the ground until a best match is obtained with the potential shadow layer. Potential shadow is retained in the final mask only for this potential best-match overlap area.*

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?

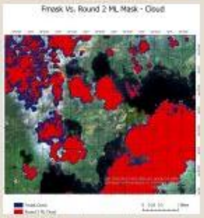


Figure 2. Comparing output for 2 rounds of training for ML algorithm to the current standard, Fmask .

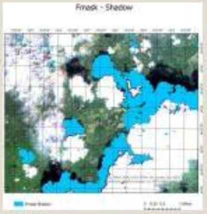


Figure 3. Glaring errors with current standard on shadow masking. Missed shadows and unfilled holes in cloud mask.

Figure 4. 2 rounds of training ML algorithm results in high shadow detection and better, more detailed coverage.

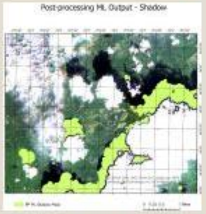


Figure 5. Post-processing results from Round 2 ML output diminishes returns, creates less accurate result.

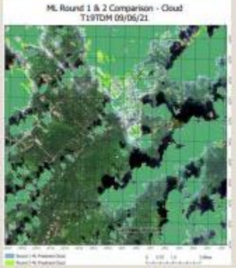


Figure 6. Both Rounds of trained ML output show little variation in cloud detection.

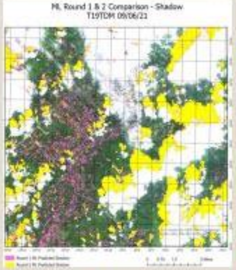


Figure 7. Both Rounds of trained ML output placed together show increased accuracy of identifying shadow pixels with a second round of training.

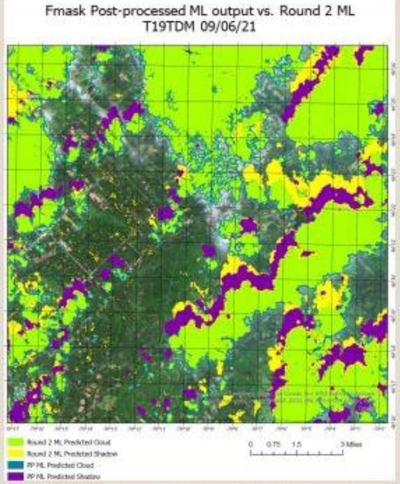


Figure 8. Round 2 ML prediction masks with Post-processed ML Fmask output. Optimal combination of layers to have the most cloud/shadow coverage would be either ML or PP ML cloud mask, paired with the Round 2 ML predicted shadow.

Fmask shadow prediction needs to be altered to allow for greater variation in where potential shadow layer is, and refinement of best-match properties.

Further rounds of training could enhance cloud detection, and reduce further shadow overestimates on areas of clear pixels.

Results

Fmask VS. Round 2 ML Output – Clouds **(Figure 2)**
- Fmask cloud coverage is greater, but less detailed (due to buffering of mask)
- Round 2 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 thin 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 wispier, 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.

Future Work

- Perform accuracy assessment & quantitative comparison on ML output
  - Compare to Fmask accuracy (between 92.4 and 96.4%, dependent on version and source of imagery)
- Evaluation of additional leaf-on imagery
- Further refinement of ML
- More training points for shadow
- Train for cloud/shadow detection on fall/spring imagery
- Software development
  - Expand service to predict all clouds on imagery within Maine
  - QGIS Plugin
    - **Continued work with an MCC student intern for a GEG239 Student Capstone Project, Spring 2023**

References & Data Sources

Images acquired from Sentinel-2.
All maps are projected in WGS 1984 UTM Zone 19N.

- 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.2011.10.028
- Zhu, Z., Wang, S., & Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. Remote Sensing of Environment, 159, 269–277. 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. Kasey Legaard, *Assistant Research Professor*
Barbara Wheatland Geospatial Lab
Dr. Daniel Hayes, *Director*
Anthony Guay, *Remote Sensing Specialist*
David Sandilands, *Aerial Survey Pilot & Remote Sensing Technician*
University of Maine – Center for Advanced Forestry Systems
Dr. Aaron Weiskittel, *Director*
Monroe Community College – Department of Chemistry & Geosciences
Jonathon Little, *GIS Professor & Principal Investigator*
Wayne Howard, *Computer Science & GIS Professor & Senior Personnel*

This experience is thanks to 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). Available for educational use only. Created 2022.

Questions/Comments and Thank you!

Email:

Jon jlittle@monroecc.edu

Wayne whoward11@monroecc.edu

Catherine cdubreck001@monroecc.edu

Heather hpierce@monroecc.edu

Geography

Geospatial Information Science and Technology (Gist)

- Certificate   Department of Chemistry and Geosciences

GET THE "G.I.S.T." ON THE GROWING FIELD OF GEOSPATIAL INFORMATION SCIENCE AND TECHNOLOGY

Start My Application >

Explore Careers >

School of Science, Technology, Engineering & Math (STEM)

Geospatial Information Science and Technology (G.I.S.T.) is used virtually everywhere. It converts remote sensing information provided by satellites and imagery into digital data.