

California Coast Ecological Conservation Coupling Satellite Observations and Machine Learning to Understand Future Fog Prevalence Along the Coast Redwood Range

Summer 2025 | California – Ames
August 8th, 2025

Authors: David Wilcox (Analytical Mechanics Associates), Aubrey Palmer (Analytical Mechanics Associates), Paul Seibert (Analytical Mechanics Associates), Macey Hartmann (Analytical Mechanics Associates)

Abstract: Fog engulfs the California coast during the dry summer months, serving as a vital water source for coast redwoods (*Sequoia sempervirens*). The future of fog is uncertain, threatening redwoods and their functions: carbon sequestration, ecosystem balance, and recreational value. Previous research has found mixed outcomes, with some studies expecting a decrease in fog, while others predict stability or even an increase. To address this uncertainty, NASA DEVELOP partnered with Save the Redwoods League to use satellite data and machine learning to better understand the future of fog. Using new satellite products, we developed a method for quantifying coastal fog presence. Then, we trained a machine learning model to use climate projection and satellite-derived fog data to predict changes in fog through the year 2090. The model predicted fog to increase along most of the coast redwood range, with an average increase of 3.4 hours of fog per day, which would benefit redwood conservation and carbon sequestration. To more closely examine these changes, we analyzed sea and land surface temperature trends over the last quarter-century, finding mixed impacts on fog. This understanding of fog—past, present, and future—informs redwood conservation for Save the Redwoods League and guides future research and monitoring efforts.

Key Terms: Coastal advection fog, fog and low cloud coverage (FLCC), machine learning, Random Forest, MODIS, GOES, coast redwoods (*Sequoia sempervirens*)

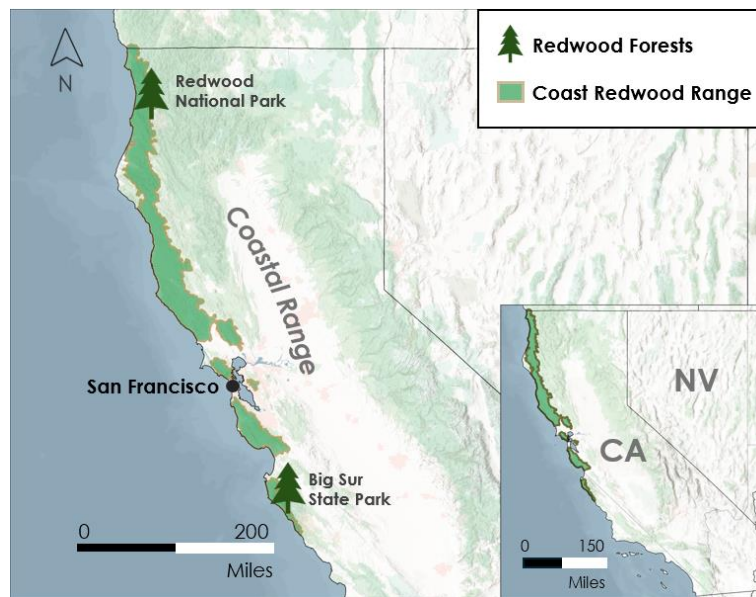
Advisors Dr. Morgan Gilmour (NASA Ames Research Center), Dr. Chris Potter (NASA Ames Research Center), Dr. Weile Wang (NASA Ames Research Center), Dr. Taejin Park (NASA Ames Research Center)

Lead: Katie Miller (California – Ames)

1. Introduction

Old growth coast redwoods (*Sequoia sempervirens*) are the tallest trees in the world, commonly exceeding 100 meters in height (Fujimori, 1977). They are among the longest living trees too, with lifespans of 1,200–1,800 years or more (National Park Service, 2007). Conserving coast redwoods is important, as they provide habitat for native flora and fauna, play a major role in mitigating climate change through carbon sequestration, and are culturally significant for many communities.

The entire coast redwood range is located within a thin strip along the Pacific coast from Monterey County, California to southern Oregon (Figure 1). This region experiences a Mediterranean climate, with most rainfall occurring between October and April (rainy season), and less than 10% occurring from May through September (dry season; Dawson, 1998). During the dry summer months, fog insulates redwoods from solar radiation, suppresses the loss of water to the atmosphere (transpiration), is directly intercepted by the canopy and fed to the understory through stemflow, and is absorbed through foliar water uptake (Burgess & Dawson, 2004; Dawson, 1998). In short, fog is a key drought buffer for coast redwoods.



Basemap Credits: FDEP, ESRI, TomTom, Garmin, SafeGraph, FAO, METI/NASA, USGS, EPA, NPS, USFWS, Esri, NASA, NGA

Figure 1. Project study area, defined by the historic range of old-growth coast redwoods (*Sequoia sempervirens*) across California, courtesy of Save the Redwoods League. Current old-growth coast redwoods only occupy about 5% of their historical range (Save the Redwoods League).

Across this range, climate change's effect on fog is uncertain. Johnstone & Dawson (2010) found that fog frequency is decreasing across the region, while a previous DEVELOP publication suggests that there was no significant increase or decrease in fog frequency in the past twenty years (Werner et al., 2022). Given the importance of redwoods, more research is needed to confidently understand future changes in fog.

Geospatial fog research is underpinned by remote sensing techniques, using satellites such as the Geostationary Operational Environmental Satellite (GOES) Advance Baseline Imager and NASA's Aqua Moderate Resolution Imaging Spectroradiometer (MODIS). The GOES satellite offers frequent imaging (every 5 minutes) at lower spatial resolution (2km), while Aqua MODIS provides higher spatial resolution (250m–1km) imaging at a lesser frequency (daily). These satellites offer a broad understanding of the extent and duration of fog.

Previous research has used these satellites to map fog across coastal California. Torregrosa et al. (2016a) used the GOES satellite to create monthly datasets of fog and low cloud cover from 1999–2009, and a previous DEVELOP project tested the ability of the Terra MODIS satellite to create similar daily and monthly datasets (Werner et al., 2022). More recently, researchers have combined satellite observations with machine learning techniques, such as random forests, to study fog (Haynes et al., 2022; Werner et al., 2022).

Building on this, our project aimed to understand how fog will change along the coast redwood range. We created fog projections to the year 2090 by training a machine learning model on historical fog datasets. We analyzed past fog frequency during June–September from 2023–2024, and dug deeper into other environmental variables, such as sea and land surface temperature, to understand underlying drivers for our future fog projections. This allowed us to create a past, present, and future understanding of fog across the California coast.

This project was designed to support conservation planning by our project partner, Save the Redwoods League, a 501(C)(3) nonprofit organization that works to protect, restore, and connect redwood forests with people by purchasing land for coast redwood conservation. Currently, the League uses a mapping tool which includes several climate variables to understand where coast redwoods will be most resilient, allowing them to prioritize their conservation efforts. The tool does not currently consider fog, so incorporating fog projection data from this project can allow for more informed conservation decisions. By combining satellite observations with machine learning, this project developed a strong methodology to project changes in fog. A deeper analysis of driving variables, like sea and land temperatures, can inform future monitoring and conservation efforts. As climate change shifts fog patterns, this insight is critical for redwood conservation.

2. Methods

To develop a comprehensive understanding of fog—past and future—we integrated datasets from a variety of sources. To understand past fog, we employed the GOES-18 satellite, Aqua MODIS sea and land surface temperature products, and weather station data. To project future fog patterns, we used historical and projected climate data and a satellite-based fog climatology dataset.

2.1 Data Acquisition

To analyze recent fog frequency, we leveraged both satellite and in situ data. We used GOES-18 Level 2 Cloud Top Phase (ABI-L2-CTP) scans over the Pacific U.S. sector at 5-minute intervals and 2km resolution, downloaded via Amazon Web Service S3 Buckets using a Python (3.12.7) script and the boto3 package. To validate these satellite-derived observations, we used weather station data from eight stations in the California Irrigation Management Information System: Smith River #263, Watsonville West II #209, Pajaro #129, Pescadero #253, Carmel #210, De Laveaga #104, Pacific Grove #193, and Ferndale Plain #259. We downloaded air temperature and dew point temperature (°C) data from each station covering the beginning of the station record through June 30, 2025, at hourly intervals via their website (California Irrigation Management Information System, 2007). Together, both data types allow for a more comprehensive assessment of fog frequency.

To examine environmental drivers of fog, we used Aqua MODIS sea and land surface temperature products. We used the Aqua satellite because it passes over California at approximately 1:30 AM and 1:30 PM local time, near the hottest time of the day, capturing the hottest, most stressful temperatures for redwoods. We downloaded daytime and nighttime land surface temperature (MYD21A2-061, 1km resolution) and daytime skin sea surface temperature (MODSA-8D4D9, 4km resolution) data from July 2002 to June 2025. These were acquired from the Physical Oceanography and Land Processes Distributed Active Archive Centers (PO.DAAC and LP.DAAC) using the LP.DAAC Application for Extracting and Exploring Analysis Ready Samples (AppEEARS) interface and the PO.DAAC data downloader programming interface, respectively. These interfaces clipped, reprojected, and formatted the data as NetCDF files. These datasets allowed us to examine changes in sea and land surface temperatures, important drivers of fog formation and longevity.

For future fog projections, we used historical climate, historical fog, and climate projection data. For historical and projected climate data, we used NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP-CMIP6), which are downscaled from the Coupled Model Intercomparison Project Phase 6 (Thrasher et al., 2022). We acquired historical data from 1999–2009 and projected data from 2080–2090, in 25 km daily resolution, focusing on the variables of surface relative humidity (%), surface specific humidity (%), near surface temperature (K), near surface temperature minimums and maximums (K), downwelling longwave radiation (W/m^2), downwelling shortwave radiation (W/m^2), surface wind speed (m/s), and precipitation rate ($\text{kg m}^{-2} \text{s}^{-1}$). We used the current-track climate projection scenario, SSP2-4.5. The dataset was provided to the team courtesy of NASA GeoNEX. These climate projections are key for training the random forest to detect fog.

We also required historical fog data. The satellite products used in our own fog analysis are only 5 years old, and more data is necessary to accurately train a machine learning model. Using GOES satellite data, Torregrosa et al. (2016a) created a 10-year dataset of monthly fog and low cloud cover hours per day from 1999–2009 at 4km resolution. This dataset originates from the California Climate Commons database (Torregrosa et al. 2016b). Two months (June 2001 and August 2006) are missing from the dataset and were not included in any further analysis.

2.2 Data Processing

To apply machine learning methods spatially, all data must match in projection, spatial resolution, grid alignment, bounding box, and time index. This allows it to easily be converted from geospatial formats (NetCDF) to a table-style dataset (pandas DataFrame) to train the model. This was a significant, time-intensive step in our methodology.

Table 1

Common values assigned to the data during the processing step

Projection	EPSG 4326 (World Geodetic System 1984)
Bounding Box	Latitude 33.995103 to 42.224943, longitude -120.011854 to -128.029946
File Format	NetCDF

For historical fog, we processed the GOES and weather station datasets using Python. First, we time indexed the GOES satellite’s Cloud Top Phase files based on the filename, grouping them in an hourly dictionary. For the weather station validation data, we combined the individual stations into a singular csv file, then graphed them to find ‘dead periods’ where sensors were not operating as expected so we could omit the data from further analysis. Using the remaining data, we calculated the dew point depression, or the difference between the air temperature and dew point temperature. When the dew point depression is at or below 2.5°C , and no rain is recorded at the station, we assume fog is present (American Meteorological Society, n.d.; National Weather Service, n.d.). Using the Python packages Pandas and NumPy, we removed any observations with rain from the dataset and added a new column indicating fog presence based on the dew point depression (Haynes et al., 2022; pandas Development Team, 2020). From there, we resampled the data from an hourly to a daily timestep to show station-based fog hours per day, then resampled to a monthly timestep to show fog hours per month. This processing allowed our GOES data to be compared to the in situ weather station data.

For a deeper investigation into fog drivers, we analyzed sea and land surface temperatures. First, we subset the Aqua MODIS sea and land surface temperature datasets to only include the months of June–September from July 2002 to June 2025. Next, we calculated the average summertime temperature for each year. Finally, we merged both datasets to highlight warming trends of both land and ocean in later analysis.

For fog projection, the historical and future climate dataset is per-day, but the future climate projections of the dataset are per-month. To prepare these for our model, the historical portion was upsampled from a daily to a monthly timestep. We then cropped both datasets to the bounding box, reprojected accordingly, and merged the historical files into one file with all meteorological variables using the Python packages xarray and rioxarray (Hoyer et al., 2016; Hoyer et al., 2017). To prepare the fog dataset (Torregrosa et al., 2016b) for our model, we upsampled the native resolution (4km) to match the climate data (25km). We then snapped it to the grid of the climate data with nearest neighbor interpolation, cropped it to the bounding box, and converted to a NetCDF using ArcGIS (version 3.4.0). Each monthly NetCDF file was combined into one time-indexed file using Python xarray (Hoyer et al., 2016). This enabled both of the datasets to be fully comparable over the same region, time, and grid.

2.3 Data Analysis

2.3.1 Historical Fog and Low Cloud Cover

To estimate historical fog frequency, we used the GOES Cloud Top Phase product (GOES-R Algorithm Working Group, 2018). This product examines data from the GOES satellite and current weather models to examine the tops of clouds. It classifies each pixel into one of seven categories: warm liquid water, supercooled liquid water, mixed phase, ice phase, clear sky, no data, or uncertain pixels. In the summertime in California, marine layer cloud tops are almost always above freezing, so warm liquid water pixels can be considered fog.

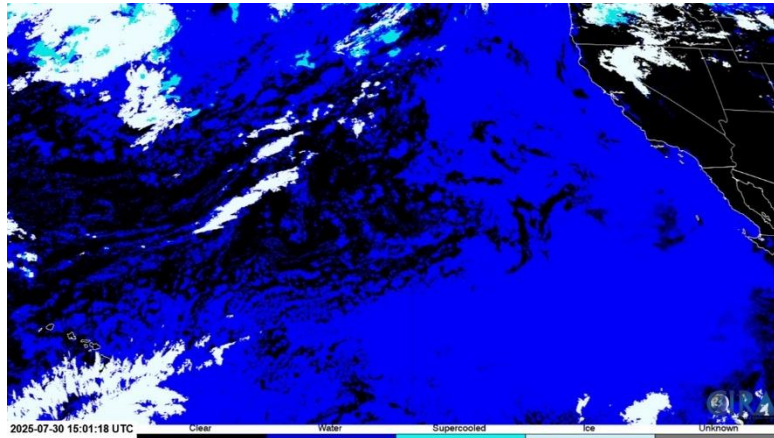


Figure 2. An example of the GOES-R Cloud Top Phase product; the large swath of blue is the marine layer of fog and low stratus across the Pacific Ocean, while areas of white and light blue show high cirrus clouds or convective thunderstorms (from GOES-18, July 30, 2025 at 8:01 AM PDT; Image Credit NOAA/CIRA).

We analyzed each time-sorted scan one month at a time, counting the amount of time each pixel was classified as warm liquid water cloud. Since the GOES-18 Pacific U.S. sector scans are taken every 5 minutes, each warm liquid water pixel represents 5 minutes of fog or low stratus cloud at that location. We wrote a script using xarray and NumPy that opened each data file, and then added 5 minutes to a ‘low cloud minutes’ counter at warm liquid cloud pixels, added 5 minutes to a ‘non-low cloud minutes’ counter at the other pixels, and ignored any clear sky, no data, or uncertain pixels (Harris et al., 2020; Hoyer et al., 2016). It is important to track non-low clouds because other, higher clouds can block the sensor from detecting low clouds. For each month, this approach provided the number of low cloud minutes, plus the number minutes with clouds that might obscure the sensor from detecting fog.

This is a new approach to studying fog. Previous detection methods, such as the one used by the fog dataset driving our predictions, are different. They utilize differencing of two infrared bands. The Cloud Top Phase product used in this new methodology does utilize this band differencing, but as one step in a broader overall

methodology. While the band differencing method is simpler, it lacks any formal validation and may be less accurate. Recent studies incorporate more complex methods of analysis into band differencing, such as dynamic temperature thresholds and spatial filtering, achieving detection probabilities from roughly 60–80% (Jahani et al., 2025; Ma et al., 2025). The Torregrosa et al. (2016a) band differencing method used simpler forms of analysis and likely falls at the lower end (or below) this range.

In addition to detection methods, sensor resolution also impacts which local details can be seen in fog trends. The GOES scans used in the Torregrosa et al. (2016a) analysis were at a larger 4km resolution and 15-minute intervals, but the newer GOES-18 satellite used in this analysis scans at 2km resolution and 5-minute intervals. This higher sensor resolution captures details in fog patterns which have not been observed before.

To validate the accuracy of this method, we used in situ weather station observations to also compute the number of fog hours. We grouped the DataFrame of observations by individual station, then added up the number of observations which met the fog hour dew point deficit threshold of 2.5°C. Observations are taken each hour, so one observation equals one hour of fog. For each station, we totaled the number of fog hours, then divided by the total number of days, yielding the average number of fog hours per day based on weather station observations.

To compare this to the GOES analysis, we extracted the average fog hours per day from the pixel nearest to the coordinates of the weather station using the Python package NumPy. We plotted each weather station's fog hours calculated from the weather station versus fog hours calculated from the GOES analysis. Comparing both datasets (where weather stations are available) increases overall confidence in the GOES fog and low cloud detection results.

2.3.2 Land and Sea Surface Temperature Trends and Fluctuation

For sea and land temperature analysis, we applied polynomial fitting to the datasets using xarray's polyfit function, yielding temperature rate of change (in °C per year). We omitted rates of change less than -0.5 °C (land) or -2.0 °C (sea), and greater than 0.1°C (land) or 2.0°C (sea) to control for outliers. To highlight sea and land temperature extremes, we created a temperature anomaly map by subtracting the average temperature across the 23-year period from the single year mean in 2022, showing how much temperature in specific regions differ from long-term averages. To depict temperature's general fluctuation patterns across the coast redwood range, we used xarray's diff function to create a map of standard deviation in land temperature. These maps show spatial trends in the processes governing coastal fog.

2.3.3 Predicting Future Fog

To project future fog frequency under changing climate conditions, we trained a random forest regression model using the Python package scikit-learn (Pedregosa et al., 2011). This is a common machine learning technique in Earth sciences for understanding and predicting complex environmental phenomena. Random forest models can capture non-linear relationships, make no prior assumptions about the data, are easily scalable, and are relatively easy to understand than more complex machine learning models (Breiman, 2001). The model works by creating a “forest” consisting of many decision trees. The forest predicts the average number of fog hours per day in each month based on meteorological variables.

We randomly divided the data, the number of fog hours per day (Torregrosa et al., 2016b) and the meteorological predictor variables (CMIP6), into a 70/30 split, using 70% of the data to train the model, and the remaining 30% to test the model. We omitted a single, random month of the fog and low cloud dataset from the training data. This setup allows us to assess the model's performance on both an unseen time period and spatial region.

Once trained, the model predicted values by stepping through each decision tree, then averaging the predicted values across all trees to generate the final prediction. This ensemble approach is effective due to the Law of

Large Numbers, which states that over a large number of random samples with the same probability, the average will converge to the expected value (Equation 1). Mathematically, this can be expressed as

$$\sum_{i=1}^n \frac{x_i}{n} = \bar{x} \quad (1)$$

where n is the sample size and x_i is the value of each sample x in row i (Breiman, 2001).

To build the most accurate model but still maintain generalizability and limit computational power, we conducted a hyperparameter grid search. The model uses several hyperparameters to prevent overfitting while still maintaining accuracy. Our search examined maximum number of trees (1–100), maximum depth of trees (1–10), and minimum samples required to be a leaf node (1–15). Importantly, we used a random state of 40 to ensure replicability. The model generated a random forest for every combination of these hyperparameters, generating 15,000 total forests. We used these optimized hyperparameters to generate the final model used to predict future fog.

3. Results

The above methods projected fog to increase along nearly the entire coast redwood range. Leveraging new products created with the GOES satellite gives a higher-resolution, higher-accuracy depiction of current fog. Ultimately, the analysis from the MODIS satellite can inform future fog monitoring efforts.

3.1 Historical Fog and Low Cloud Cover

Using the GOES-18 Cloud Top Phase product offered valuable insight into fog presence along the California coast at a higher resolution and higher temporal frequency. While previous analysis of fog, such as the Torregrosa et al. (2016b) dataset, reveals trends with lots of spatial homogeneity, this new analysis encapsulates many more of the local influences affecting fog at any one point along the coast. This is important when applying observations, such as in redwood conservation, as conditions can change drastically over a short distance.

In Figure 3, areas shaded in green represent 3+ hours of fog per day. While it is not a settled fact, collaboration with Save the Redwoods League and previous studies (Dawson, 2004) decided to use this value as a meaningful amount of fog for coast redwoods. The higher-resolution fog detection method using the Cloud Top Phase product highlights local maximums of fog not seen in other methods, such as thermal differencing.

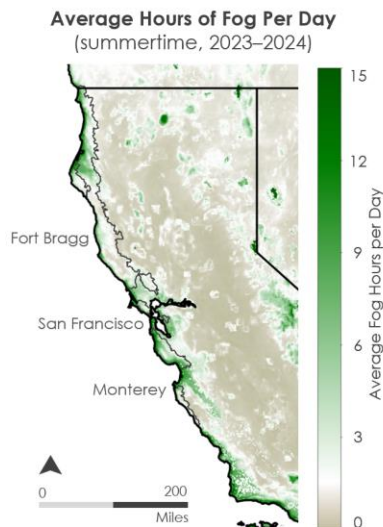


Figure 3. Map showing average hours of fog per day during the summertime across the coast redwood range, calculated for 2023 and 2024. Green areas represent 3+ hours of fog per day. For a comparison between this product and Torregrosa et al. (2016b), see Appendix C.

One example of this local insight is along the North Coast (Figure 4). During the summer, fog is steered by the predominant northwest wind caused by the North Pacific high pressure system. The angle of the wind pushes heavy amounts of fog onshore along the North Coast, particularly north of Fortuna (highlighted in red). Cities in this region are among some of the foggiest in our analysis, with many averaging 10+ hours of fog per day.

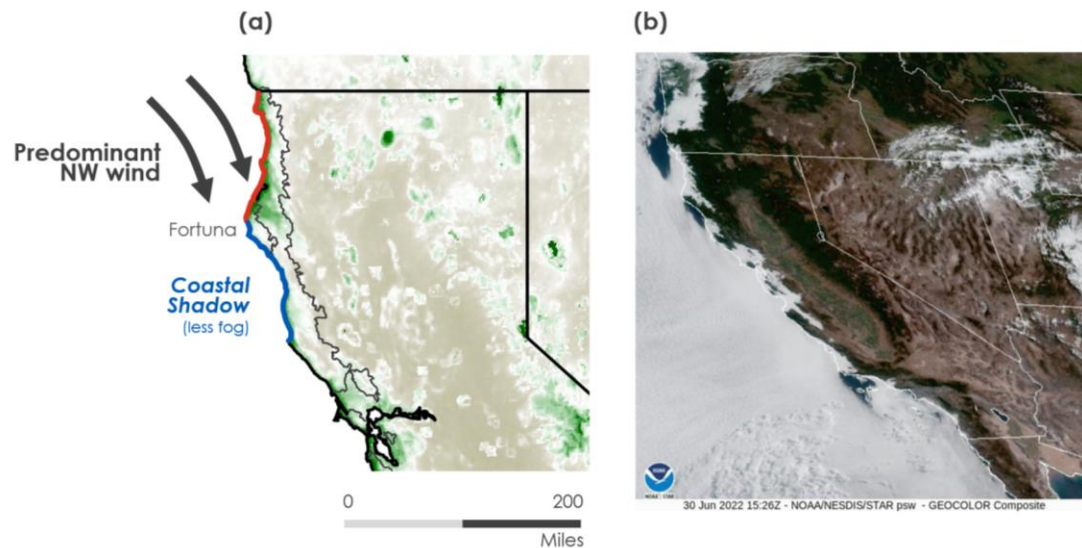


Figure 4. The predominant wind direction, caused by the North Pacific high-pressure system, creates “fog shadows” south of protruding portions of the coastline. This can be seen via satellite (from GOES-18, June 30, 2022, at 8:26 AM PDT; Image Credit NOAA).

In contrast, coastal areas south of Fortuna (highlighted in blue) experience considerably less fog, living in the “shadow” of the foggyer north portion (red) of the coast. While this effect can be seen in the data—the red portion of the coastline is “greener” than the blue—it is also evident on individual satellite images taken on foggy days. While methods like those employed by Torregrosa et al. (2016a) paint broad-brush climatologies of fog, these higher-resolution, higher-accuracy methods reveal important local differences in fog.

To ensure these finer distinctions are accurate, we analyzed ground-based weather stations from the California Irrigation Management Information System. Our analysis sorted through these stations to find usable data. During the 2023–2024 period, only 6 stations within or near the study area were viable to use.

Despite the small sample size, these results do offer valuable insight into the accuracy of the new method (Figure 5). Four of the stations are clustered around the 1:1 line, meaning that the weather stations and Cloud Top Phase detection method were in relatively good agreement, and two stations are significant outliers. For the bottom-most station on the plot, Pajaro #129, the GOES method only recorded an average of ~2 hours of fog per day, which is a clear anomaly. While we did not have time to thoroughly investigate, this might occur due to coastal pixel effects within the Cloud Top Phase product. For the right-most station on the plot, Pacific Grove #193, the GOES method recorded a reasonable 13.4 average fog hours per day, while the weather station recorded a less likely 18.2 average fog hours per day. This could be caused by an over-sensitive hygrometer. Excluding extreme outliers, the GOES Cloud Top Phase methodology for detecting fog matched closely with in situ weather station observations, creating a promising new way to study fog.

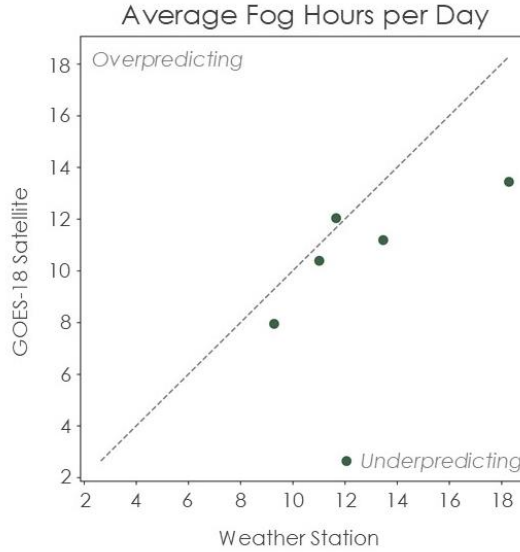


Figure 5. Plot depicting average fog hours per day (June–September 2023–2024) according to the GOES satellite Cloud Top Phase product versus the ground-based weather stations. Each point represents one weather station location. Locations along the dotted (one-to-one) line show agreement between the satellite method and ground-based observations, while locations further from the line show disagreement. For a station map and individually labeled points, see Appendix D.

3.2 Predicted Future Fog and Low Cloud Cover

The model was able to predict future fog and low cloud cover hours, which we averaged over the decade of 2080–2090 (Figure 6, center). There are large differences in the number of fog hours between the historical and future fog maps, as illustrated in the right of Figure 6. The model predicted an average increase of 3.4 hours across the coast redwood range.

In general, our model predicted increased fog along the coast redwood range, with the notable exceptions of the Point Reyes and Monterey Peninsula. The largest increases, up to 7.5 hours, are predicted at the southern end of the range near Pfeiffer Big Sur. This result differs significantly from past research on current trends of fog along the coast of California. Whereas past research has shown either a decline in fog (Johnstone & Dawson, 2010) or no significant trend in fog (Werner et al., 2022), this research predicts a widespread increase in fog. These predicted changes in daily fog and low cloud cover are fairly drastic, and future research should focus on making the predictions more robust for use in conservation efforts.

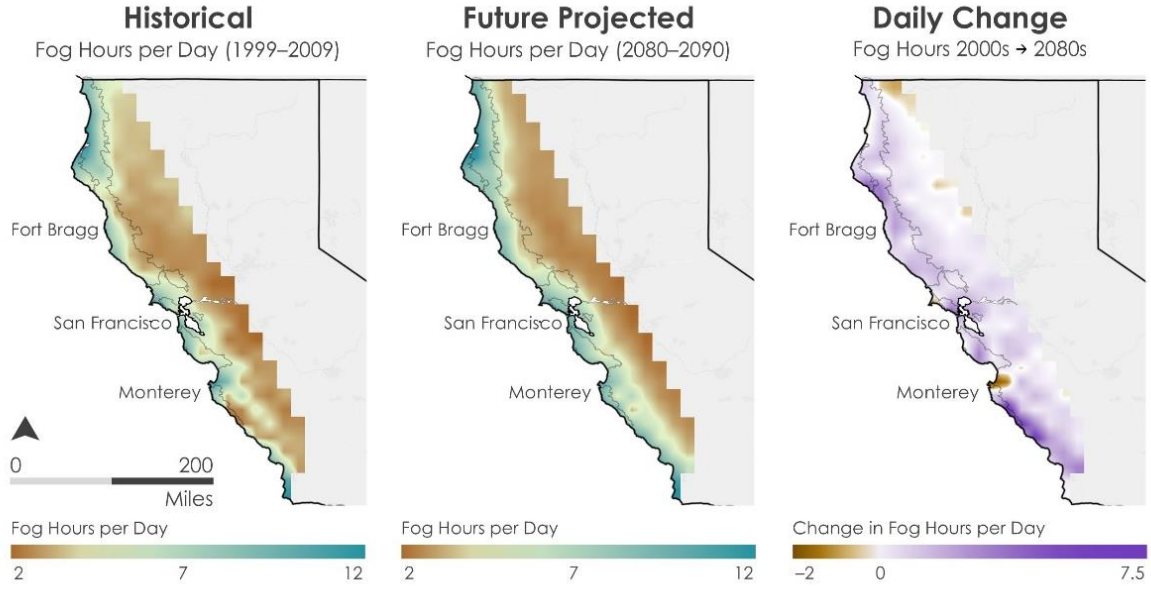


Figure 6. Maps of fog and low cloud cover over the coast redwood range, outlined in grey. Left shows the average daily hours of fog and low cloud cover from June–September 1999–2009 (Torregrosa et al., 2016b).

Middle shows the random forest’s predicted daily average hours of fog and low cloud cover in June–September 2080–2090. Right is the difference in the future and past fog maps, showing projected increases (purple) and decreases (brown). Note, these maps are smoothed using bilinear interpretation; see appendix X for the native resolution maps.

3.3 Random Forest Regressor Hyperparameters

Hyperparameters allowed us to tune the model’s accuracy and reduce computational effort. From the hyperparameter search, we found that the most accurate model had 99 trees, a maximum tree depth of 9, and 2 minimum samples required for a leaf node. Maximum number of trees and maximum depth of each tree had a much higher impact on the model accuracy than the minimum samples for a leaf node. In this use case, the accuracy of the model plateaus at 15 trees (Figure 7a) and a tree depth of 9 (Figure 7b), and the minimum samples for a leaf node had a minimal effect on accuracy (Figure 7c). The hyperparameter search was both time and computationally intensive but knowing which hyperparameters were most important helped us to better understand how the model works and limit computational effort in future hyperparameter searches.

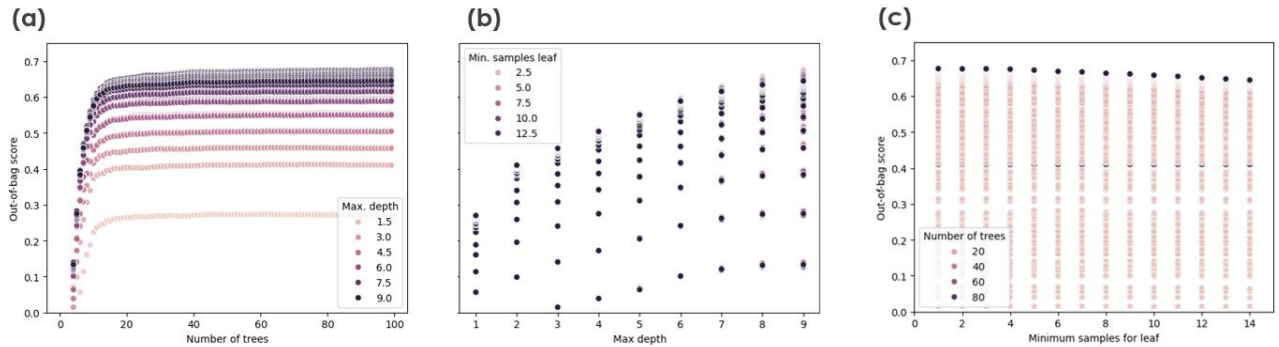


Figure 7. Scatter plots showing the influence of random forest hyperparameters on the accuracy of the model.

(a) Scatter plot of the effect of the number of trees in a random forest on the out-of-bag-score (R^2 on previously unseen test data), colored by the maximum depth of decision trees within the forest. (b) Scatter

plot of the effect of the maximum depth of decision trees within the forest on out-of-bag score, colored by the minimum samples in the leaf node of a decision tree. **(c)** Scatter plot of the effect of the minimum samples for a leaf node on the out-of-bag score.

3.4 Random Forest Regressor Accuracy

Using the optimized hyperparameters, we built a relatively accurate model. The coefficient of determination between the true daily fog hours (Torregrosa et al., 2016b) and the model’s predicted daily fog hours is 0.68 (Figure 8a). The root mean square error (RMSE) is 2.1 hours, relative to a calculated historical average of 5.5 hours within the coast redwood range. Given this, the accuracy of our model is suitable for the goals of this study. In addition to these metrics, we also left out a single month of data (June 2003) in our training data to analyze the spatial accuracy of the model (see Figure A1).

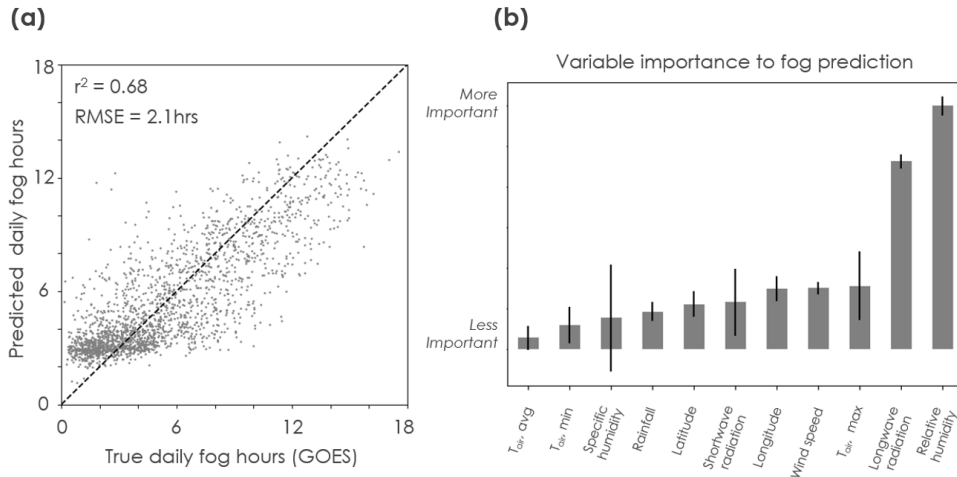


Figure 8. (a) Scatter plot of “true” average June–September fog hours per day using (Torregrosa et al., 2016b) versus the model-predicted average June–September fog hours per day. The dashed black line indicates the one-to-one line, where every point would lie if the model was perfect. **(b)** Variable importance (mean $\pm$ standard deviation) in the climate projection dataset for predicting fog and low cloud cover.

To understand the drivers of our model, we analyzed variable importance (Figure 8b). Specifically, we examined Gini importance, a measure of how well a single variable improves a predictive machine learning model. Relative humidity emerged as the single most important variable for predicting fog, followed by downwelling longwave radiation. This makes sense—high relative humidity (i.e., 100%) usually indicates fog presence, plus downwelling longwave radiation is associated with low clouds reflecting the Earth’s heat back toward the surface. Maximum daily air temperature was the third most important variable. Fog acts as a shield to sunlight, keeping daytime temperatures cooler than they otherwise would be. Air temperatures also dictate how much moisture the air can hold.

The future projections of these three variables driving our future fog and low cloud cover predictions can be seen in Figures B1, B2, and B3. For a more comprehensive look at the model-predicted relationships between relative humidity, downwelling longwave radiation, maximum air temperature, and fog predictions, see appendix Figures C1, C2, C3, respectively. Given that the model behaves in a manner in line with the known meteorology behind fog formation, we have confidence that the general patterns driving the model are correct.

3.5 Land and Sea Surface Temperature Trends and Fluctuation

The key factor in producing coastal advective fog is the temperature difference between the warm waters of the North Pacific and the cooler waters of the California current along the U.S. West Coast. As warm, moist air from the Pacific is pushed across the California current, the air cools and condenses, forming fog that

spills into the coast. Given the importance of this ocean temperature difference to fog formation, our temperature analysis reveals important trends that could impact coastal fog formation.

When interpreting our results, it's important to note that this analysis examines skin sea surface temperature and land surface temperature. Surface, or skin, temperature is the temperature at the land surface or within the top 3cm of ocean, not actual air or ocean temperatures. The latter of which are much more difficult to accurately analyze using remote sensing methods.

Within our analysis, we found general warming trends in ocean temperature along the North Coast, with a notable hotspot near Humboldt County, CA (Figure 9). Further south, we found small, clustered cooling patches ~200 miles off the coast of Monterey, CA. These may be associated with the projected cooling trend in fog around the Santa Cruz Mountains at the southern end of the redwood range. The magnitude of the temperature difference which produces fog, ~10 °C, is two orders of magnitude larger than the magnitude of the warming trend, 0.1 C°/yr. Although these trends are notable, this early look suggests that they are not enough to have a significant impact on coastal fog production.

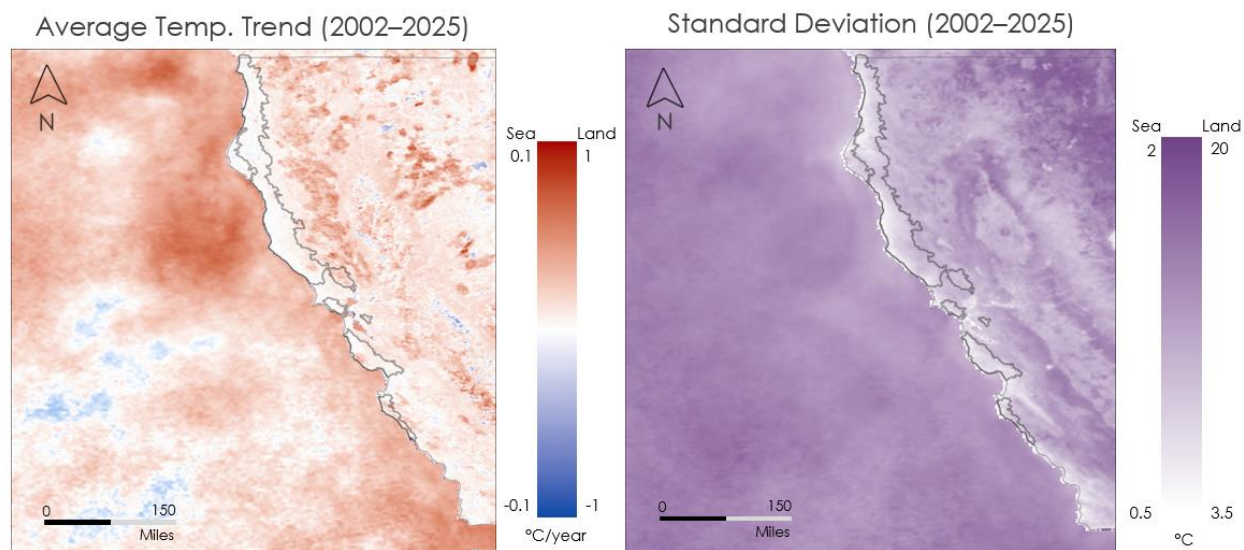


Figure 9. Combined and processed Aqua MODIS Land and Sea Surface Temperature Maps. On the left, a map of the average annual temperature change from 2002–2025. On the right, a map of the standard deviation of temperature over 2002–2025.

Fog also depends on land temperatures, which dictate how long it takes for fog to “burn off” as it moves inland. Analysis revealed warming inland, but relative stability along the coast redwood range over the last quarter century. The standard deviation across this region is low, too, meaning that this area has relatively stable temperatures both day-to-day (diurnally) and, according to the trend analysis, season-to-season. This is expected in a region where fog frequently insulates the land surface from the warming sunlight.

3.6 Errors & Uncertainties

While we developed a strong methodology to study future fog, the results are limited by the quality and quantity of the data used by the model. The downscaled climate projections drive our model, but the coarse pixels limit the scale at which we are able to project fog, erasing local, small-scale details, particularly topography. As with any physical climate model, the inherent uncertainty of the climate projections is carried through to our results. While climate projections show an increase in humidity in arid regions, like the interior of California, observations from the last four decades have not followed this expectation (Simpson et al., 2023). The coast redwood range is not an arid region, but even proximity to errors like these can skew results.

The fog analysis dataset is the other main model input, and thus a source of uncertainty. Its 1.4 million data points were reduced to just a few thousand during the required upscaling. Increasing the amount of data points available to the model would produce a more accurate result. The overall method, however, is an effective way to project fog, and changes to input data, such as a longer-timespan fog history or different climate datasets, can be easily incorporated.

4. Conclusions

4.1 Interpretation of Results

Our results offer an early look at fog patterns along the California coast after the next 60 years. The model predicted a widespread increase in fog under the most likely future climate scenario. A few notable exceptions to this are the Monterey Peninsula and Point Reyes. The predicted decreases in fog in these areas may be associated with difficulties in the underlying climate projection model in discerning between the ocean and land surface (i.e. mixed pixels). Pixels that contain a mix of land cover types are notably difficult for climate models to predict accurately, and that may be influencing our results.

These results bode well for redwood conservation outcomes, but they are not the only factor to consider. Fires, drought, and a variety of other weather and climate extremes can hurt redwoods, and research into climate change's effect on these extremes is still ongoing. A better understanding of future fog allows for more comprehensive planning in the context of all of these extremes.

4.2 Feasibility & Partner Implementation

The insights this research offers will shape ongoing conservation efforts. Our deeper dives into underlying drivers, like sea and land temperature, inform future ground-based monitoring, and the development of this new methodology creates exciting opportunities for further research. Save the Redwoods League can incorporate these results into their digital mapping tool, allowing them to be considered with a host of other factors when choosing areas to preserve.

When using data to make conservation decisions, a cautious approach is necessary. With lower spatial resolution and confidence, these future fog projections are no exception. The pixel size is large—155,000 acres, about the size of Chicago. Fog conditions vary greatly across this area. The projections only consider one climate scenario, the current-track (SSP2-4.5), limiting the range of outcomes to anticipate. When making conservation decisions, to plant, or not to plant, these uncertainties can limit the clarity the data actually offers.

Similar to previous studies, we found that it is feasible to use Earth observations and machine learning to study fog and guide redwood conservation. Fog is challenging to study, as it relies on complex interactions between atmospheric, oceanic, and terrestrial systems. This methodology is highly adaptable, and can easily incorporate new, more advanced climate projections, variables, and fog datasets.

5. Acknowledgements

The California Coast Ecological Conservation team would like to thank our science advisors, Dr. Morgan Gilmour, Dr. Chris Potter, Dr. Weile Wang, and Dr. Taejin Park for their time, insight, and collaboration. We would like to thank our Lead, Katie Miller, for her guidance, support, and encouragement, as well as DEVELOP Fellow Maya Hall for her valuable input. We would also like to thank Amanda Clayton and Sarah Hafer-Martin (for their feedback and project coordination support. Finally, we express our gratitude towards Vrinda Suresh and Laura Lalemand with our partner, Save the Redwoods League.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Aeronautics and Space Administration.

This material is based upon work supported by NASA through contract 80LARC23FA024.

6. Glossary

Earth observations – Satellites and sensors that collect information about the Earth’s physical, chemical, and biological systems over space and time

DAAC – Distributed Active Archive Center, where NASA Earth observation data is stored

GOES – Geostationary Operational Environmental Satellite

LST – Land surface temperature

MODIS – Moderate Resolution Imaging Spectroradiometer

NetCDF – Network Common Data Form, a standard file format for geospatial data

NEX-GDDP-CMIP6 – NASA Earth Exchange Global Daily Downscaled Projections using Phase 6 of the Coupled Model Intercomparison Project

SSP – Shared Socioeconomic Pathway, a way of categorizing a particular greenhouse gas emission scenario

7. References

- Amazon Web Services. (n.d.). AWS SDK for Python (Boto3). Retrieved from <https://boto3.amazonaws.com/v1/documentation/api/latest/index.html>
- American Meteorological Society. (n.d.). Fog. In Glossary of meteorology. <https://glossary.ametsoc.org/wiki/Fog>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Burgess, S. S. O., & Dawson, T. E. (2004). The contribution of fog to the water relations of *Sequoia sempervirens* (D. Don): Foliar uptake and prevention of dehydration. *Plant, Cell & Environment*, 27(8), 1023–1034. <https://doi.org/10.1111/j.1365-3040.2004.01207.x>
- California Irrigation Management Information System (CIMIS). (2007). T-REX: CIMIS Weather Station Data. (Version 1.0) [Dataset]. UCAR/NCAR - Earth Observing Laboratory. <https://doi.org/10.26023/ECJN-RTD9-020X>
- Dawson, T. E. (1998). Fog in the California redwood forest: Ecosystem inputs and use by plants. *Oecologia*, 117(4), 476–485. <https://doi.org/10.1007/s004420050683>
- Fog guide. (n.d.). National Weather Service. https://www.weather.gov/media/zhu/ZHU_Training_Page/fog_stuff/fog_guide/fog.pdf
- Fujimori, T. (1977). Stem biomass and structure of a mature *Sequoia sempervirens* stand on the Pacific Coast of northern California. *Journal of the Japanese Forestry Society*, 52(11), 435–441.
- GOES-R Algorithm Working Group & GOES-R Series Program. (2018). NOAA GOES-R Series Advanced Baseline Imager (ABI) Level 2 Cloud Top Phase (ACTP) (Version 4) [Dataset]. NOAA National Centers for Environmental Information. <https://doi.org/10.7289/V5NP22QW>
- Harris, C. R., Millman, K. J., Walt, S. J. van der, Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. van, Brett, M., Haldane, A., Río, J. F. del, Wiebe, M., Peterson, P., Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Haynes, J. M., Noh, Y.-J., Miller, S. D., Haynes, K. D., Ebert-Uphoff, I., & Heidinger, A. (2022). Low cloud detection in multilayer scenes using satellite imagery with machine learning methods. *Journal of Atmospheric and Oceanic Technology*, 39(3), 319–334. <https://doi.org/10.1175/JTECH-D-21-0084.1>
- Hoyer, S., Fitzgerald, C., Hamman, J., & others. (2016). xarray: (Version 0.8.0). <https://doi.org/10.5281/zenodo.59499>
- Hoyer, S., & Hamman, J. (2017). xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, 5(1). <https://doi.org/10.5334/jors.148>
- Hulley, G. (2021). MODIS/Aqua Land Surface Temperature/3-Band Emissivity 8-Day L3 Global 1km SIN Grid (Version 061) [Dataset]. NASA Land Processes Distributed Active Archive Center. <https://doi.org/10.5067/MODIS/MYD21A2.061>

- Jahani, B., Karalus, S., Fuchs, J., Zech, T., Zara, M., & Cermak, J. (2025). Algorithm for continual monitoring of fog based on geostationary satellite imagery. *Atmospheric Measurement Techniques*, 18(8), 1927–1941. <https://doi.org/10.5194/amt-18-1927-2025>
- Johnstone, J. A., & Dawson, T. E. (2010). Climatic context and ecological implications of summer fog decline in the coast redwood region. *Proceedings of the National Academy of Sciences*, 107(10), 4533–4538. <https://doi.org/10.1073/pnas.0915062107>
- Ma, H., Chen, C., Yi, Z., Feng, H., & Wu, X. (2025). Himawari-8 satellite detection of morning terrain fog in a subtropical region. *Climate Services*, 38, 100551. <https://doi.org/10.1016/j.cliser.2025.100551>
- NASA/JPL. (2020). *MODIS Aqua Level 3 SST Thermal IR 8 Day 4km Daytime* (Version 2019.0) [Dataset]. NASA Physical Oceanography Distributed Active Archive Center. <https://doi.org/10.5067/MODSA-8D4D9>
- National Park Service. (2007, February 2). Sequoia-Kings Canyon NPs: The Giants of Sequoia and Kings Canyon. https://www.nps.gov/parkhistory/online_books/seki/stagner/sec2.htm
- pandas Development Team, The. (2020). pandas-dev/pandas: Pandas (Version latest) [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.3509134>
- Pavolonis, M., & Calvert, C. (2020, June 1). *Enterprise Algorithm Theoretical Basis Document for Cloud Type and Cloud Phase*. National Oceanic and Atmospheric Administration Center for Satellite Applications and Research. https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Enterprise/ATBD_Enterprise_Cloud_Type_v3_2020-06-01.pdf
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Simpson, I. R., McKinnon, K. A., Kennedy, D., Lawrence, D. M., Lehner, F., & Seager, R. (2023). Observed humidity trends in dry regions contradict climate models. *Proceedings of the National Academy of Sciences of the United States of America*, 121(1), e2302480120. <https://doi.org/10.1073/pnas.2302480120>
- Stagner, H. (1958). *The Giants of Sequoia and Kings Canyon* (5th ed.). Sequoia Natural History Association. <https://npshistory.com/publications/seki/stagner/sec2.htm>
- Thrasher, B., Wang, W., Michaelis, A., Melton, F., Lee, T., & Nemani, R. (2022). NASA Global Daily Downscaled Projections, CMIP6. *Scientific Data*, 9(1), 262. <https://doi.org/10.1038/s41597-022-01393-4>
- Torregrosa, A., Combs, C., & Peters, J. (2016a). GOES-derived fog and low cloud indices for coastal north and central California ecological analyses. *Earth and Space Science*, 3(2), 46–67. <https://doi.org/10.1002/2015EA000119>
- Torregrosa, A., Combs, C., & Peters, J. (2016b). GOES-derived fog and low cloud indices for coastal north and central California ecological analyses [Dataset]. California Climate Commons. <https://climate.calcommons.org/datasets/summertime-fog>

Werner, Z., Hin Choi, C. T., Winter, A., Vorster, A. G., Berger, A., O'Shea, K., Evangelista, P., & Woodward, B. (2022). MODIS sensors can monitor spatiotemporal trends in fog and low cloud cover at 1 km spatial resolution along the U.S. Pacific Coast. *Remote Sensing Applications: Society and Environment*, 28, 100832. <https://doi.org/10.1016/j.rsase.2022.100832>

8. Appendices

Appendix A: June 2003 spatial accuracy test of random forest

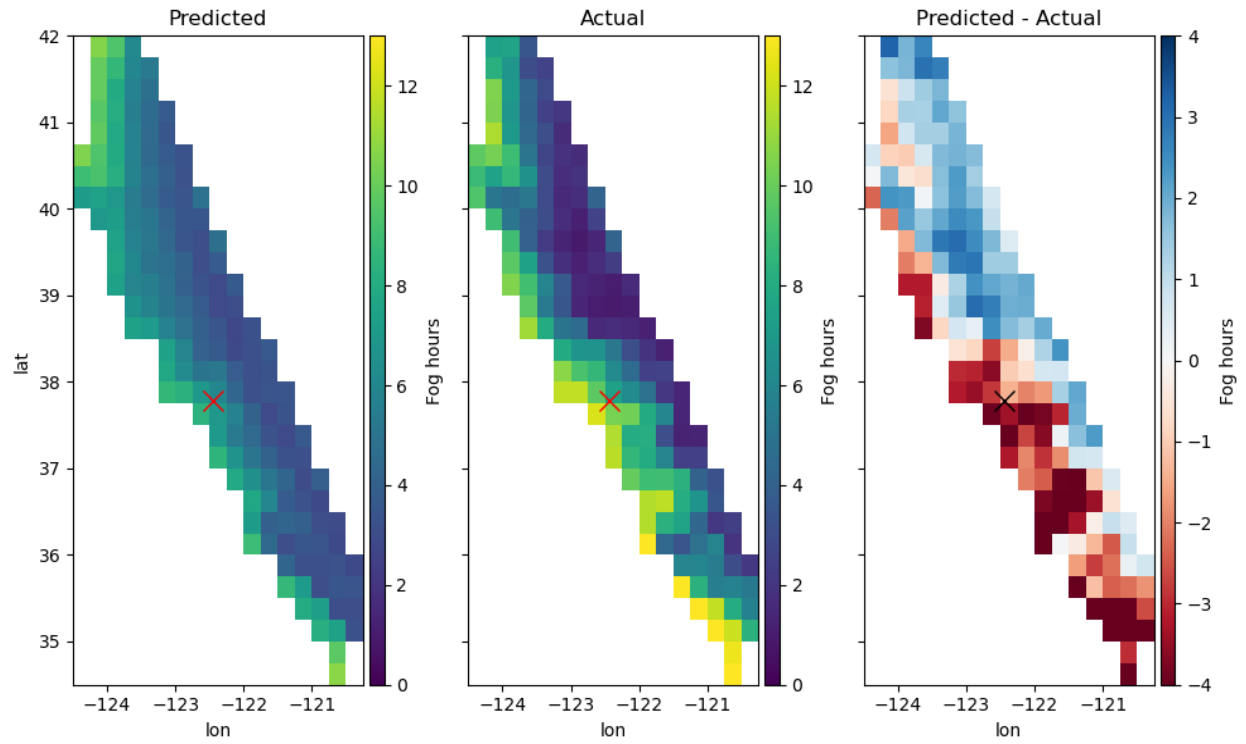


Figure A1: Maps of predicted (left), actual (center), and predicted minus actual (right) fog and low cloud hours, for the month of June 2003. The location of San Francisco is indicated by an X in each map.

Appendix B: NEX-GDDP-CMIP6 climate variable maps

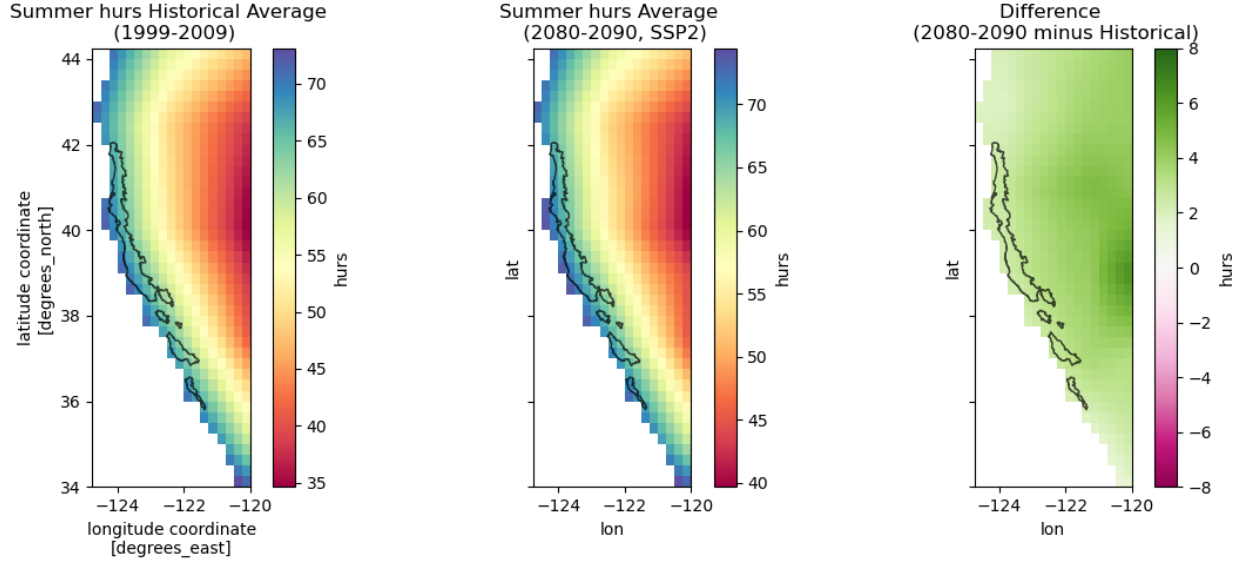


Figure B1: NEX-GDDP-CMIP6 Relative humidity (%) projections under SSP2-4.5. Averaged over 1999-2009 (left), averaged over 2080-2090 (center), and 2080-2090 average minus 1999-2009 average (right).

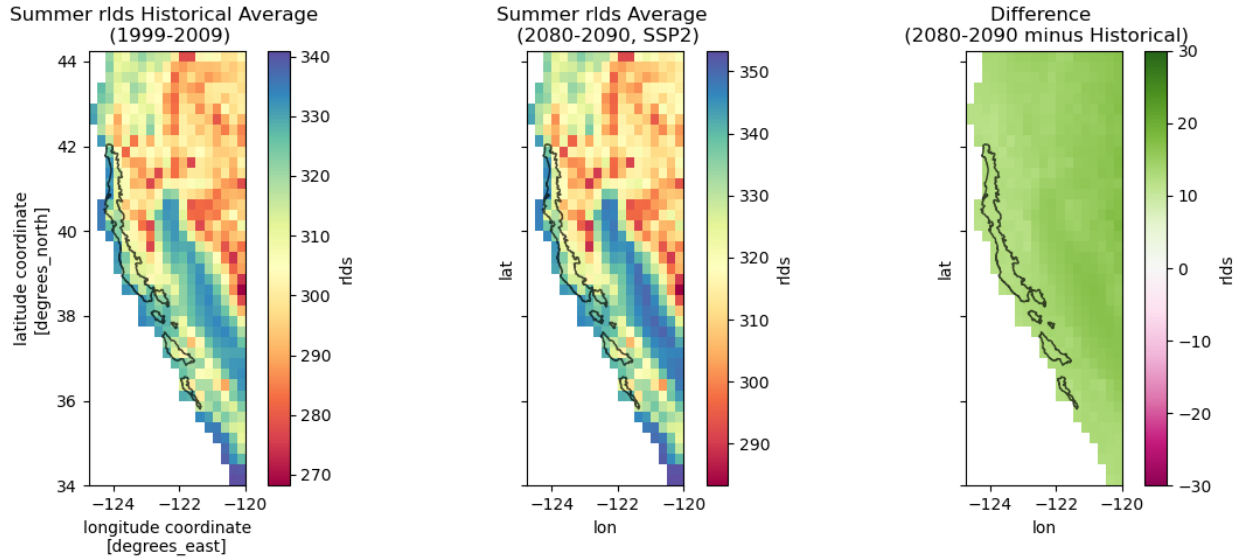


Figure B2: NEX-GDDP-CMIP6 Downwelling longwave radiation (W/m²) projections under SSP2-4.5. Averaged over 1999-2009 (left), averaged over 2080-2090 (center), and 2080-2090 average minus 1999-2009 average (right).

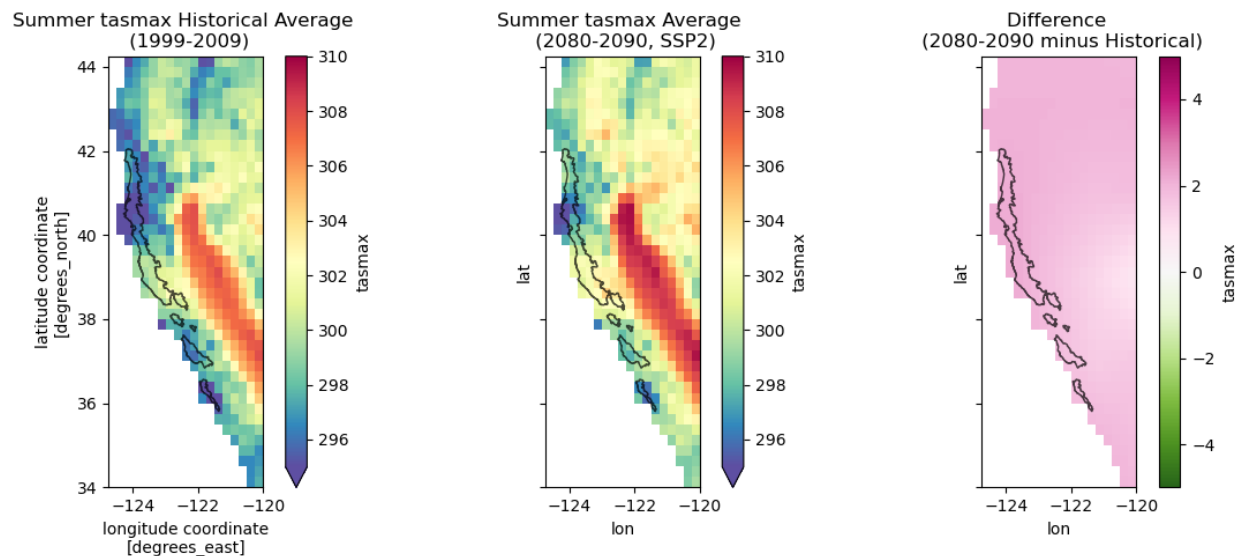


Figure B3: NEX-GDDP-CMIP6 Maximum daily near-surface air temperature (K) projections under SSP2-4.5. Averaged over 1999-2009 (left), averaged over 2080-2090 (center), and 2080-2090 average minus 1999-2009 average (right).

Appendix C: Scatterplots between fog and low cloud cover and climate variables

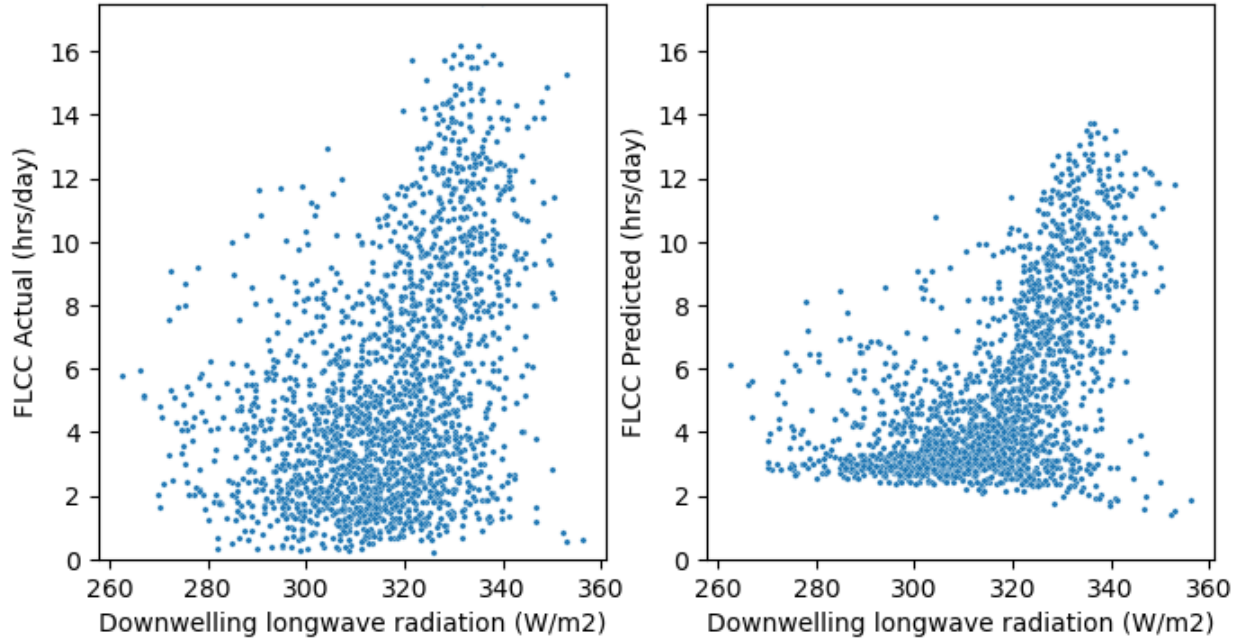


Figure C1: Effect of historical reanalysis daily averaged downwelling longwave radiation on fog and low cloud cover. Left shows a scatterplot of daily-averaged downwelling longwave radiation versus fog and low cloud cover hours using the Torregrosa et al. (2016b) data product. Right shows a scatterplot of daily-averaged downwelling longwave radiation versus fog and low cloud cover hours predicted by the random forest regressor.

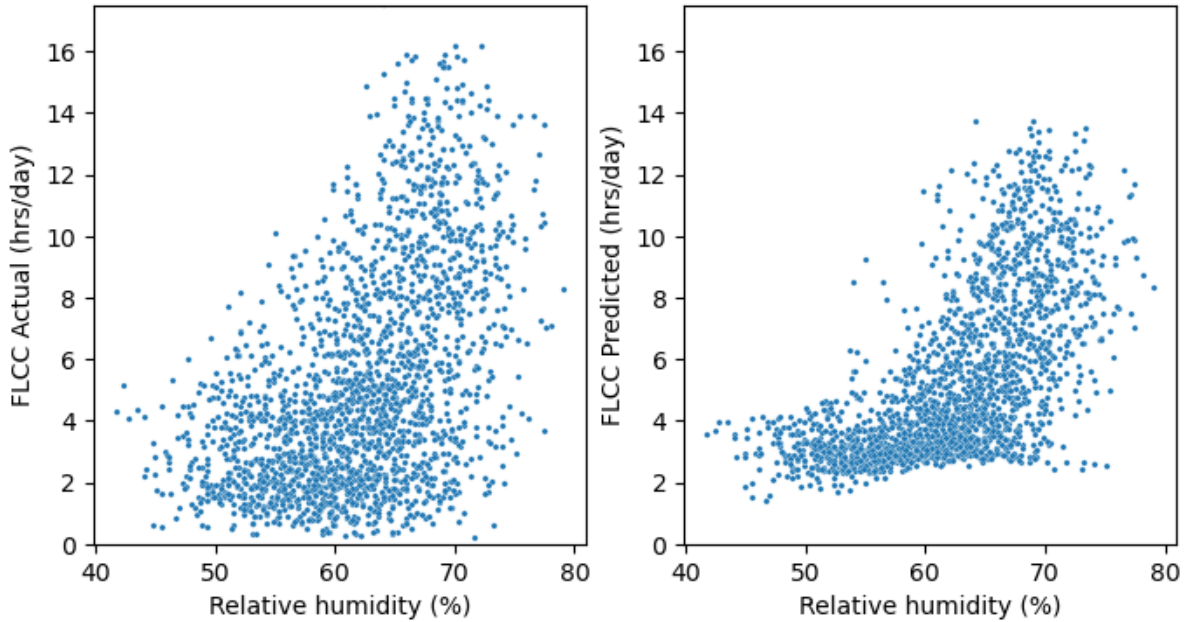


Figure C2: Effect of historical reanalysis daily averaged relative humidity on fog and low cloud cover. Left shows a scatterplot of daily averaged relative humidity versus fog and low cloud cover hours using the Torregrosa et al. (2016b) data product. Right shows a scatterplot of daily averaged relative humidity versus fog and low cloud cover hours predicted by the random forest regressor.

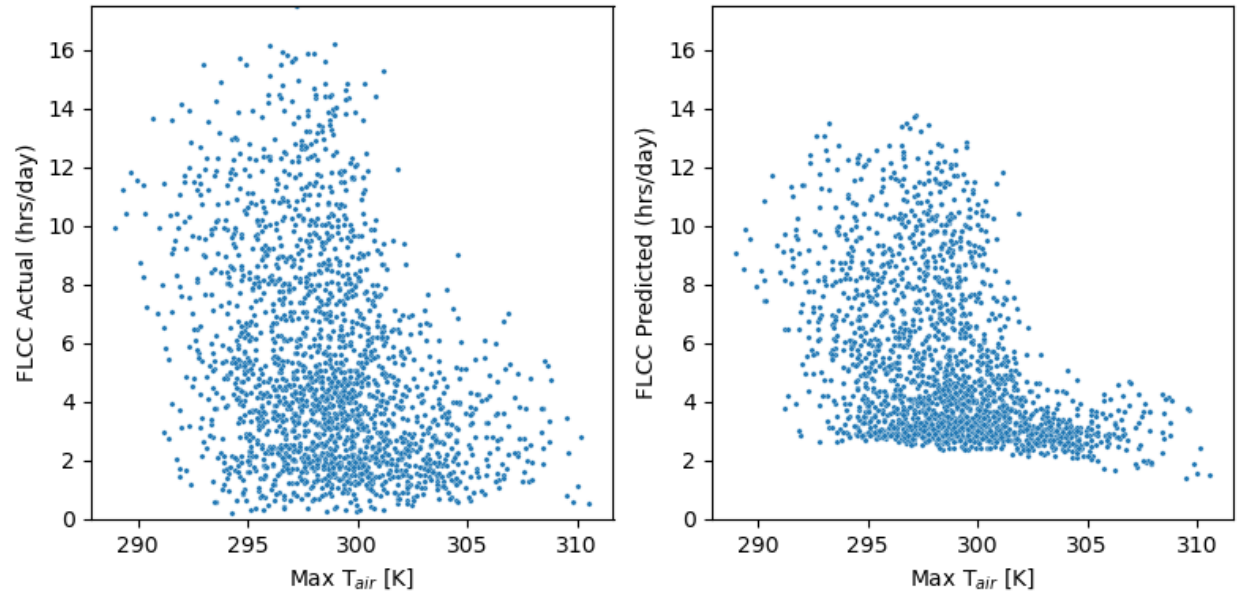


Figure C2: Effect of historical reanalysis daily maximum near-surface air temperature on fog and low cloud cover. Left shows a scatterplot of daily maximum near-surface T_{air} versus fog and low cloud cover hours using the Torregrosa et al. (2016b) data product. Right shows a scatterplot of daily maximum near-surface T_{air} versus fog and low cloud cover hours predicted by the random forest regressor.

Appendix D: Ground-based weather station map and labeled comparison plot

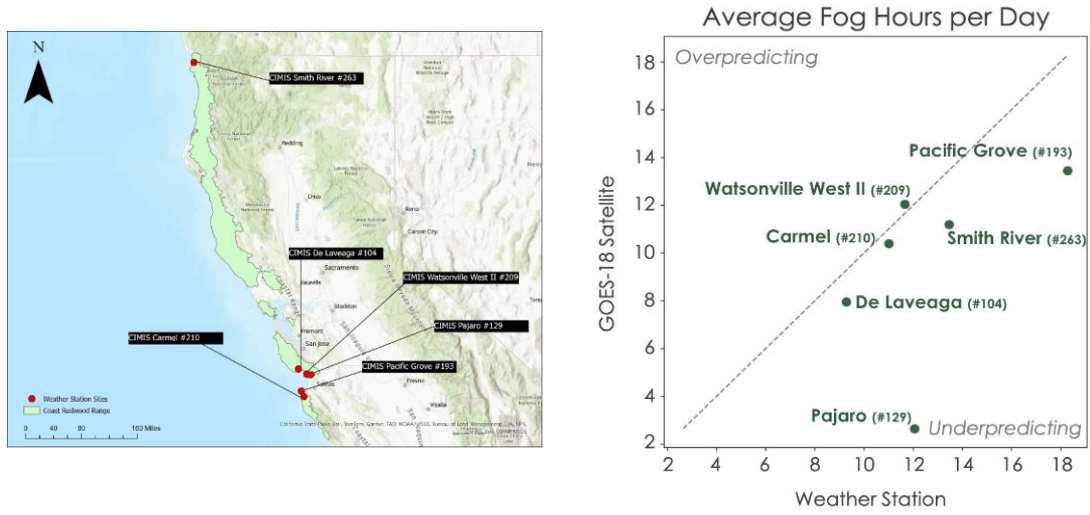


Figure D1: Left: Map showing locations of the 6 weather stations used for analysis. Right: Plot depicting average fog hours per day (June–September 2023–2024) according to the GOES satellite Cloud Top Phase product and the ground-based weather stations (labeled). Each point represents one weather station location. Locations along the dotted (one-to-one) line show agreement between the satellite method and ground-based observations, while locations further from the line show disagreement.

Appendix E: Native resolution random forest fog maps

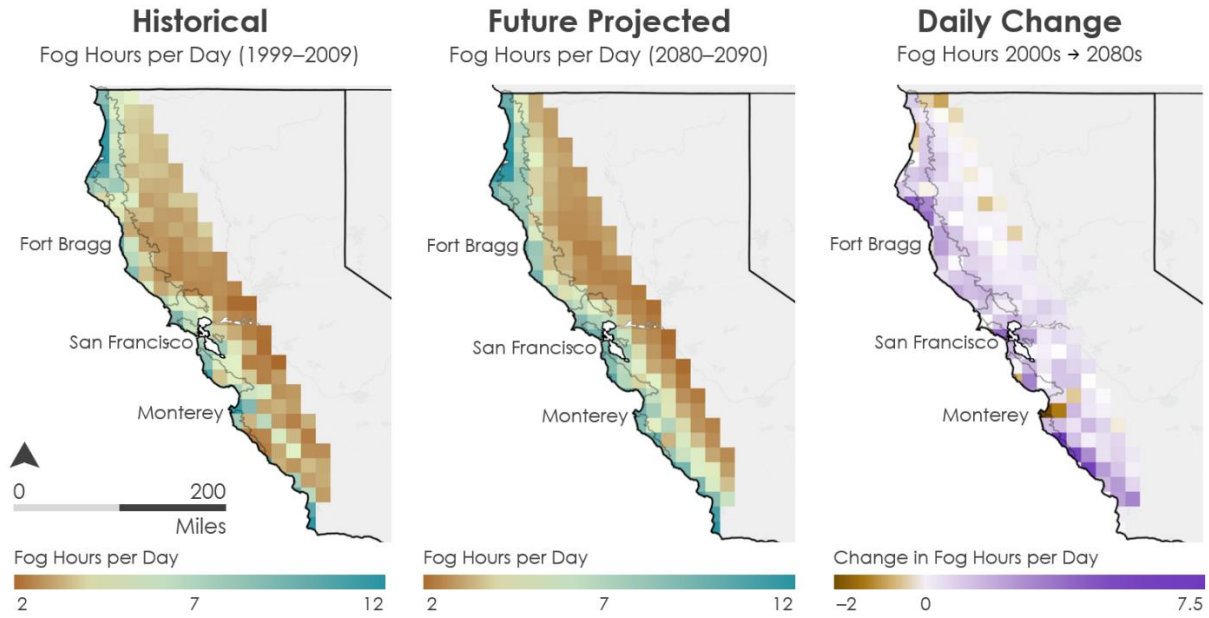


Figure E1. Native resolution (25km) maps of fog and low cloud cover over the coast redwood range, outlined in grey. Left shows the average daily hours of fog and low cloud cover from June–September 1999–2009 (Torregrosa et al., 2016). Middle shows the random forest’s predicted daily average hours of fog and low cloud cover in June–September 2080–2090. Right is the difference in the future and past fog maps, showing projected increases (purple) and decreases (brown).