

Summer Fog

Frequency Patterns and Impact

on Intertidal Organisms around Washington Coast

from GOES-17 Satellite Imagery, Field Photos, and Field Sensors

Autumn Nguyen^{a,b},
Jessica Lundquist^c,
Steven Pestana^c,
Eli Schwat^c

NSF REU-Blinks Program 2022
Summer 2023

^aFriday Harbor Laboratories REU-Blinks 2023

^bMount Holyoke College, South Hadley, MA 01075, USA

^cCivil and Environmental Engineering, University of Washington, Seattle, WA 98195-2700, USA

Contact information:

Autumn Nguy n

Mount Holyoke College, South Hadley, MA 01075, USA

ngoc54n@mtholyoke.edu

autumn.yngoc@gmail.com

Keywords: coastal fog, fog and low clouds cover, fog frequency,
GOES-17, satellite, cloud top height, machine learning, intertidal organisms, Washington

Table of Contents

Abstract

1. Introduction

Significance of fog

Knowledge gap about fog in Washington

Motivation from intertidal organisms

Research methods: satellite, field cameras, and field sensors

Research questions:

1. Frequency patterns of fog and low clouds cover (FLCC) around WA coast
2. Comparison of FLCC from satellite and from camera photos
3. Impact of FLCC on intertidal organisms

2. Background

2.1) Definition

2.2) How fog protects intertidal organisms

2.3) Other ecological roles of fog

2.4) Fog frequency trend in previous studies, and anecdotally

3. Methods

3.1) Satellite data processing

3.1.1) Download

- 3.1.2) Reproject
- 3.1.3) Load, stack, chunk
- 3.1.4) Calculate and plot FLCC frequency

3.2) Field sensor data processing

- 3.2.1) Field cameras and sensors
 - Support Vector Machine model
 - Criteria to label training photos
 - Process of creating the training set
 - Confusing instances
 - Model parameters and cross-validation accuracy
 - Output data frame

3.3) Calculate midday low tide hours

4. Results

- 4.1) Overall frequency map from satellite data
- 4.2) Local fog frequency from satellite data and from camera photos
- 4.3) FLCC frequency by months

4.4) FLCC frequency by times of the day

4.5) Comparing satellite-detected cloud height and camera-derived fog presence

4.6) FLCC-protected hours during midday low tide around San Juan Island

4.7) Impact of FLCC on temperature

4.8) Changes in cloud top height between the first and second half of the summer

5. Discussions

6. Conclusions

7. Future steps

Acknowledgement

Bibliography

Appendix

Abstract

Fog can offer protection for intertidal organisms around the coast of Washington when low tides occur coincidentally with warm temperatures. This protection becomes more important as extreme heatwaves are expected to increase with global warming. Using Cloud Top Height products from NASA and NOAA's GOES-17 satellite, we created frequency maps and timeseries of fog and low clouds cover (FLCC) during the summer months of 2022 over Washington's coastal areas. Our results showed that FLCC increased significantly in both coastal and ocean locations from May to August. On San Juan Island, FLCC was least frequent in the Northeast area, which contained the town of Friday Harbor. Intertidal organisms around False Bay and Cattle Point had protection from FLCC against heat stress during around 50% of the cumulative midday low tide hours in summer 2022, while it was only around 14% for the ones around the UW Friday Harbor Lab. We also used photos from our field cameras to acquire more accurate fog presence at local sites on the island. We trained an SVC Machine Learning model to do binary classification on all the photos, and the results were compared with the FLCC presence results that we inferred from satellite data. Comparisons suggested that we have yet to be able to use the satellite's Cloud Top Height products to make conclusions about FLCC presence in a small-scale area at a specific timestamp, but we can use them to make good estimation of FLCC patterns over time at large spatial scales from the products.

1. Introduction

One of the most common phenomena along the West Coast of the United States in the summer is fog. Fog, or very low clouds, are the cause of many marine hazards, but they also have quite a few important roles. They are the reason why coastal redwoods, pines, salamanders, among others, can survive the dry season, and the factor that makes a lot of coastal areas much cooler than average (Dawson T.E. 1998). We believe that fog can also offer protection for intertidal organisms by blocking sunlight and providing humidity. As climate changes, people are concerned about fog declining. A lot of studies have investigated fog patterns in California, but little has been known about fog in Washington. However, the importance of coastal fog to intertidal organisms may be more profound in Washington, as the intertidal organisms there may suffer more damage from the warming climate than their counterparts in the southern latitudes. This is because during the summer months in Washington, the tides are usually lowest during midday, exposing the organisms to the hottest air and sunlight of the day, whereas during midday in California, the tides are usually often high enough to submerge the organisms in water, providing significant protection against heat stress for them (Helmuth et al 2002).

To fill in that knowledge gap, we set out to study fog and low clouds cover (FLCC) patterns around the Washington coast. To look at them from top down, we used the Cloud Top Height imagery from the GOES-17 satellite, which captured the coastal area of Washington state at 10 km spatial resolution and 5-minute temporal resolution. To study them from the ground, we set up cameras as well as temperature and humidity sensors at multiple sites on San Juan Islands to detect microscale fog and analyze its

effect on the local ecosystems. Our main research goals are: (1) finding the frequency patterns of summer FLCC around the Washington coast using GOES-17 Cloud Top Height products, (2) comparing the satellite data with the field data to see how we can best interpret them, and (3) quantifying the impact of FLCC on intertidal organisms in Washington, in terms of the number of midday low tide hours with FLCC protection and of the temperature difference created by FLCC presence.

Section 2 provides background information on fog from previous studies. Section 3 describes the methods we use to obtain and analyze our data from the satellite, the field, and other sources. Section 4 presents the results we get. Section 5 further discusses a few implications of some results, section 6 concludes this part of a hopefully ongoing research topic, and section 7 talks about a few future steps.

2. Background

2.1) Definition

As defined by the American Meteorological Society

(<https://glossary.ametsoc.org/wiki/Fog>), fog is "Water droplets suspended in the atmosphere in the vicinity [of] the earth's surface that affect visibility." Simply put, fog is a cloud very near the ground that reduces visibility. In our study, we define fog in three ways: (1) in satellite images, fog is the clouds lower than a certain height from the ground, (2) in field camera photos, it is fog when objects which are often visible are now obscured behind a white or grey shield, and (3) from sensor data, it is fog when the air

temperature and dewpoint match. Fog appears when the temperature and dewpoint of the air at a location become identical, i.e., when relative humidity reaches 100%, and fog dissipates when air temperature rises above the dewpoint temperatures, i.e., relative humidity decreases.

2.2) How fog protects intertidal marine organisms

Because its water droplets reflect light, fog can shield everything inside and underneath it from a lot of solar radiation, the primary heat source. High humidity also keeps organisms from losing too much water.

Organisms in the rocky intertidal zones are assumed to live very close to their thermal limit (Helmuth et al 2002, [Fields et al 1993](#), [Glynn 1991](#)) . Therefore, the warming climate can be unbearable for them, especially when the tides are low enough to expose the organisms to the air. Studies have found that intertidal organisms' body temperatures are often much higher during aerial exposure than during water immersion, and that those in temperate regions like Washington suffer thermal damage of proteins almost exclusively during aerial exposure (Helmuth et al 2002). In the summer of 2001, along the U.S. West Coast, sites in Washington had the highest numbers of hours when the lowest tides occurred during midday. The intertidal organisms exposed to direct sunlight during those hours were therefore at highest risk of heat damage (Helmuth et al 2002). Our calculation from NOAA's tide height prediction for summer 2022 also showed that the tide around San Juan Island in Washington was in the lower half of its maximum range during over 97% of the midday hours (refer to our section 4.4), which indicates that

the trend of low tide during midday in the summer in Washington that Helmuth saw in summer 2001 is still true today.

2.3) Other ecological roles of fog

Fog helps California coastal redwood conserve water, especially during the dry summer season in California, by reducing their transpiration rates (Johnstone and Dawson 2010). In fact, California coastal redwoods are only present in places with frequent coastal fog. Summer fog frequency had significant positive correlations with endemic coastal California pines' annual growth. (Johnstone and Dawson 2010 *) Water from fog that drips off the trees is also an important water source for other species, from understory plants to salamanders, from lichens to salmon (Dawson 1998, Branch et al 2022).

2.4) Fog frequency trend

While Johnson and Dawson's statistical calculations indicated a very slight decreasing trend in fog frequency over the previous 60 years, they also inferred a 33% reduction in fog frequency since the early 20th century (Johnson and Dawson 2010), and newspaper have been writing about people claiming that fog is definitely much less frequent than it used to be (Branch et al 2022).

3. Methods

3.1) Satellite data processing

3.1.1) Download raster data

We downloaded the Cloud Top Height product from GOES-17, a satellite operated by NOAA and NASA, which takes images of the western United States every 5 minutes every day. We ran the download-goes Python script in Steven Pestana's goes-ortho repository (github.com/spestana/goes-ortho) (Mello and Pestana 2022) to download the raster files of all images taken between May and September of 2022—the months of summer or near-summer weather in Washington—from NOAA's GOES repository on Amazon Web Services (registry.opendata.aws/noaa-goes/). A raster file is a file that represents a two-dimensional picture as a grid of pixels. Though an image appears to be 2-dimensional, they also usually have multiple layers. These layers are commonly referred to as bands or channels, and they represent the observations of the reflectance or emittance under different wavelengths of light. Each pixel has one or more numbers associated with each layer that the pixel has. Each number typically ranges from 0 to 255, where "0" represents black (no light) and "255" represents very bright light, in the wavelength of light associated with each band.

In all files, we downloaded the area of the Washington coast -- between 46- and 49-degrees latitude (North), and -125 to -122 longitude (West).

Each pixel in a Cloud Top Height image has information about the height of the cloud top (HT) in the area that the pixel represents. Specifically, the value of the HT variable of

each pixel indicates the highest altitude of the visible portion of the cloud seen from top down in meters. Figure 1 is the plot of the HT variable of one Cloud Top Height image, after it has been reprojected and orthorectified, which is explained in the following section.

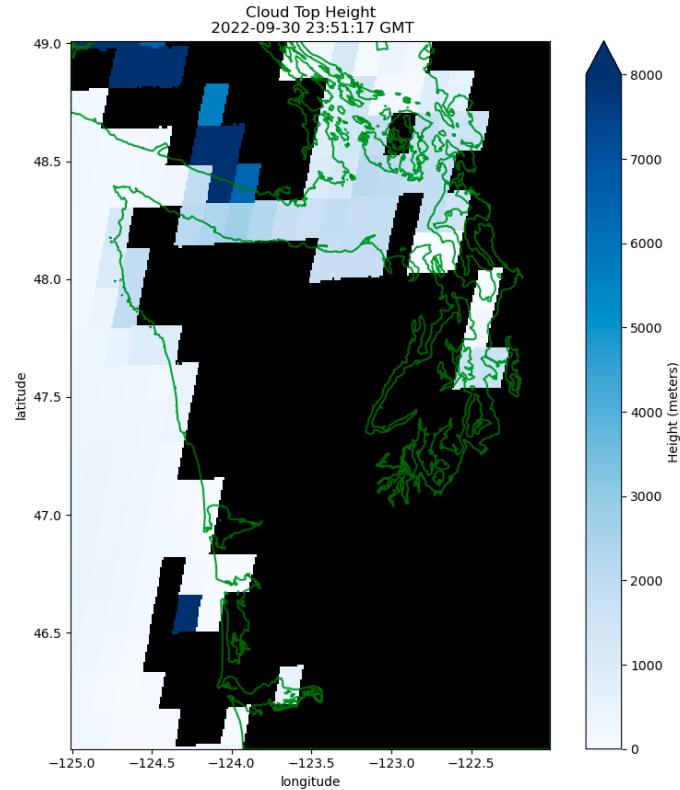


Figure 1. Plot of a reprojected Cloud Top Height images with shorelines added. The color bar for the Height variable on the right indicates that the darker the blue hue that a pixel has, the higher the cloud top in the area represented by that pixel is. In contrast, the lighter the blue hue that a pixel has, the lower the cloud top in that area is. Black indicates no cloud detected in that pixel.

It is important to note that when a high cloud is detected for one pixel, it does not necessarily mean that there are no lower clouds beneath it.

This Cloud Height algorithm first derives data from the Cloud Mask algorithm, which determines which pixel is cloudy and which pixel is clear sky. Then, the Cloud Height algorithm will process only the cloudy pixels (Heidinger et al 2020). The clear sky pixels will have NaN (not a number) values, and they are represented as black pixels in the figure above. There may be pixels that are determined as cloudy by the Cloud Mask algorithm, but when the Cloud Height algorithm processes them, no clouds are detected, and so those pixels have a value of 0.

To confirm our understanding of this algorithm, we compared a Cloud Top Height image with other images taken in different bands, or different wavelengths of light, of the same area at the same time. We also checked the Data Quality Flags of our Cloud Top Height image. Details about those are in Appendix C.

The cloud height product is based on the thermal bands in GOES, which are sampled at 2 km resolution at the equator and near 5 km resolution in Washington State. The cloud height product is produced at 10 km resolution, but in the process of reprojecting it, we downsampled to 1 km resolution to better resolve the effects of topography. In each 10-kilometer-squared area, some parts can have very clear sky, and some can have some clouds, which makes it very difficult to use one number to represent the cloud top height of the whole area. Moreover, the error range for the reported height in a pixel in a Cloud Top Height raster is around 1-2 km, so a pixel with a value of 0 may have clouds that are

somewhere between 0-2 km high. These limitations were considered when we chose the cloud top height threshold for fog and low clouds cover.

3.1.2) Reproject, parse datetime, clean

We ran the goes-ortho Python script in Steven Pestana's goes-ortho repository (<https://github.com/spestana/goes-ortho>) to reproject the images from ABI Fixed Grid coordinates to latitude and longitude coordinates. The goes-ortho code also orthorectified the images, which means that the images were adjusted to represent the correct terrain. The motivation for orthorectification came from the fact that terrain in the images was distorted, as GOES-17, viewing the Earth from right above the equator, took images of Washington state from a very strong (Pestana, Lundquist 2022).

We extracted datetime information from each raster file name and added them as a variable to the dataset of the image file. We also removed all variables and dimensions but the Cloud Top Height variable in each raster file, thereby reducing the size, or disk usage, of the folder of 40785 files from 63GB to 14GB.

3.1.3) Load, stack, chunk

To work with all the data at once, we need to load them from our computer's storage to the computer's memory. Because over 40,000 images are too much for one computer to hold in memory, we utilized Dask, a Python library for parallel computing. Dask can divide a dataset into many smaller pieces, called *chunks*; each chunk is small enough to

load into memory. Dask increases efficiency because multiple computer cores can work on multiple chunks at the same time.

We chunked the dataset in a way that made it easy to grab the time series data of a single point on the map. We stored the chunked data in Zarr, a file storage format. The essence is that while the original dataset is made up of many raster files, each file representing an image from a single timestamp, the Zarr dataset is a file that stores data in time series, each timeseries containing all the values of each pixel, or each small group of pixels, at all timestamps that the raster files originally come in. That Zarr file only took up 1.9GB of our disk usage when storing all the data that took up 14GB in the folder of raster files.

```

# Open all the files
ds = xr.open_mfdataset(nc_files, chunks={'time': 500})
✓ 9m 12.9s

ds['HT']
✓ 0.0s

xarray.DataArray 'HT' (time: 40785, latitude: 284, longitude: 284)

Array      Chunk
Bytes      12.25 GiB   315.06 kiB
Shape      (40785, 284, 284) (1, 284, 284)
Dask graph 40785 chunks in 81571 graph layers
Data type  float32 numpy.ndarray

Coordinates:
longitude  (longitude) float64 -125.0 -125.0 ... -122.0 -122.0
array([-125.001936, -124.991363, -124.980789, ..., -122.030782, -122.020209,
       -122.009635])

latitude  (latitude) float64 49.0 48.99 48.98 ... 46.02 46.01
array([49.004789, 48.994216, 48.983642, ..., 46.033636, 46.023062, 46.012489])

time      (time) datetime64[ns] 2022-05-01T00:01:17 ... 2022-09-...
array(['2022-05-01T00:01:17.000000000', '2022-05-01T00:06:17.000000000',
       '2022-05-01T00:11:17.000000000', ..., '2022-09-30T23:41:17.000000000',
       '2022-09-30T23:46:17.000000000', '2022-09-30T23:51:17.000000000'],
      dtype='datetime64[ns]')

```

Figure 2a. A visual representation of how 40785 raster image files, each of which has 284 pixels along the longitude and latitude dimension, are stacked on top of each other.

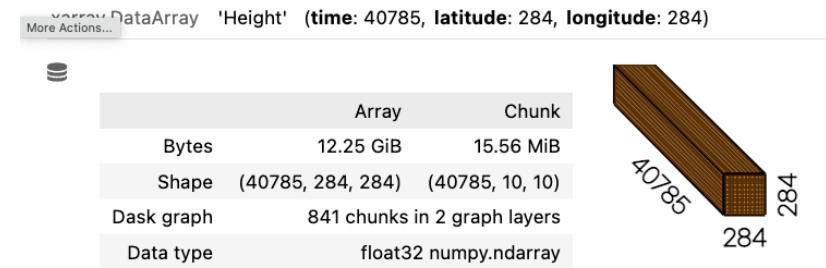


Figure 2b. A visual representation of how the stack is chunked along the time dimension so that each chunk still has 40785 pixels along the time dimensions, but only 10 pixels along the longitude and latitude dimensions.

3.1.4) Calculate and plot fog and low clouds cover frequency

First, we had to choose a height threshold for FLCC. If a region, or a pixel on the GOES-17 imagery, had a cloud top height detected below that threshold, we would say that there was FLCC detected in that pixel. The typical inversion height, which is the height where fog is usually trapped below, that was found in the northern California coast is 400 meters (Johnstone and Dawson 2002). We chose 2000 meters, rather than 400, to be our initial threshold, because GOES Cloud Top Height algorithm has an error range of 1-2 kilometers (Table 6, p.64, Heidinger et al 2020), and we didn't want to miss any fog event. Moreover, 2000 meters is also the upper threshold for the category of low clouds in temperate regions as defined by NOAA (<https://www.noaa.gov/jetstream/clouds/four-core-types-of-clouds>).

At each pixel, if the detected cloud height value is below that threshold, we classify it as having fog and low clouds cover (FLCC). If the height value is above the threshold, we classify it as having high clouds. If the height value is NaN, we classified it as having no clouds. Then, across the 40785 timestamps, we calculated the percentage of time that FLCC were detected at each pixel and created a fog and low clouds cover frequency plot from those percentage values. Because a pixel with high cloud top height detected can still possibly have lower clouds or fog beneath, this frequency map is the frequency of FLCC without higher clouds.

3.2) Field sensor data processing

3.2.1) Field cameras and sensors

We deployed Wingscapes timelapse cameras around the island. During May-September 2022, the cameras were set to take pictures every 30 minutes from 6am to 7:30pm. In this study we specifically looked at four cameras at these locations:

1. Friday Harbor Lab, on NOAA Weather Station facing east.
2. Cattle Point, on a tree facing east.
3. False Bay, on a tree facing southwest.
4. Mount Dallas, on a dead tree facing east.



Figure 3. Locations of cameras on San Juan Island.

We also deployed HOBO Pro v2 sensors at various locations, including at these four, to measure temperature and humidity at the same locations with the cameras.

3.2.2) Classify field photos

Over two thousand photos have been taken at each site during summer 2022, and more are still being taken every 30 minutes until today. To know how often and what time fog was present at each site over time without having to go through each photo manually, we trained a Computer Vision model to do the image classification task. We chose Support Vector Machine because it is easy to train and it can handle high-dimensional data, such as our photos, which are three dimensional arrays of numbers. Moreover, compared with other popular algorithms for image classification such as neural networks, SVM is generally less prone to overfitting, more interpretable, and more computationally efficient for small datasets like ours.

First, we set the criteria for humans to label a photo, or an instance, as having fog or not. A photo is classified as having fog if it meets one or more of the criteria in Table 1 below.

Table 1. Criteria for labeling fog presence in camera photos

Location	Criteria	Examples of no fog	Examples of fog
Mount Dallas	<p>1) The whole image is blurry in white or grey.</p> <p>2) The mountain range and the water on the right are not visible.</p> <p>3) There's a lot of white patches on the ground.</p>	  	 
False Bay	<p>1) The whole image is blurry in white or grey.</p> <p>2) The range of trees on the left is not visible.</p>	 	 

Cattle Point	1) The whole image is blurry in white or grey.		
	2) The island across the water is not visible, either partly or completely.		
Friday Harbor	1) The whole image is blurry in white or grey		
	2) The stretch of land across the water is not visible, either partly or completely.		

We manually labeled around a subset of photos to use as training data for each site. First, we scrolled through all the photos and picked the ones that looked like there was fog to look at more closely and label. After we had got a set of fog photos that were diverse

enough to represent what fog looked like at different times and in different forms at that site, we scrolled through the remaining photos and picked the ones that looked like there was no fog to look closely and label. We tried to have a similar number of photos between the two classes so that our model wouldn't bias against either class. This was not easy, however, as there were very few foggy times at Friday Harbor Lab, which made it much harder to find enough fog photos for the training set.

The process of classifying photos as having fog or not having fog for training was very qualitative. There was a small fraction of photos that we were not confident if our classifications were correct or not. We showed a few of which in Figure 4 below.



Figure 4. Confusing photos. a is more blurry than other photos around that timestamp. b seems to have a few white stretches at the horizon. c seems to have fog out of the bay, but the trees were still visible. d has blurry spots which we were not sure if they were signs of fog or just the water condensation on the camera lens.

For the photos that are too confusing for us to determine like these four photos above, we didn't manually label them but let the model determined. The Computer Vision model can sometimes detect the similarities or differences between those confusing photos and other photos that humans don't see right away.

For example, for this confusing picture at Cattle Point when it looked like there was fog approaching from far away, but it was still clear enough at the current location for us to see the island across the water, the model was also very uncertain, but it eventually labeled the photo as having no fog, as shown in Figure 5. Our hope is that the model can be more consistent in deciding the labels of confusing photos than humans are.



Figure 5. The model’s probabilities of this instance belonging to either class were very close, which indicated that it was uncertain about its final prediction of fog presence in this photo.

As we scrolled through the photos to label a subset for training, we also manually deleted many photos that were totally black, because they were taken before the sun rose earlier in the summer or after the sun set later in the summer. We couldn’t conclude about fog presence in those pictures, and we didn’t want them to affect the accuracy result of the model.

There were more photos in our training set for Mount Dallas and Cattle Point than for the other two because the photos at these sites were more variable. Fog there could have many different looks, because of the landscape, seascape, and sunlight in relation to the angle and altitude of camera. Therefore, when we had the same number of photos in the training set for Mount Dallas like for False Bay, the cross-validation accuracy was very low. We then tried to include pictures at all times of the day and in many different weather conditions in the training set. We also tried to identify the photos that the model classified wrong and add them to our training set.

Table 2. Dataset size, and model's parameters and accuracy of each site

Site	Number of instances in the training set, and in total	Modell parameters	Cross-validation accuracy
False Bay	21, 2192	C=1, gamma=1, kernel='linear'	100%
Mount Dallas	72, 2043	C=1000, gamma=0.0001, kernel='rbf'	84%
Cattle Point	73, 2429	C=5, gamma=0.0001, kernel='rbf'	94%
Friday Harbor	34, 2469	C=5, gamma=0.001, kernel='rbf'	82%

We ran the models on all the photos at each respective site. We then added the predicted labels of the photos to a data frame consisting of the photo files' name, path, datetime information, which was parsed from their metadata. This data frame after being sorted in

a chronological order could be used to create timeseries plots of fog presence at each location.

Even though we still had to manually classify many photos to use for training and validate some of the predicted labels, the benefits of these Machine Learning models are that we can now use them on a lot more photos at these sites, and that we can have a very neat table with all the resulting labels organized in chronological order.

3.3) Calculate midday low tide hours

We downloaded hourly tide data from May 1st to September 30th 2022 from NOAA (<https://tidesandcurrents.noaa.gov/waterlevels.html?id=9449880>, accessed on June 22 2023). We used the predicted tidal height data, because the verified tidal height data were missing values in some periods, and because the predicted height and the verified height values were not very different when we plotted them together. We calculated the maximum tidal range by subtracting the lowest predicted tidal height from the highest. After extracting only the midday hours, we categorized them in tidal height groups to see, of all midday hours, which proportion there were low tides and which proportion there were high tides.

Because NOAA's tidal heights were predicted for every hour, while GOES-17 Cloud Top Height was measured every 5 minutes, we did linear interpolation on them to count the number of midday low tide hours when there was FLCC present.

4. Results

4.1) Overall frequency map from satellite data

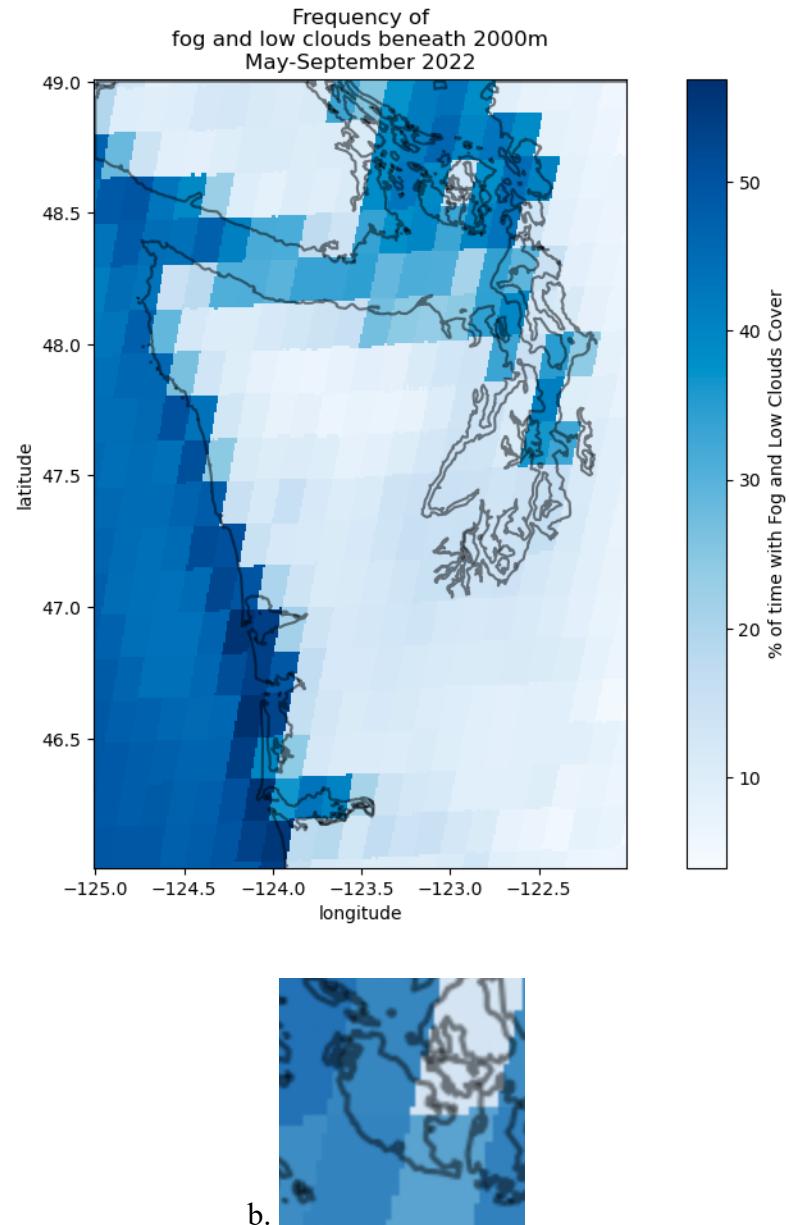


Figure 6. Frequency of FLCC below 2000m from May through September 2022 around

(a) the Washington coast and (b) San Juan Island.

FLCC frequency is the highest along the coast and is much higher in ocean areas than inland. Coastal locations around latitude 46.3 to 47, the southernmost coast of Washington, have the highest FLCC frequency in the state – around 56% of the time. Around the San Juan Islands, the area containing Shaw Island, Orcas Island, and the Friday Harbor area of San Juan Island, has the least FLCC frequency. On San Juan Island, the Southwest part is shown to have the most frequent FLCC – around 40% of the time.

4.2) Local fog frequency from satellite data and from camera photos

The FLCC frequency value of a location from the satellite is for the whole satellite pixel where that location point is in. Because FLCC are usually localized, there may be difference in FLCC frequency between the point-scaled individual location and the whole area of 10x10 km (downscaled to 1x1 km) that each satellite pixel represents. On the other hand, the frequency value from the camera photos classified as having fog should be the accurate fog frequency value of that place. We chose two locations, False Bay and Mount Dallas, to investigate those frequency values in Table 3.

Table 3. Percentage of the time there was local fog or FLCC at False Bay and Mount Dallas.

	Fog seen in camera photos	FLCC detected in satellite imagery, when the cloud top height threshold for FLCC being below 600 meters	FLCC detected in satellite imagery, when the cloud top height threshold for FLCC being below 2000 meters
False Bay looking Southwest, from May 1st to July 18 th 2022 ¹ , 6am-8pm everyday	3%	4.7%	23.9%
Mount Dallas looking East, from July 19 th to September 30 2022, 6am-8pm everyday	30%	26.9%	55.4%

¹ The reason why none of our camera records at these locations both started on May 1st and ended on September 30th, which is the summer period we are studying, is because of logistic difficulties. Our camera at False Bay broke on July 18th 2022, so we didn't have record there for the rest of that summer. We deployed the camera at Mount Dallas on July 19th, so we didn't have record there earlier that summer.

4.3) FLCC frequency by months

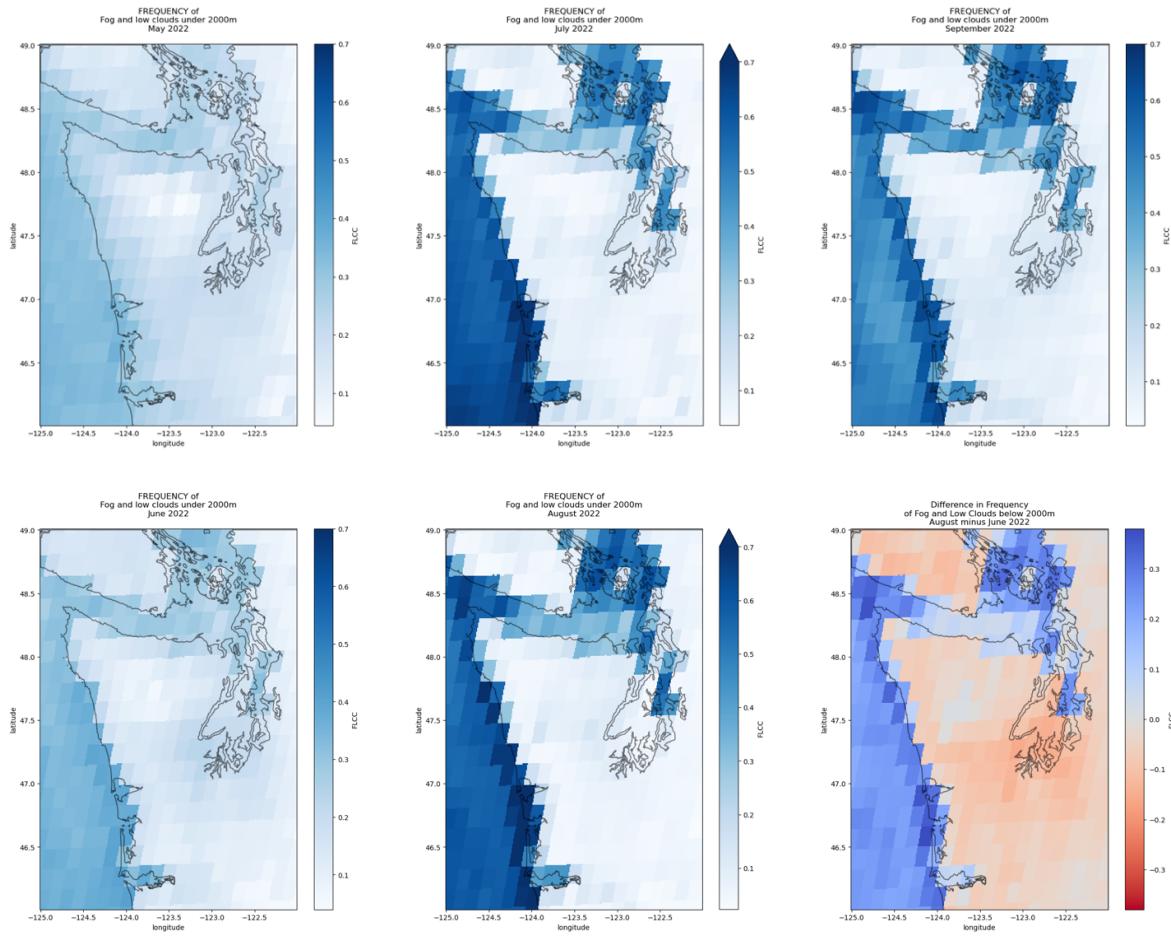


Figure 7. Frequency of FLCC under 2000 meters by each summer month during May–September 2022, and the difference map in frequency of June and August, around the Washington coast.

Continuously from May to August, the FLCC frequency increased significantly in coastal and ocean locations and decreased slightly in inland locations. These two trends peaked in August and reversed in September. The difference map between June and August showed the FLCC frequency contrast between ocean and island locations — ocean

locations have more frequent FLCC in August than in June, while inland locations have more frequent FLCC in June than in August.

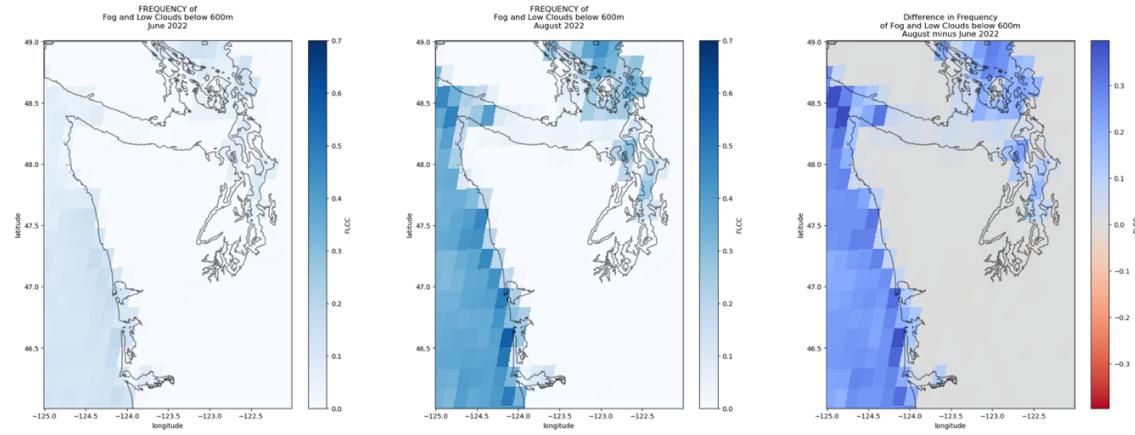


Figure 9. Frequency of FLCC under 2000 meter around the Washington coast by each month of June and August 2022, and the difference map between the two.

When the threshold for FLCC was set to be 600 meters, there was much less FLCC overall, and almost no FLCC on land, but there was still the overall pattern of increasing fog later in the summer.

4.4) FLCC frequency by times of the day

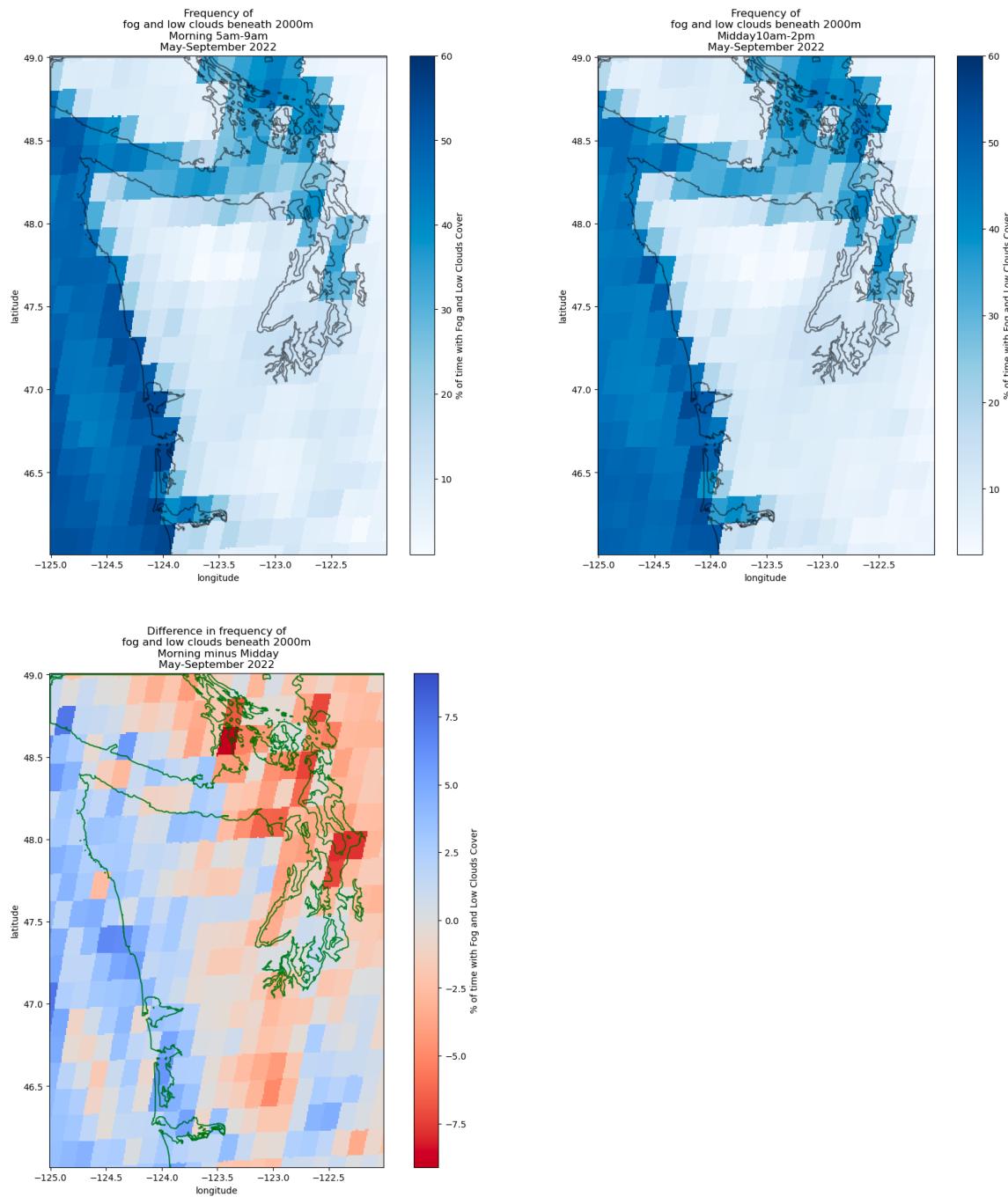


Figure 10. Frequency of FLCC under 2000 meter around the Washington coast by times of the day — the morning hours (5-9am) and midday hours (10-2pm) — and the difference map between the two.

The ocean locations along the coast to the south of Washington have more FLCC in the morning than in the midday hours, and the reverse for the locations around the San Juan Islands, Vancouver Island, and Seattle have more FLCC. Note that midday FLCC is still the most frequent along south coast of Washington.

We have not calculated the exact numbers, but while investigating fog presence through camera photos, we noticed that there was most frequent fog presence in the early mornings throughout the summer. This observation agreed with what the frequency maps above indicated – most of the pixels were lighter in the midday map, meaning less frequent FLCC. This phenomenon is true for a lot of places around the world, as early morning is the time when the air temperature drops to the dew point, leading to the formation of fog.

4.5) Comparing satellite-detected cloud height and camera-derived fog presence

We wanted to see how much our interpretation of FLCC from the satellite-detected cloud height and the camera-derived fog presence agreed with each other. For example, at some timestamp when there was fog presence seen in the camera photos of a location, we wanted to know if our interpretation of FLCC from the satellite imagery at that timestamp – like there was FLCC if the detected Cloud Top Height for the pixel containing that

location was below 2000 meters, and there wasn't FLCC otherwise – indicated the same thing.

We investigated the period of late June because our air temperatures sensor indicated that that period was warmer than normal, which meant that FLCC presence was important.

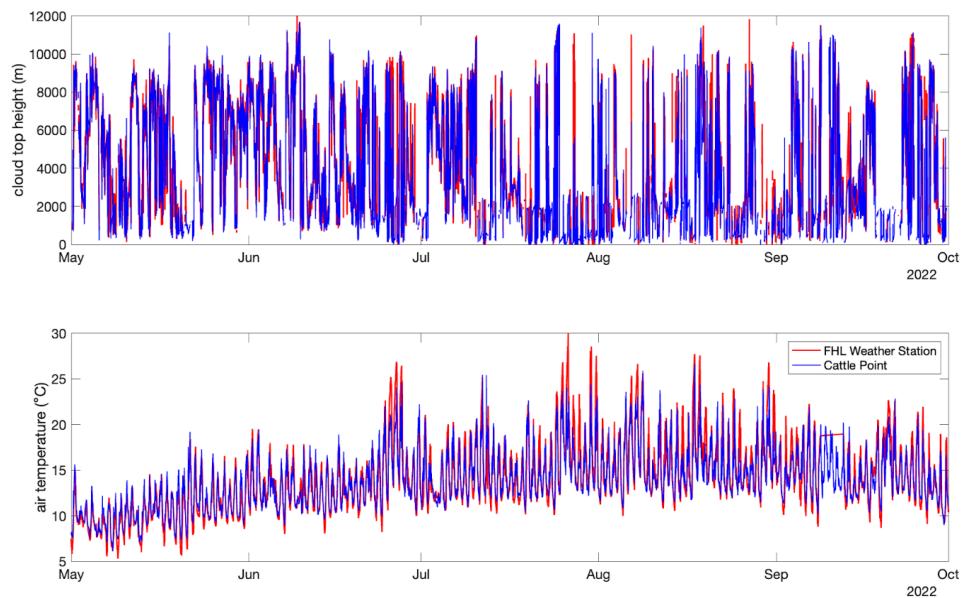


Figure 11. Cloud top heights from GOES-17 and air temperatures from field sensors through the summer at Friday Harbor Weather Station and Cattle Point.

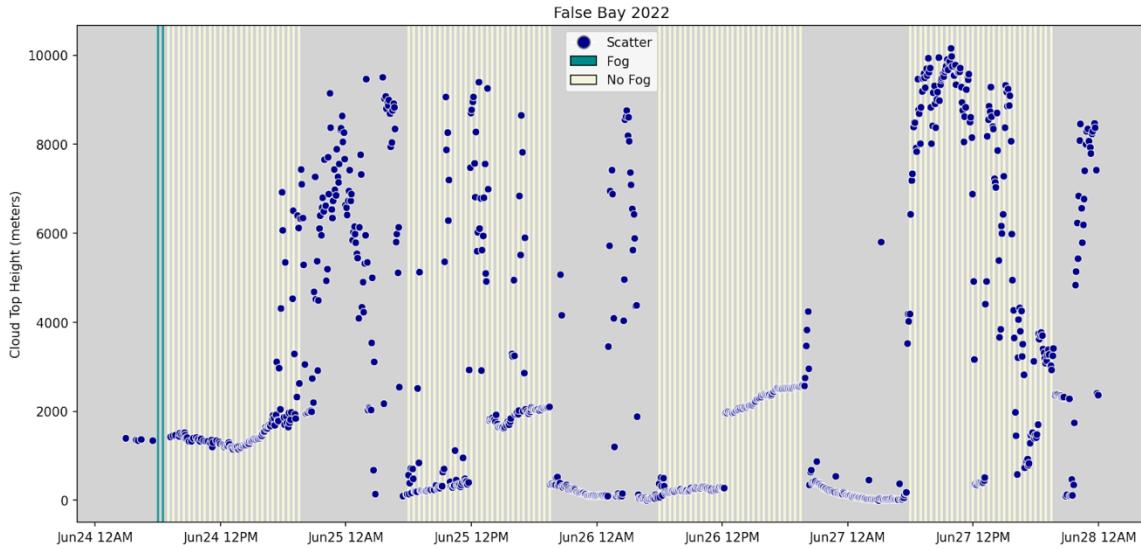


Figure 12. A plot showing cloud top height values from the satellite and fog presence from the field camera on top of each other for comparison. The purple scatter dots indicate the cloud top height values for the satellite pixel containing False Bay. The green vertical lines indicate that the photos taken at the site during those times are classified as having fog by the Machine Learning model, and the yellow vertical lines indicate no fog. This plot is for the late June period of 2022 at False Bay.

Though there were a lot of cloud top height detected below 2000 meters, there were barely any times with fog detected by the computer vision model on the photos. Manually

checking the photos, we saw that the sky in the photos were mostly clear during the times that the cloud height detected was below 2000 meters, like during the day of June 24.



On the other hand, when the cloud height detected was high, around the range of 6000-10000m, there were clouds seen in the pictures, like in the morning of June 27.



Similar inconsistencies happened in mid-July, when the temperature was also high. For example, in the morning of July 12th, there were no fog or any clouds seen in the camera photos after 7am, but the cloud top heights detected by GOES-17 for the area containing

False Bay were consistently around 500 meters, which were below the 2000-meter threshold for FLCC.

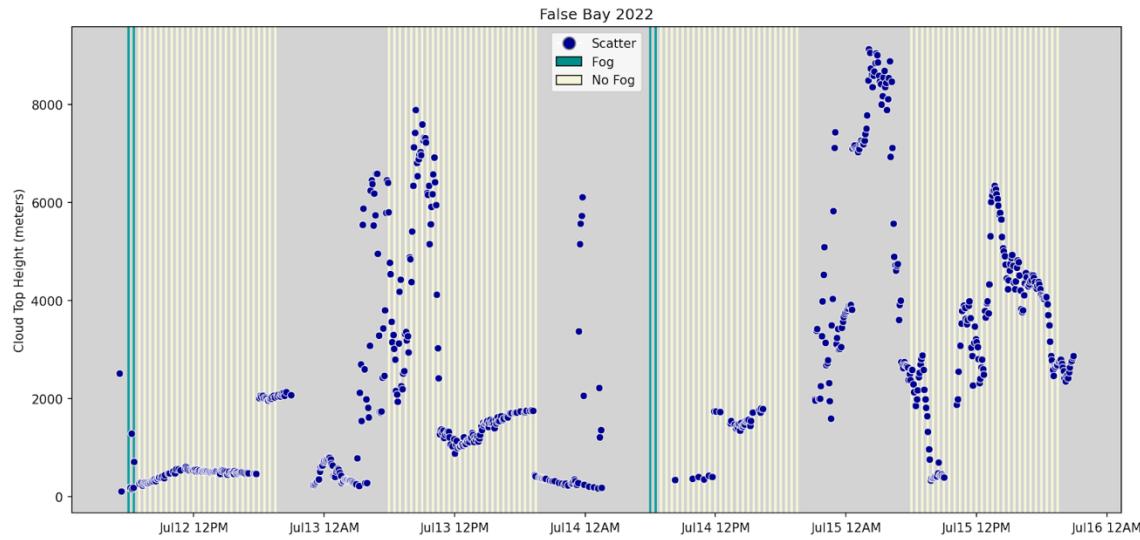


Figure 13. A plot showing cloud top height values from the satellite and fog presence from the field camera on top of each other for comparison. The purple scatter dots indicate the cloud top height values for the satellite pixel containing False Bay. The green vertical lines indicate that the photos taken at the site during those times are classified as having fog by the Machine Learning model, and the yellow vertical lines indicate no fog. This plot is for the mid-July period of 2022 at False Bay.



Figure 14. Camera photos at False Bay in the morning of July 12th 2022.

We chose a specific timestamp – noon local time – to look more closely into. We also referred to measurements from nearby weather stations as well as FLCC in nearby pixels to get more information about FLCC on that day, specifically at noon local time.

- Field photo: classified as no fog, because we could still see the left trees and the bay clearly. It looks like there might have been fog further out by the Olympic peninsula, but at False Bay it definitely was clear.



- GOES cloud top height value for the pixel containing False Bay was around 550 meters.
- GOES Cloud Top Height map on July 12 at 12pm Pacific Time:

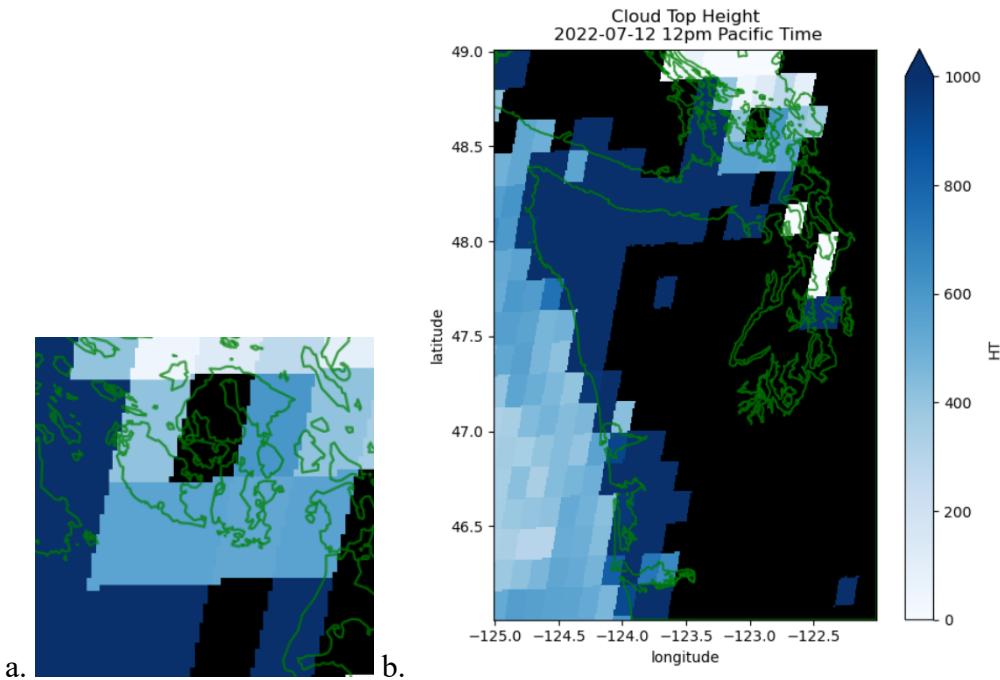


Figure 15. Cloud Top Height plot at 12pm July 12th 2022 around (a) the San Juan Islands and (b) around the coast of the whole Washington state.

The whole region south of San Juan Island has clouds around 500 meters top height detected.

- Other pictures around this time that day are very similar (and the picture at the same time in FHL is also a very clear day), yet in satellite height, some are NaNs and some are 500-600 meters.

- Friday Harbor Airport's weather_summary_set_1d is clear, cloud_layer_1_code_set_1 is 1 = clear. Friday Harbor field photo also shows very clear sky.
- Whidbey Island Naval Air Station (48.35525, -122.66352): cloud_layer_1 is 106 meaning thin scattered cloud. Cloud base height is 304 meters. Visibility is 10 miles.

4.6) FLCC-protected hours during midday low tide around San Juan Island

FLCC is believed to be most important to intertidal organisms in Washington state during the hours of low tide, which are usually around the middle of the day during the summer. In this study, we chose the midday hours to be from 11:00 (11am) through 13:00 (1pm) to be consistent with the hours studied by Helmuth et al (2002). We also categorized our tidal range into the lowest 25%, 35%, and 50% like they did. Figure 16 was from Helmuth's study, and Figure 17 is where we showed the proportion of midday low tide hours at specific locations when there were FLCC present to protect the intertidal organisms.

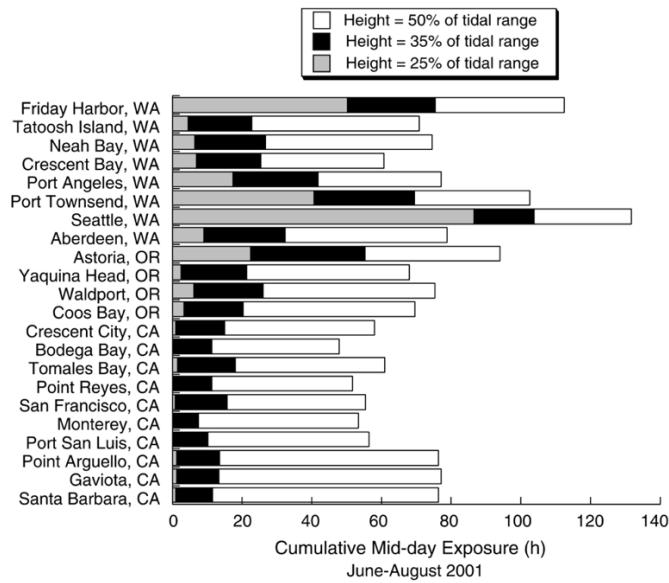


Figure 16. Number of midday hours, categorized by predicted tidal heights, during which intertidal organisms were exposed, at different sites along the West Coast of the US (Helmuth 2002).

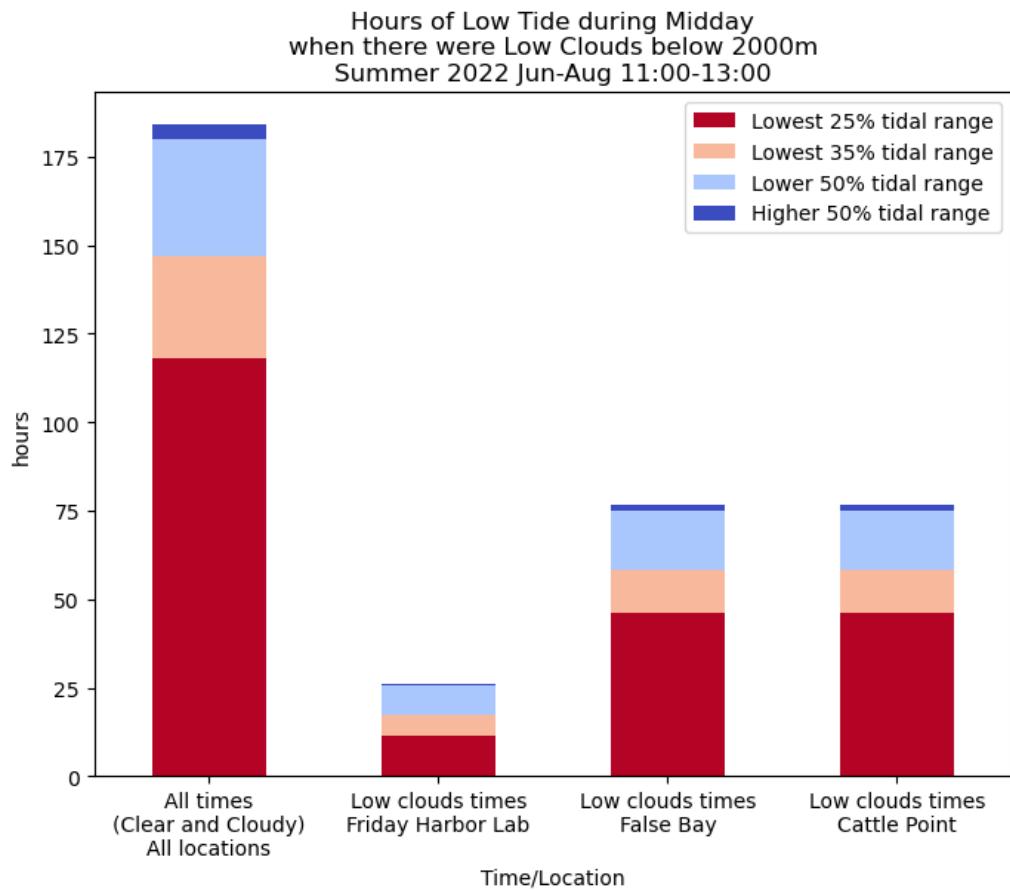


Figure 17. Cumulative midday hours categorized by the height of the tide. Low clouds times were derived from the satellite, and they implied FLCC presence.

Friday Harbor Lab has FLCC beneath 2000 meters in around 14% of the midday low tide hours. False Bay and Cattle Point have more than double the hours of FLCC during midday low tide compared to Friday Harbor. During those times, the intertidal organisms here had the protection against midday heat and direct sunlight.

4.7) Impact of FLCC on temperature

Figure 18 showed a rise in the difference between air temperature at Friday Harbor Lab and Cattle Point towards the end of July 2022, which was also when the cloud top heights detected by the satellite were consistently lower than the previous period. We investigated this more closely in the histograms in figure _ to quantify the impact of FLCC on air temperature.

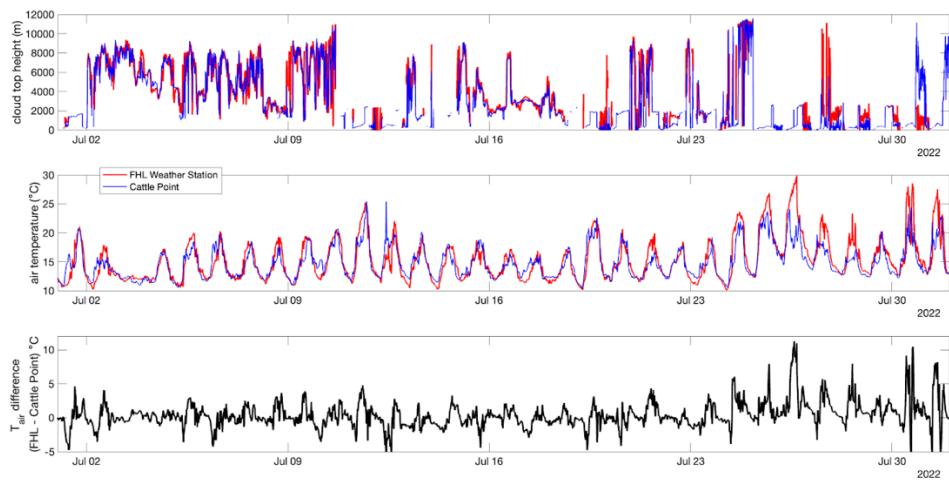


Figure 18. Cloud top heights from GOES-17, air temperatures from field sensors, and difference between temperatures at Friday Harbor Weather Station and Cattle Point during July 2022.

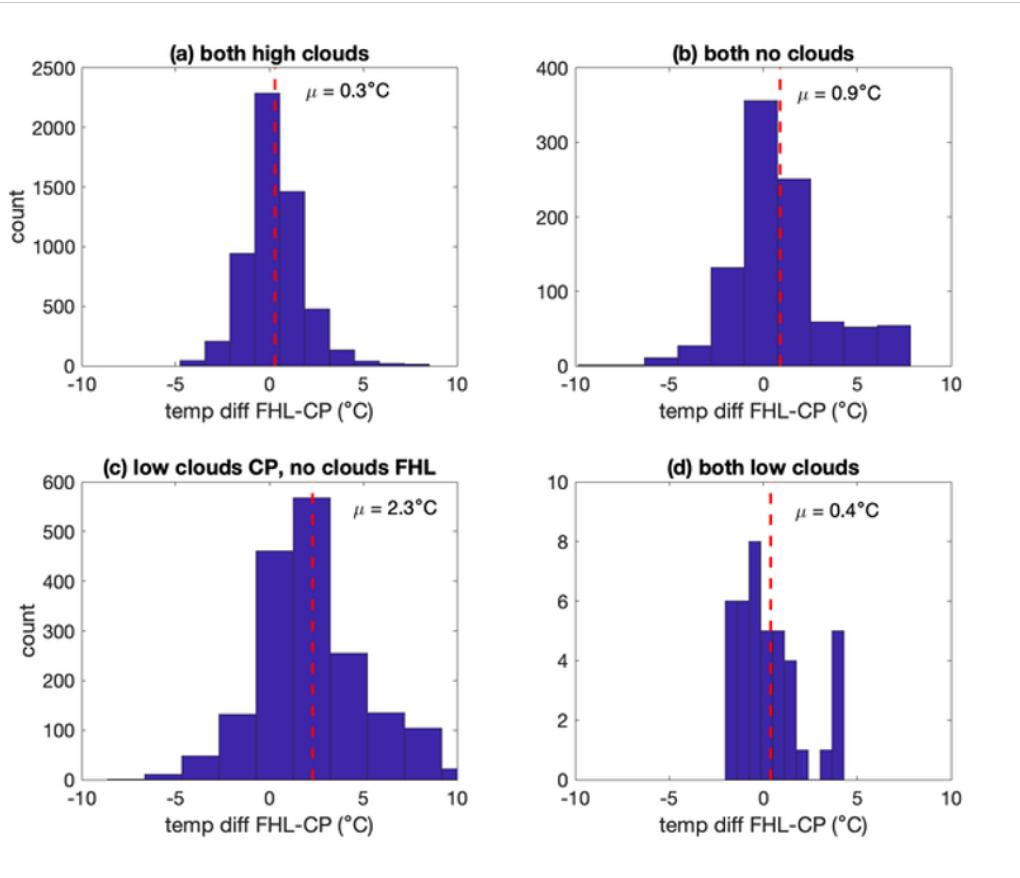


Figure 19. Difference between air temperature at Friday Harbor Lab and at Cattle Point during May-September 2022.

Cattle Point was almost always a little cooler than Friday Harbor Lab, but the difference was most pronounced when Cattle Point had FLCC and Friday Harbor didn't. During those times, Cattle Point was cooler than Friday Harbor Lab by 2.3 Celsius degrees on average. During the times that both locations had the same kind of cloud cover, or both had no clouds, the temperature difference between two locations was not as big.

4.8) Changes in cloud top height in first and second half of the summer

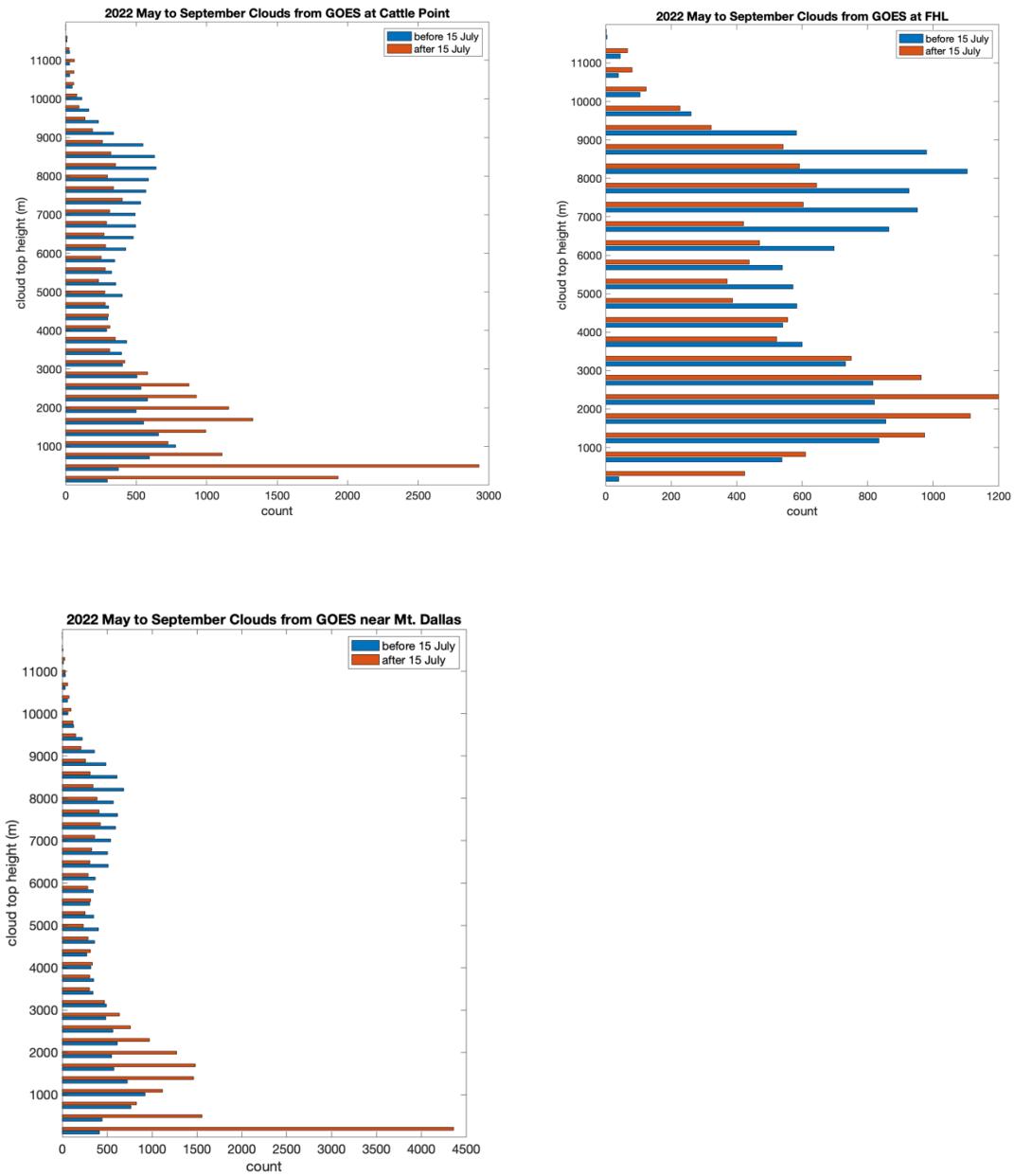


Figure 20. Changes in the frequency of cloud top heights detected by GOES-17, between the first and second half of the summer, in the pixels containing Cattle Point, False Bay, and Friday Harbor Lab.

There are many more low clouds after mid-July than before mid-July at around the three locations on San Juan Island. At around Friday Harbor, there was a balanced change between the frequency of high clouds around 7000-9000m and low clouds around 1000-3000m.

5. Discussions

In section 4.2, we looked at a few local fog frequency numbers that we calculated from satellite data and from camera photos. FLCC frequency over time from the satellite when the threshold is 600 meters matches more closely with the fog frequency from camera photos than when the threshold is 2000 meters. However, we did not do an exhaustive sensitivity test, and there may be other thresholds that are a better match. Moreover, the result that best matches with the fog frequency at a few point locations cannot guarantee that it will also best match the fog frequency at other places around the Washington coast. Further work is needed to determine the best threshold with confidence.

In section 4.5, we observed the inconsistency between what GOES-17 cloud top height data and the camera photos indicated about FLCC presence. There are a few possible explanations for these inconsistencies. First, there are likely to be differences between fog presence at point-scaled individual locations and FLCC seen for the whole area of 10x10 km that each satellite pixel represents. Second, this Cloud Top Height algorithm only outputs the value of the top of the highest cloud, even when there is FLCC beneath that cloud. Nonetheless, these findings did not necessarily mean that GOES-17 made faulty

detections of FLCC. There may very likely be low clouds or fog far away out of the range that the inland camera could capture. What the findings confirmed to us was that, if FLCC frequency is calculated by classifying FLCC presence for the areas with cloud heights below a certain threshold, then it may accurately reflect the macroscale pattern of FLCC over a period of time, but it cannot help us conclude FLCC presence at a point-location at a specific time.

6. Conclusions

This study has helped increase our knowledge of fog. From our analysis of Summer 2022 data, we can infer that in the summer in Washington state, fog increases significantly in ocean and coastal locations from May to August, and the reverse trend happens in inland areas. We now know that the location with most frequent summer FLCC in Washington in the south coast of the state. We also now know that the south coast has more frequent FLCC during morning hours than during midday hours, and vice versa for the San Juan Islands.

On San Juan Island, places along the Southwest coast, particularly Cattle Point and False Bay, are cooler than other parts of the island thanks to having frequent FLCC. Especially during the midday hours, frequent presence of FLCC means that there were less hours when the intertidal organisms at these sites suffered heat damage due to aerial exposure during low tide.

The techniques we developed to calculate the number of FLCC-protected hours during midday low tide at Cattle Point, False Bay, and Friday Harbor can also be used on other sites to find out the places that can become “climate refugia” for intertidal organisms and humans. Knowledge about “climate refugia” thanks to FLCC can be valuable in informing decisions about where to build marine conservation areas, shellfish farms, etc. The method of collecting camera photos of coastal areas and classifying the photos using a Computer Vision model can be applied to other locations to effectively acquire and process data about FLCC over time. The techniques we used to create FLCC frequency maps can also be used to create maps for other states. It will be valuable to compare the results with those of other techniques that people use to analyze FLCC, especially in states with abundant data and previous studies like California. In this study, through comparing with field cameras' photos, we know that, at least for locations like in Washington which were viewed by GOES-17 from strong angles, the current Cloud Top Height product's resolution is too coarse for inferring fog or low clouds presence at point-scaled locations at some specific time.

7. Future steps

To figure out how we can best derive fog presence and fog frequency from the cloud top height data, like which threshold to choose for example, we think it will be helpful to create the distribution plots of cloud top height values when there is fog seen and when there is no fog seen in the photos. We can classify more photos from other locations so that we can compare them and the cloud height in those areas. It would be good to try out

other Machine Learning classification models like tree-based models or neural network models and try to interpret the result from the model that we choose to use.

When we have camera photos from multiple summers, we can see if there is any change in summer fog frequency through the years. NOAA and NASA also have archived data of GOES-17 that we can use to look at changes in Cloud Top Height, hence FLCC frequency, through the years. Besides, we can also study the reasons behind the difference in FLCC frequency between the locations.

Acknowledgement

This research was supported by a grant from the National Science Foundation (DBI-2149705).

We are very grateful for the Friday Harbor Lab REU team for having made our summer research experience a very pleasant and memorable one.

We also want to thank Megan and David, George and Peggy, Scott and your family, Gary and April, Doug, and Peter and Lisa for letting us deploy our cameras and sensors on your properties, and Emily Carrington for your weather station data.

From Autumn: A special thank you to my family and friends who have shared with me your thoughts and questions about the study — when I practiced giving a presentation to you, when I shared with you the happiness of figuring out some code, or when I expressed to you how overwhelmed I was feeling about the research. Thank you to my sister, Yến Ngọc, for having drawn the graphics of the field camera icons for me to use in Figure 3.

Bibliography

Dawson, T. E. "Fog in the California Redwood Forest: Ecosystem Inputs and Use by Plants." *Oecologia* 117, no. 4 (December 22, 1998): 476–85.

<https://doi.org/10.1007/s004420050683>.

Helmuth, Brian, Christopher D. G. Harley, Patricia M. Halpin, Michael O'Donnell, Gretchen E. Hofmann, and Carol A. Blanchette. "Climate Change and Latitudinal Patterns of Intertidal Thermal Stress." *Science* 298, no. 5595 (November 2002): 1015–17. <https://doi.org/10.1126/science.1076814>.

Heidinger Andrew, Yue Li, and Steve Wanzong. "Enterprise AWG Cloud Height Algorithm (ACHA) - NOAA / NESDIS / Star Website." GOES-R SERIES DATA PRODUCTS, September 2020.

https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Enterprise/ATBD_Enterprise_Cloud_Height_v3.4_2020-09.pdf.

Horning, N. 2004. Understanding pixels, bands and channels, Version 1.0. American Museum of Natural History, Center for Biodiversity and Conservation. Available from <http://biodiversityinformatics.amnh.org>. (accessed on July 10 2023)

Johnstone, James A., and Todd E. Dawson. "Climatic Context and Ecological Implications of Summer Fog Decline in the Coast Redwood Region." *Proceedings of the National Academy of Sciences* 107, no. 10 (March 9, 2010): 4533–38.

<https://doi.org/10.1073/pnas.0915062107>.

Mello, P.A., Pestana, S., 2022. spestana/goes-py: v0.2.2 (0.2.2). Zenodo.

<https://doi.org/10.5281/zenodo.6455060>.

O'Brien, Travis Allen. "The Recent Past and Possible Future Decline of California Coastal Fog." Order No. 3471817, University of California, Santa Cruz, 2011.
<https://www.proquest.com/dissertations-theses/recent-past-possible-future-decline-california/docview/887106198/se-2>.

Pestana, Steven, and Jessica D. Lundquist. "Evaluating GOES-16 ABI Surface Brightness Temperature Observation Biases over the Central Sierra Nevada of California." *Remote Sensing of Environment* 281 (November 2022): 113221.
<https://doi.org/10.1016/j.rse.2022.113221>.

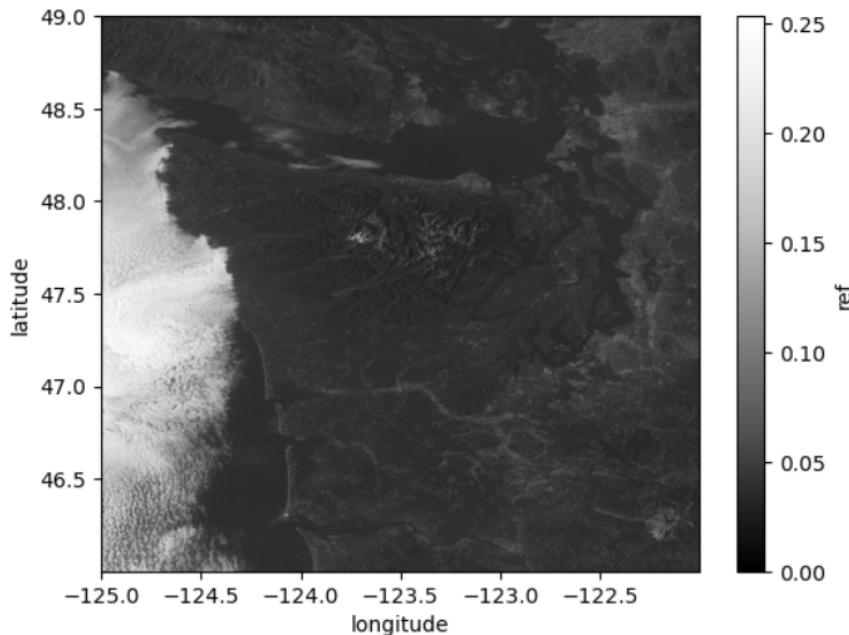
Torregrosa, Alicia, Cindy Combs, and Jeff Peters. "GOES-derived Fog and Low Cloud Indices for Coastal North and Central California Ecological Analyses." *Earth and Space Science* 3, no. 2 (February 2016): 46–67.
<https://doi.org/10.1002/2015EA000119>.

Appendix A: Code

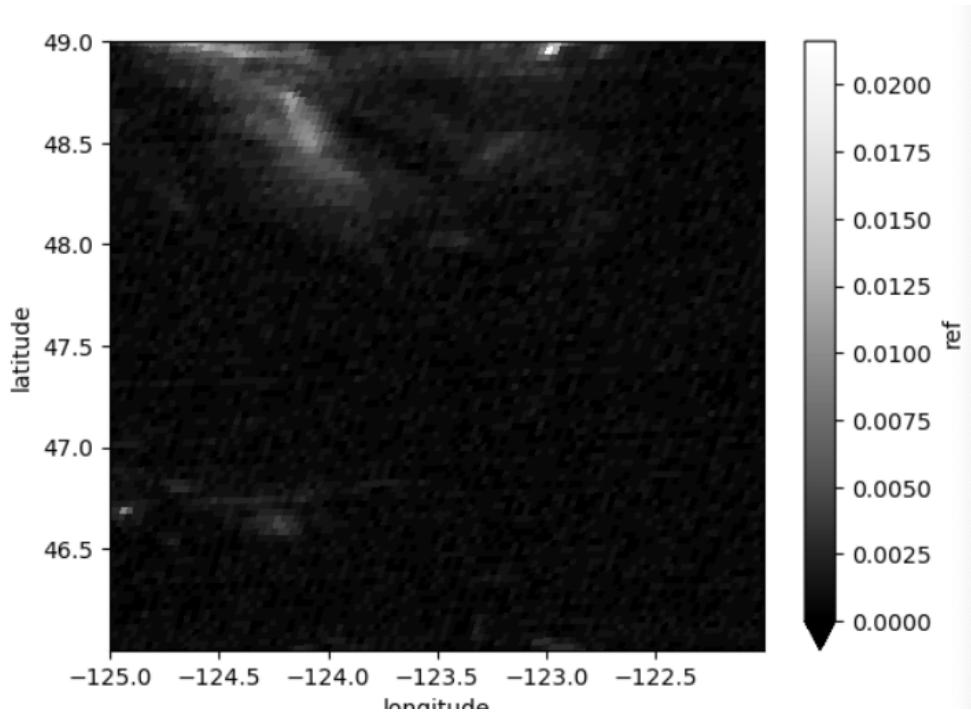
The code we wrote for this study, as well as most of the data used and produced, except for the camera photos which were too heavy to upload, are in this github repository:

<https://github.com/autumn-yng/summerfog>

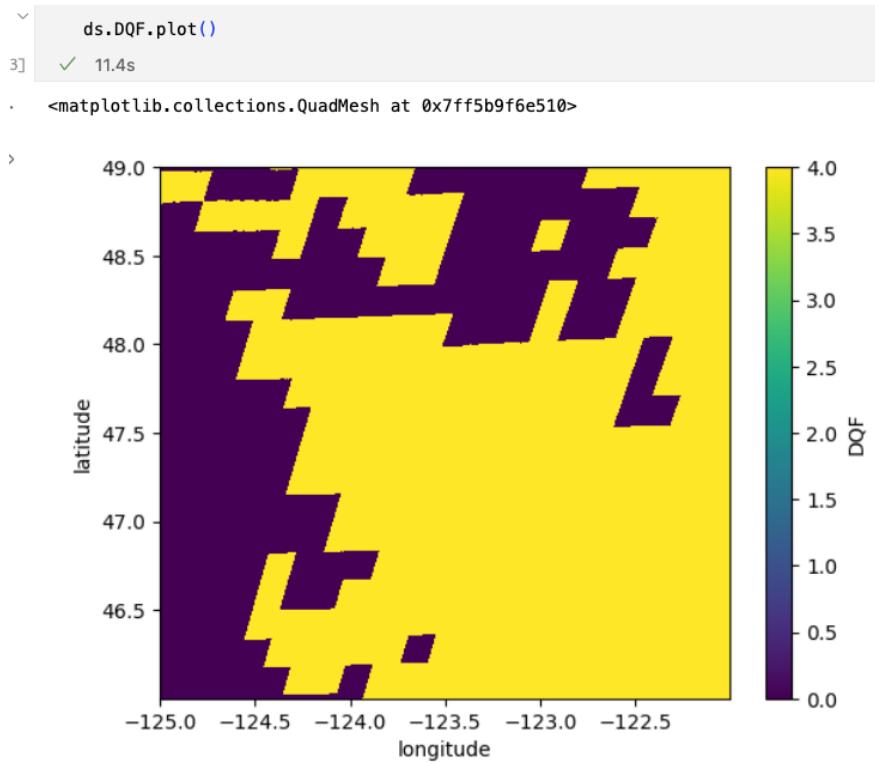
Appendix B: Relevant GOES-17 products



An image taken in Band 2, which is for detecting the ground and low clouds. The detection of low clouds in this image was consistent with the detected cloud height that indicated low clouds in the Cloud Top Height product at the same timestamp.



An image taken in Band 4, which is for detecting the cirrus clouds. The detection of cirrus clouds in this image was consistent with the detected cloud height that indicated high clouds in the Cloud Top Height product at the same timestamp.



Data Quality Flags. Comparing this plot with the Cloud Top Height product for the same timestamp confirmed that the NaN pixels meant no cloud was detected by the Cloud Mask and hence wasn't processed by the Cloud Top Height algorithms.