Long Island Sound Study
A Partnership to Restore and Protect the Sound

Using Statistics to Predict Embayment Water Temperature in the LIS

Ben Lawton, EPA Region 2 ORISE Fellow August 2022-Sept. 2023

Background

- Research began with the **Data Visualization and Analytics Challenge**
  - Program launched by the Office of Mission Support
  - **Goal:** development of new tools for sharing data and analysis

- Team consisted of
  - Cayla Sullivan (LISS/EPA R1)
  - James Ammerman (LISS/NEIWPCC)
  - Melissa Duvall (LISS/EPA R1)
  - Phil Colarusso (EPA R1)
  - Nathaniel Merrill (EPA Atlantic Ecology Division/ORD)
  - Darryl Keith (EPA Atlantic Ecology Division/ORD)
  - And… myself

- Underlying code for model can be seen as github project at https://github.com/blawton/long-island-sound-gpr
  - Should be updated in the future for usability/clarity

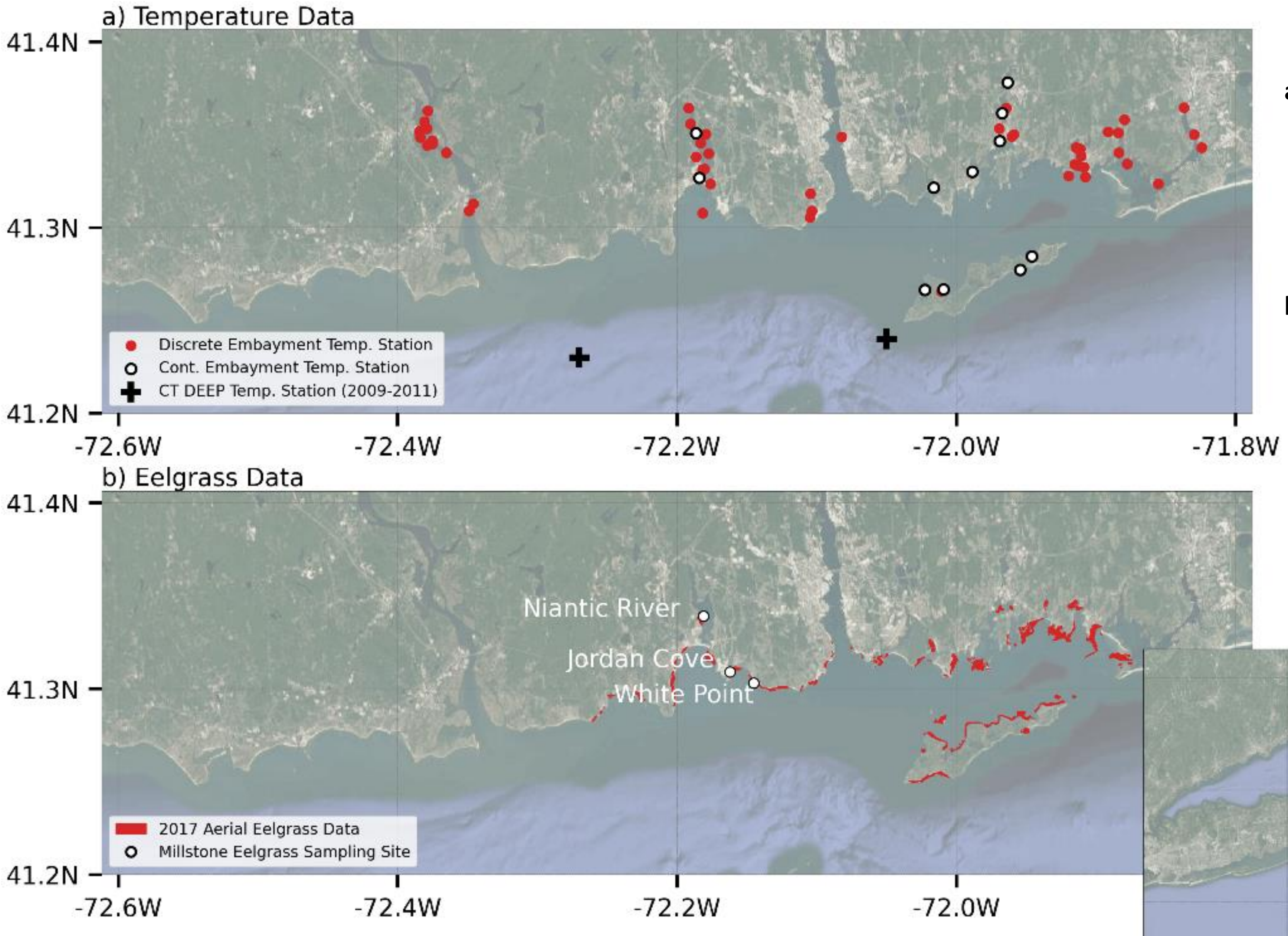
Research Question/Problem

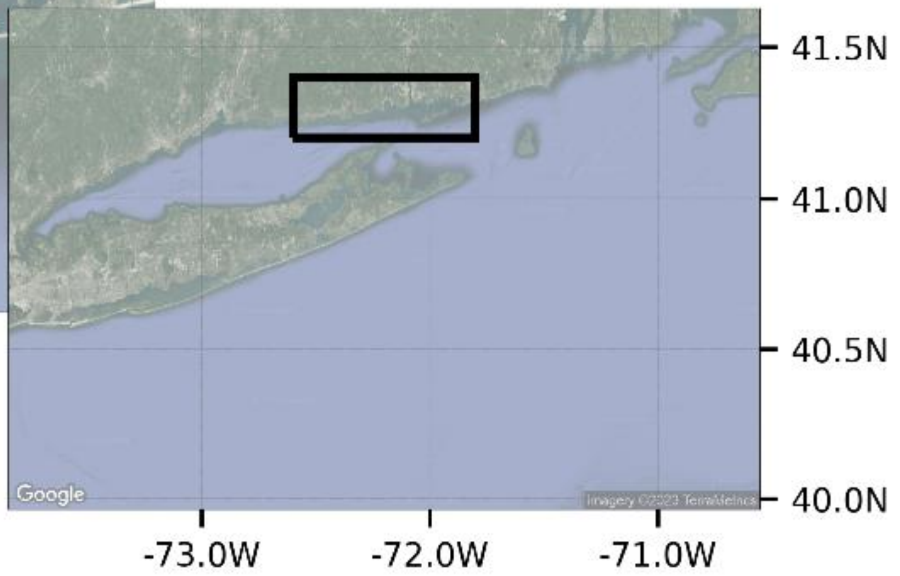
Long Island Sound Study
A Partnership to Restore and Protect the Sound

Most temperature monitoring in the LIS has been carried out with temperature sensors located in the open sound and operated by CTDEEP

- An exception to this is the temperature monitoring carried out for the last 40+ years in the Thames River estuary by Project Oceanology
- This has resulted in not capturing a significant amount of variation from temperatures in embayments
- **BUT** most eelgrass in or near embayments, and this is where significant concern over temperature arises
- Additionally, most eelgrass that is currently studied in the Long Island Sound is in the Eastern waters
    - More water clarity
    - After the wasting disease Labyrinthula Zosterae decimated LIS eelgrass populations in the 20th century, it only recovered in the Eastern Sound

Comparison of Temperature Monitoring and Eelgrass

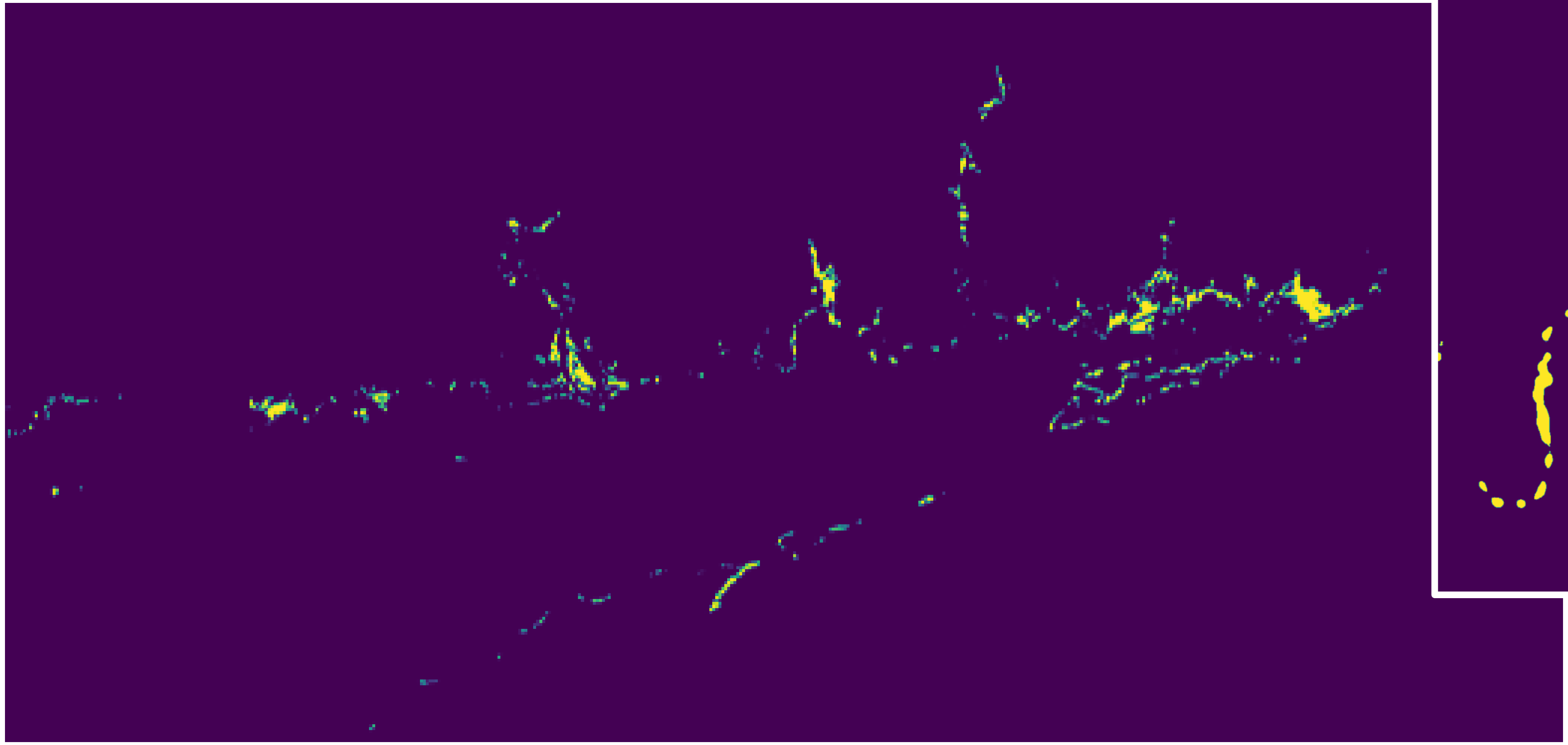




a) Discrete (red dot) and continuous (black circle) embayment monitoring stations as well as open water CTDEEP stations (plus signs)

b) Eelgrass monitoring sites (white dots) in Niantic River (NR), Jordan Cove (JC), and White Point (WP). Red areas are eelgrass extent based on 2017 aerial survey (Bradley and Patton, 2018)

With a satellite algorithm, we see similar trends in eelgrass distribution



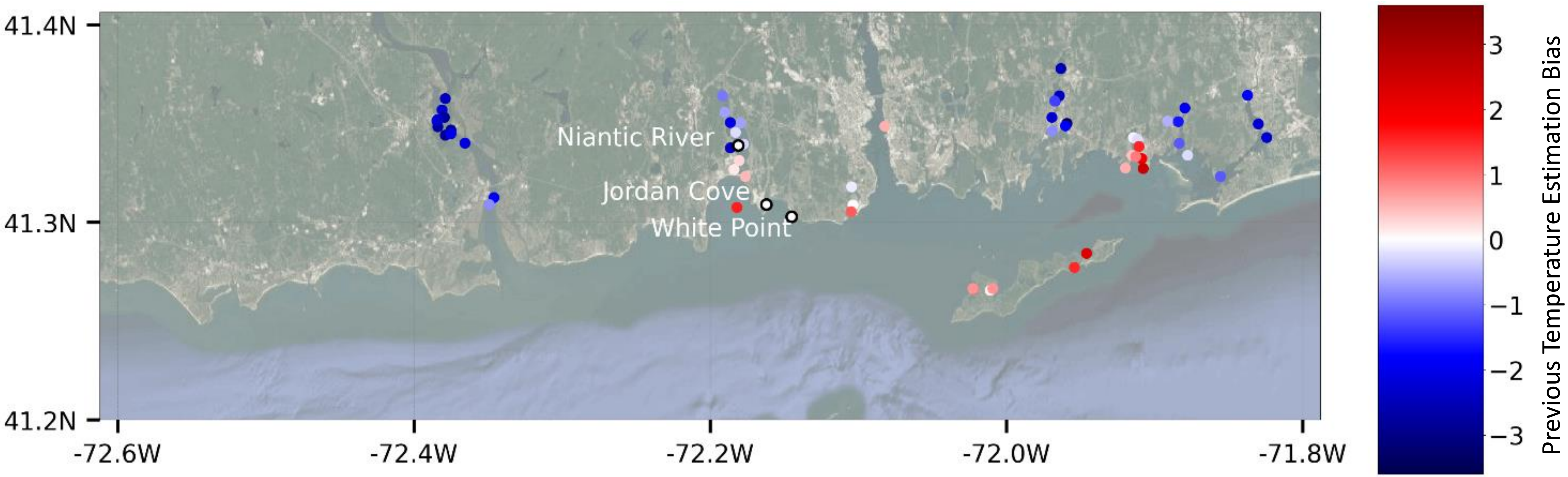
a) Overall LIS

b) Niantic detail

- Goal of the satellite algorithm is to identify meadows in locations not monitored for eelgrass
- Credit to Nate Merrill and Daryl Keith at EPA ORD for the development of the satellite algorithm

As a result of being geographically isolated from eelgrass meadow locations, CTDEEP temperature measurements pick up a significant amount of bias

- Shown below is the bias of inverse distance weighting interpolation of open sound CTDEEP data to estimate temperature at embayment measurement stations
- We focus on the Eastern Sound, where most of the measurement stations are located



Research Problems and Goal (cont.)

- **Another challenge is that most of the temperature data in the Long Island Sound is sampled a maximum of once every two weeks**
  - Can try time series methods, but this ignores the possibility of using continuous data to **spatially interpolate** to discrete sampling locations

Potential bias of using data from different regions of the estuary

Challenge of drawing conclusions from continuous and discrete data

GOAL: use a Gaussian Process to statistically interpolate an embayment heatmap

Overview of Data Coverage

Long Island Sound Study
A Partnership to Restore and Protect the Sound

- **For the embayment model, we rely on the following temp. data providers:**
  - Save the Sound Tier I monitoring (discrete)
  - Save the Sound Tier II monitoring (continuous)
  - Fisher’s Island Seagrass Management Coalition
  - University of Rhode Island
  - USGS in various upstream embayment locations in the sound (discrete)
  - USGS in the mystic river area (continuous)

Overview of Data Coverage (cont.)

- The following are counts of the monitoring stations that recorded temperature each year from 2019 to 2021 in the Eastern Sound Window
- The Eastern Sound Window is defined as being from -72.6 degrees to -71.81 degrees longitude and from 40.97 to 41.54 degrees Latitude

	Continuous			**Discrete**			
	Fisher's Island	STS Cont.	USGS Cont.	Millstone Env. Lab	STS Discrete	URI	USGS Discrete
Year:							
2019	4	2	0	2	31	16	0
2020	4	2	0	2	31	16	8
2021	4	2	3	2	27	16	6

Differences in Data Interpolation Methods

Old Model -

- Inverse Distance Weighting of Temp. Data
- Extrapolation as opposed to Interpolation (different geography)
- All data was discrete with measurement freq. ~/2 weeks

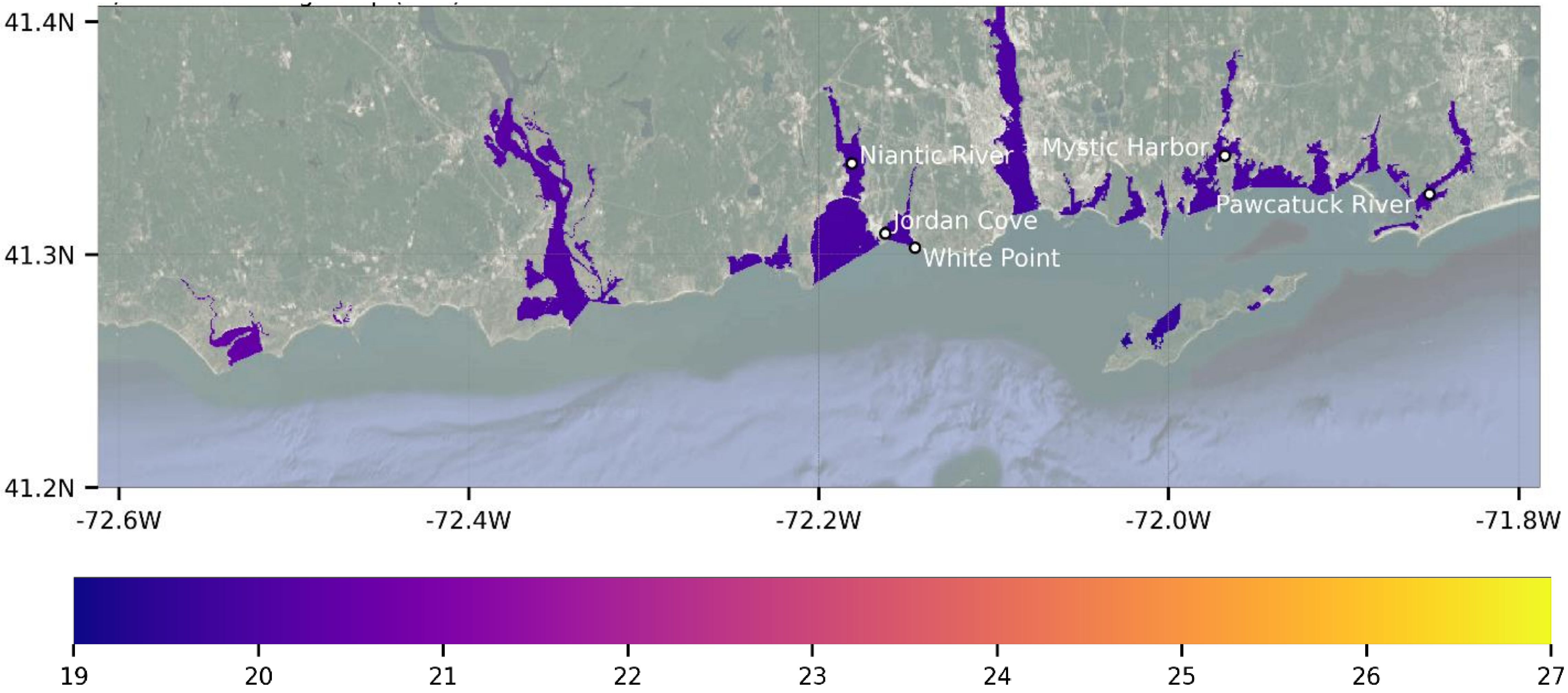
New Model -

- Gaussian Process, nonparametric machine learning method
- Uses coastal features (dist. into embayment) to predict
- Leverages continuous data to interpolate discrete data

Results

Visual Results and Comparison of Distributions

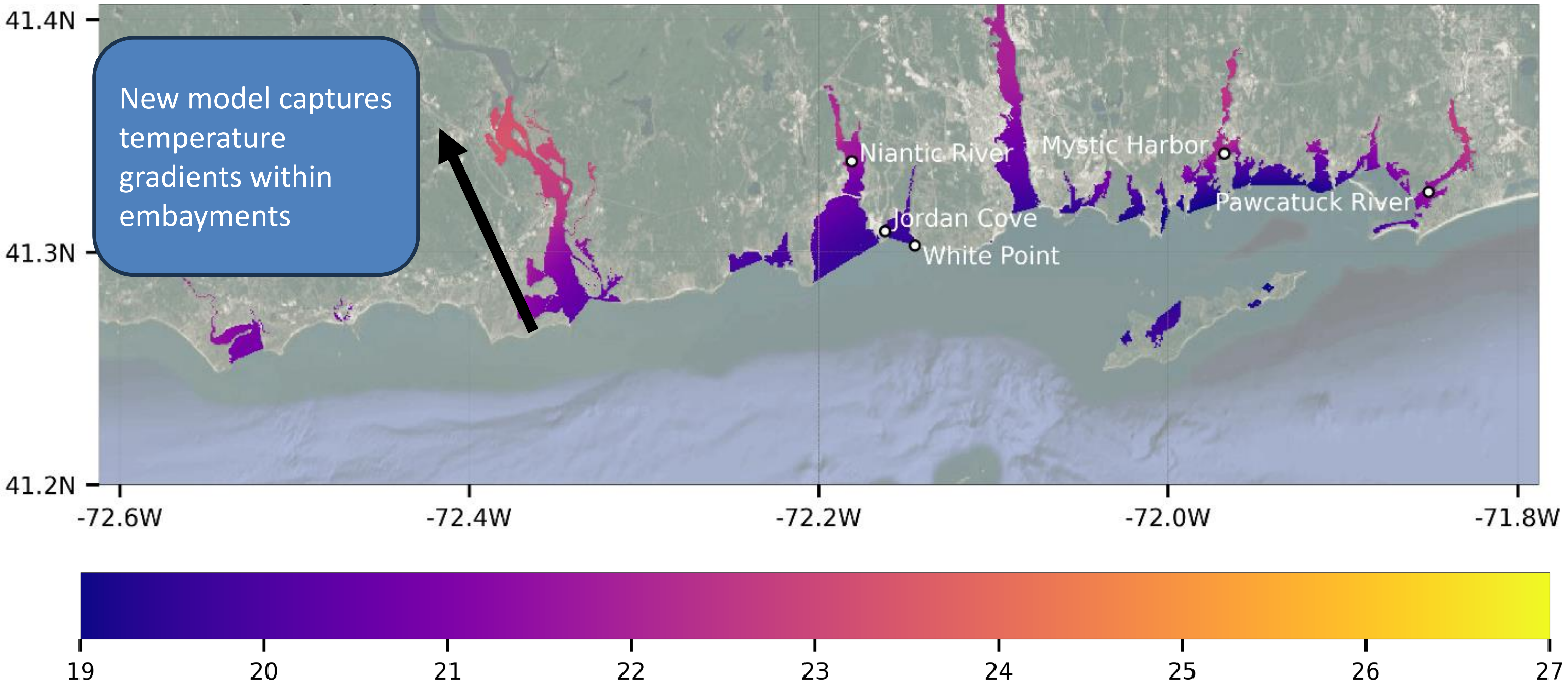
Visual Results



Heatmap from Previous Model (Avg of 2019-2021 July/August temp.)

- **Uses measurements 2-3m deep**
- **Very little variation within embayments**

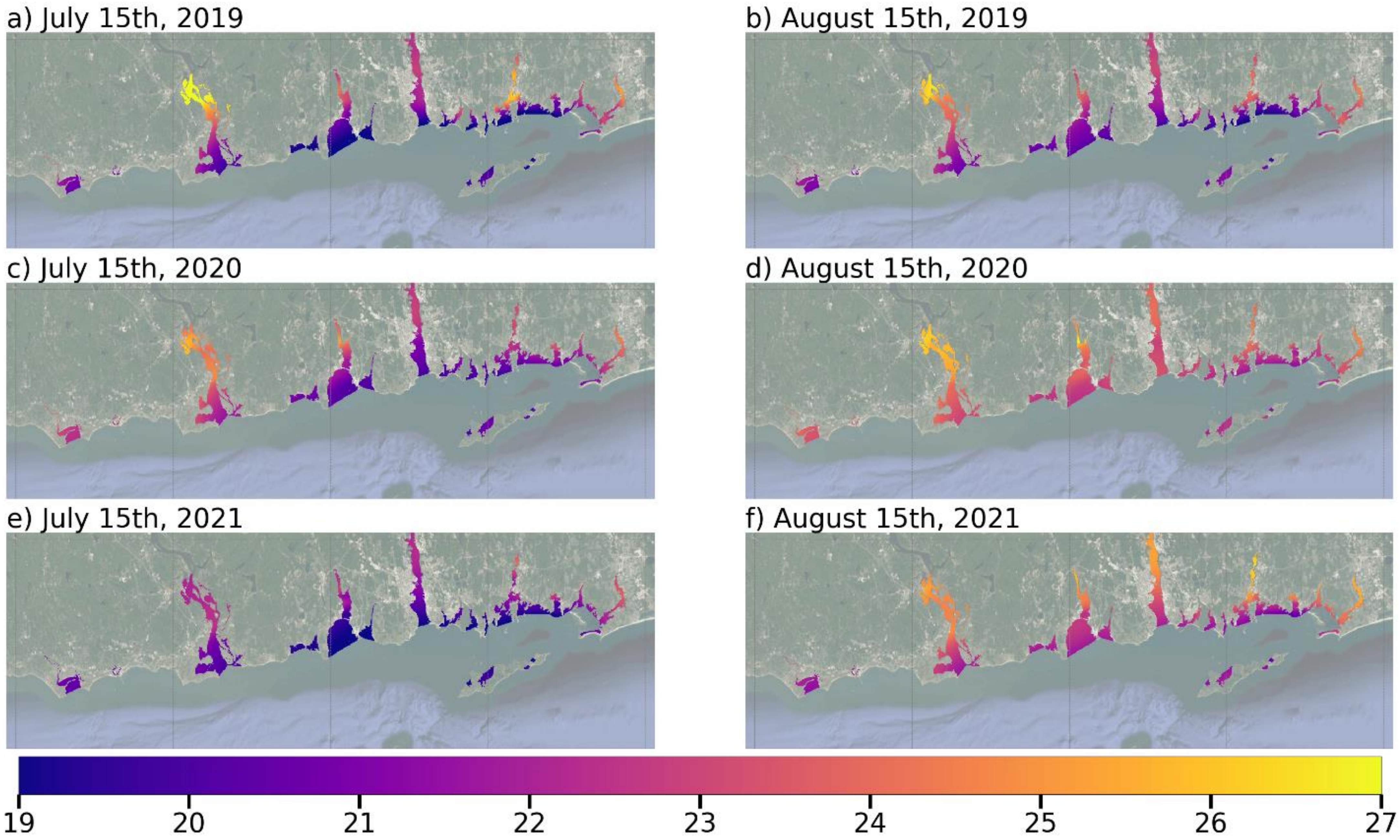
Visual Results



Heatmap for new model (Avg of 2019-2021 July/August temp)

- Uses bottom water temperatures within embayments
- Average of 2019-2021 July/August data

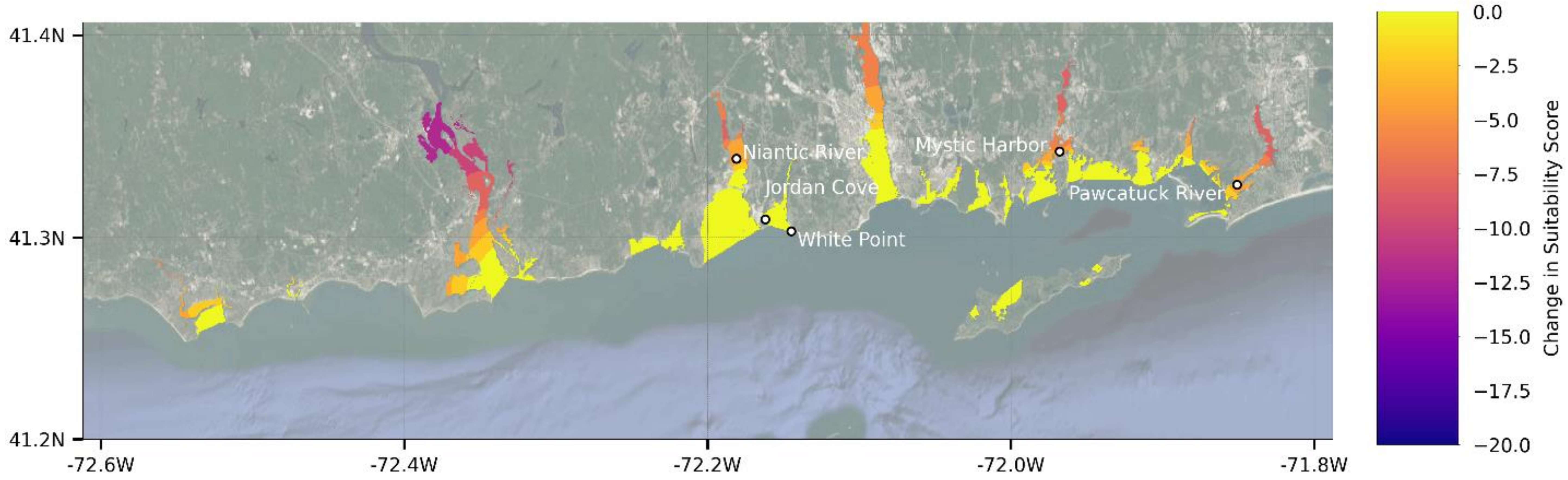
Visual Results



GP makes predictions for any day of year (DOY) within range of training data

- June 1st to October 1st
- Captures spatial **AND** temporal variance

Visual Results



Increase/Decrease in EHSI Temperature score from 2013 index from embayment data & Gaussian Process

- **Decreased score of 5 – 7.5** observed for Niantic River, Mystic Harbor, Pawcatuck River
- **Corresponds to approximately +2^{o} C change in temperature**

We assess the accuracy of our model using:

- Alignment of dist. of summer average temperatures with underlying data
  - Histogram
  - Average of all days in range for each monitoring station
  - Captures spatial variation of mean
- Days over temperature threshold comparison
  - Captures temporal variation (extreme temperatures) among all monitoring locations

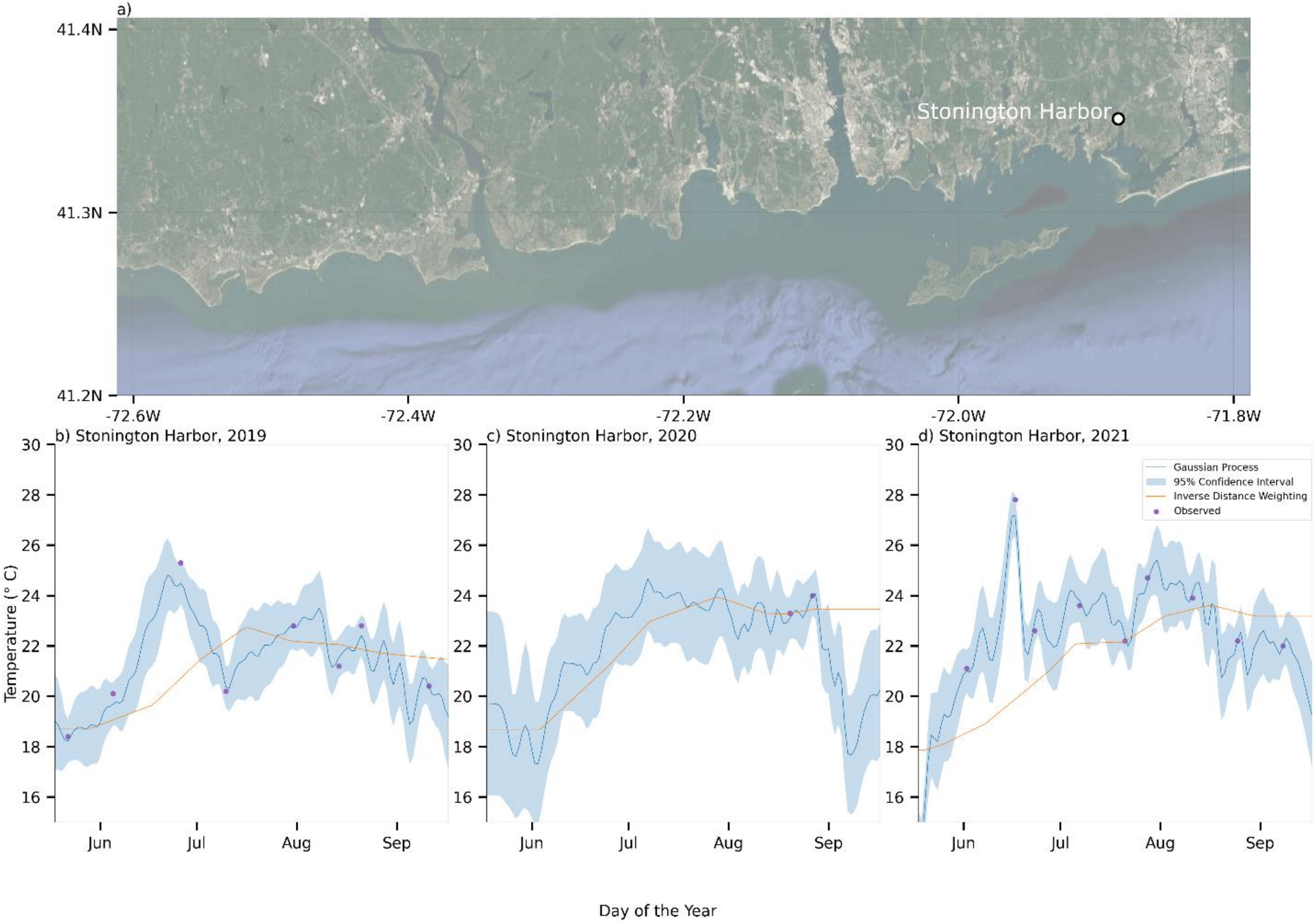
Carried out over entire range of ground truthing season: DOY 162-268

- Visual Examination of predicted time series
- RMSE of prediction on cross-validated data
  - Compared directly to RMSE of extrapolating CTDEEP temp. to embayments
  - Gaussian Process shows **45.03% improvement over the IDW interpolation**

Carried out over July/August for comparison with Vaudrey EHSI: DOY 1182-243

Time Series

Long Island Sound Study
A Partnership to Restore and Protect the Sound

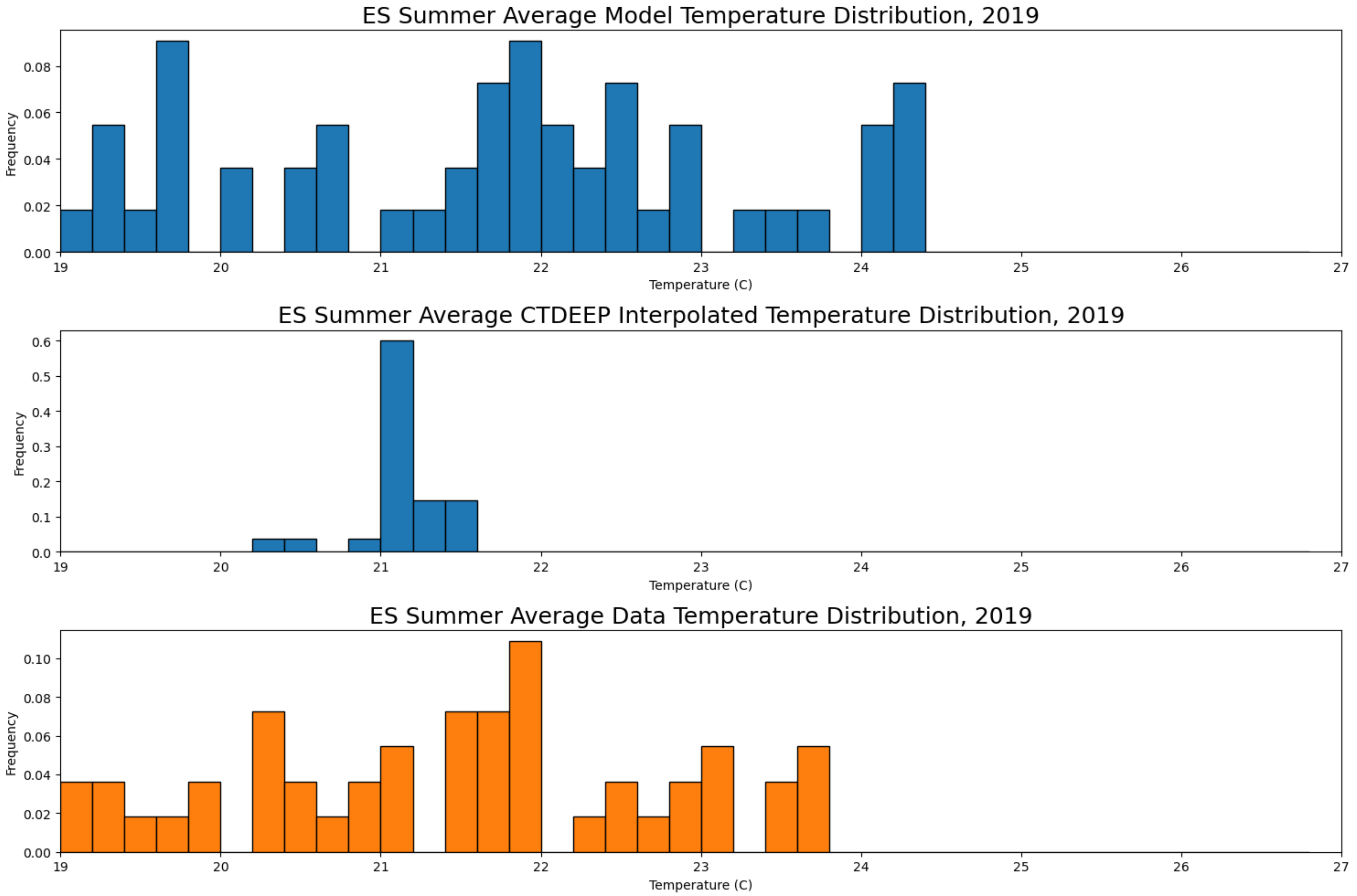


Time series for individual locations can be compared with:

- CT DEEP data, orange line (linearly interpolated)
- Underlying discrete data, purple points
- Shaded region is 95% confidence interval

Alignment of Avg Temp. Histograms (2019)

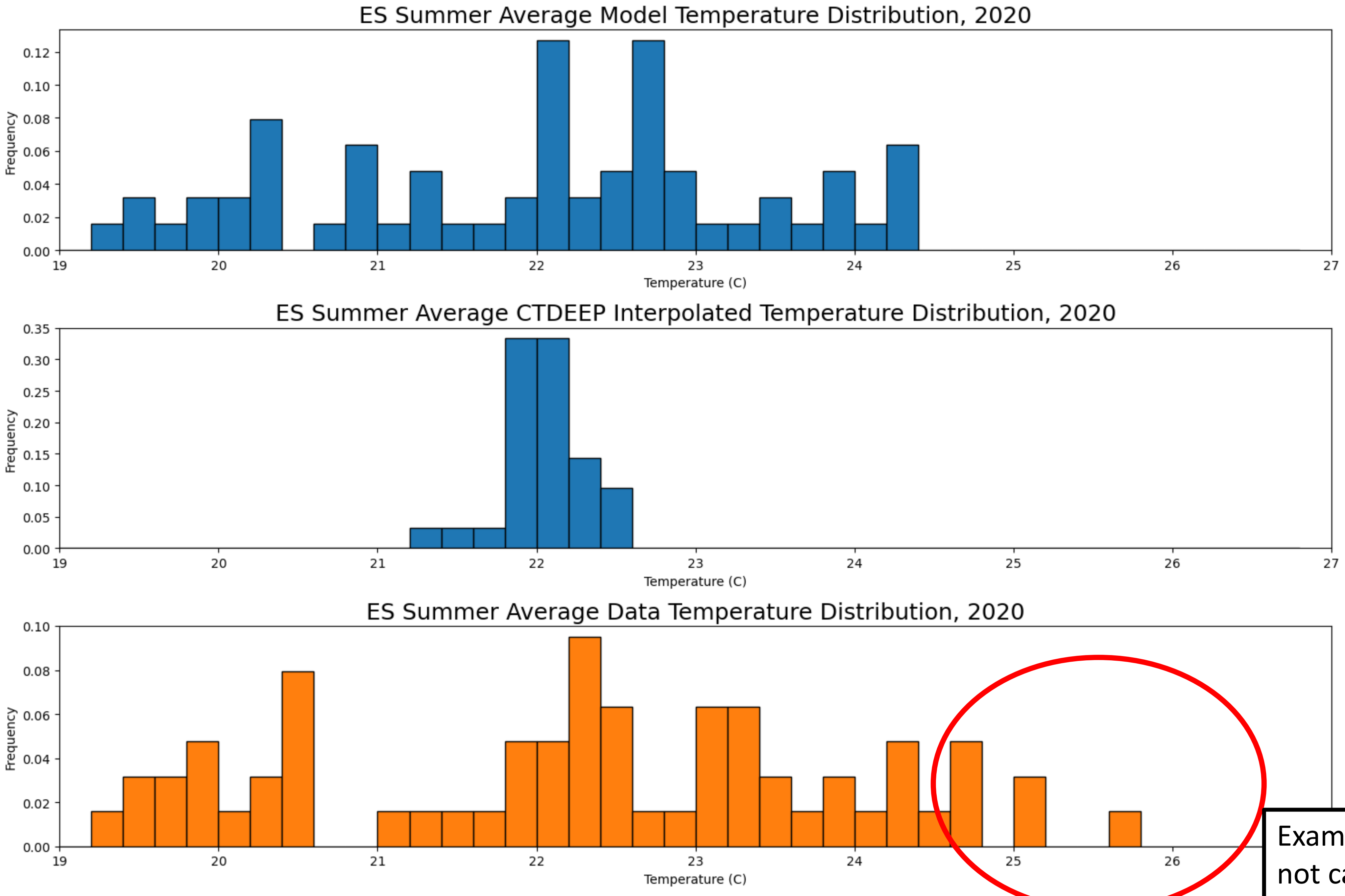
Long Island Sound Study
A Partnership to Restore and Protect the Sound



2019

- Gaussian Process Effectively Captures Temp. Distribution between stations
- CTDEEP Interpolation Collapses Distribution

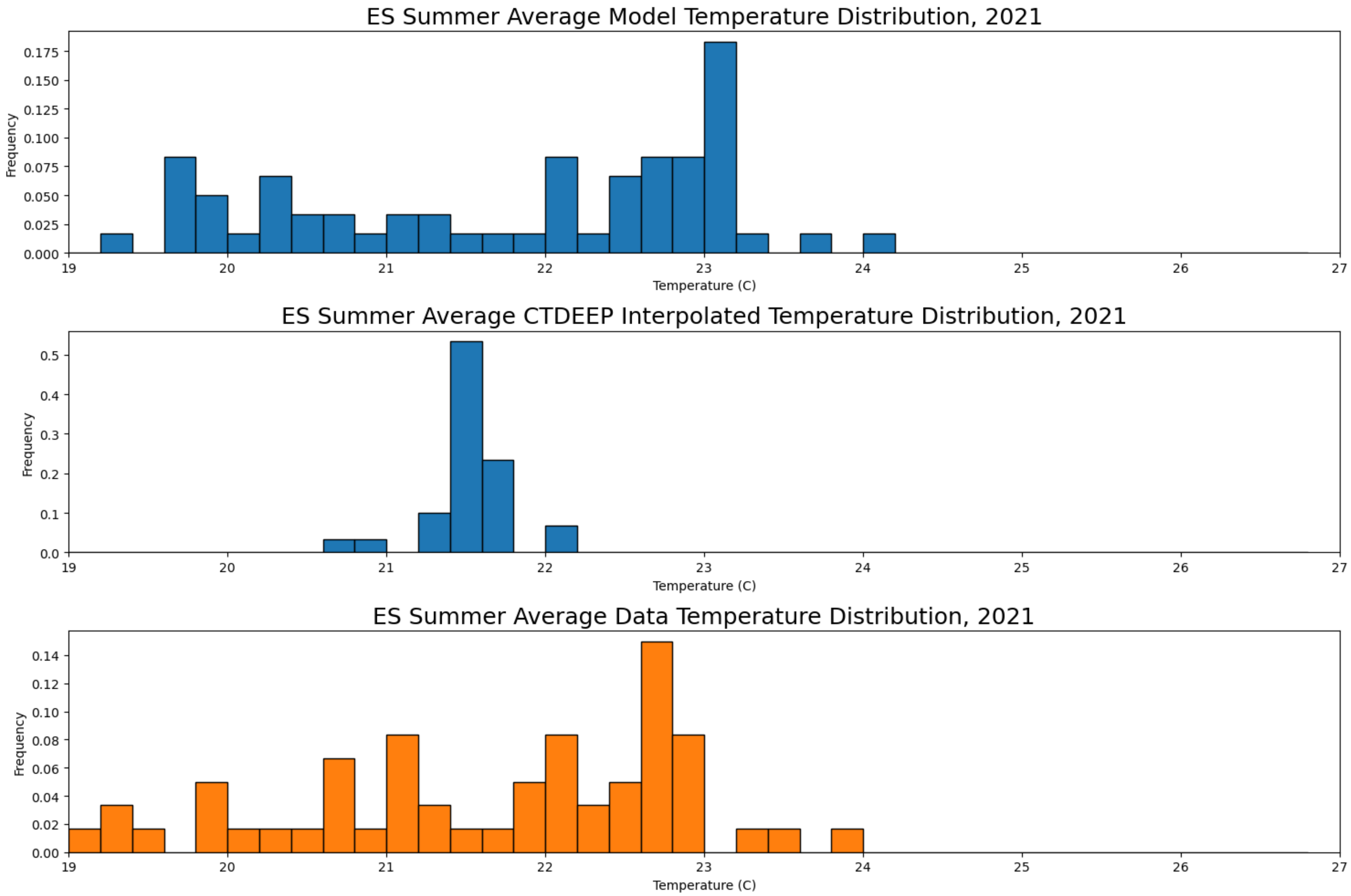
Alignment of Avg Temp. Histograms (2020)



2020

Example of extrema not captured by Gaussian Process Model

Alignment of Avg Temp. Histograms (2021)



2021

Eelgrass Temperature Thresholds

To try to assess metrics most related to eelgrass success, we also analyzed the days over temperature thresholds (as opposed to summer avg. temp.)

- 23°C and 25°C Temperature thresholds were determined from literature as the threshold for the cessation of growth, and die-offs, respectively (Moore et al., 2013; Reusch et al. 2005; Greve et al., 2003)
  - **Temperature thresholds vary by geography and study**, so refinement is important

- The area that experienced stretches of time above these thresholds was **low, but not insignificant**

- A marine heatwave (MHW) was defined following the literature (Magel et al., 2022) as 5 consecutive days over 90th percentile of temp
  - For LIS, 90th percentile = 24.75 °C
  - In modeled estuaries, 9.5-10 km2 of estuarine area experienced MHW
  - Between 15 and 16% of area

Geographic Variability of Optimal Eelgrass Temp.

Long Island Sound Study
A Partnership to Restore and Protect the Sound

Location	Country	Latitude	Optimal Eelgrass Temp. (° C)
Bahia Todos Santos	Mexico (lagoon)	31°45′N	17
Bahia Todos Santos	Mexico (open coast)	31°45′N	21
San Diego Bay	USA	32°36' N	<20
Mission Bay	USA	32°48'N	<20
Newport River estuary	USA	34°N	22
Chesapeake Bay	USA	37°16′N	19

Location	Country	Latitude	Optimal Eelgrass Temp . (° C)
York River (Chesapeake Bay)	USA	37°24′N	22
Great Harbor	USA	40°31.5′N	16–20
Great Harbor	USA	40°31.5′N	25–30
Coos Bay	USA	43°24'N	15
Willapa Bay	USA	46°30′N	15
Puget Sound	USA	47°54'N	5-8
Izembek Lagoon	USA (subtidal)	55°N	30
Izembek Lagoon	USA (intertidal)	55°N	35

Geographic Variability of Optimal Eelgrass Temp.

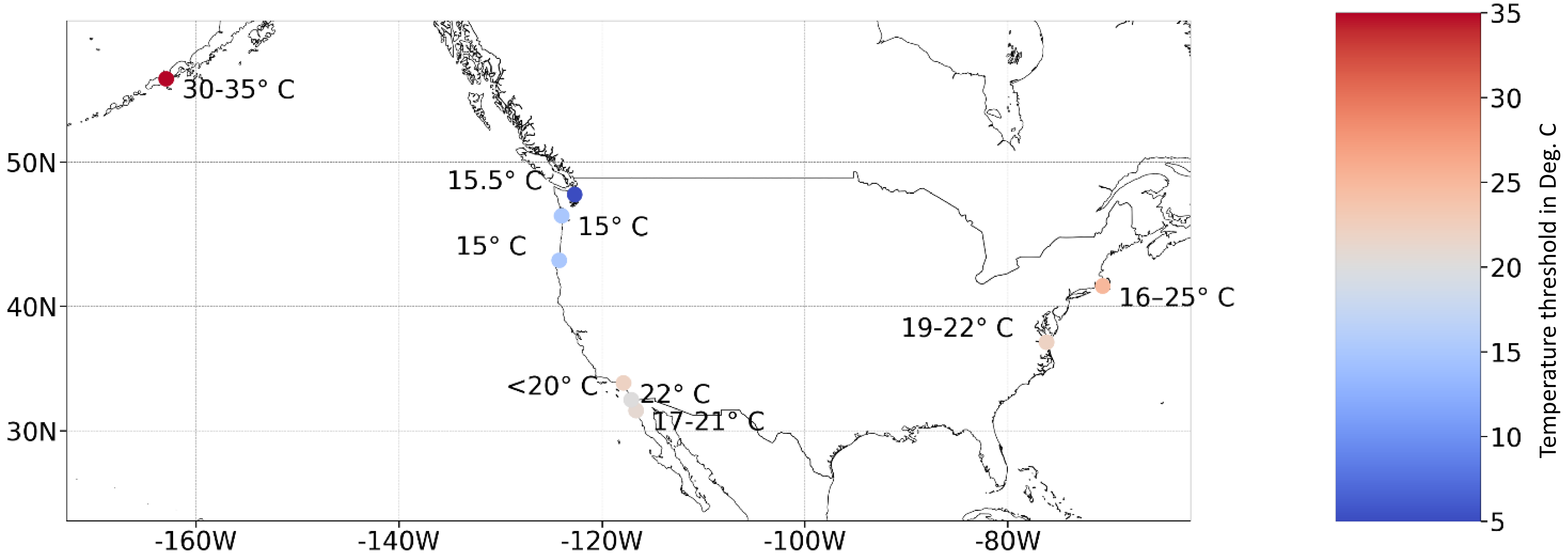


Diagram showing the temperature thresholds from the previous slide on a North American map

Predicted effect of Updated Temp. on Eelgrass

In addition to updating the temperature heatmap and the EHSI, we modelled eelgrass sensitivity to differences in summer average

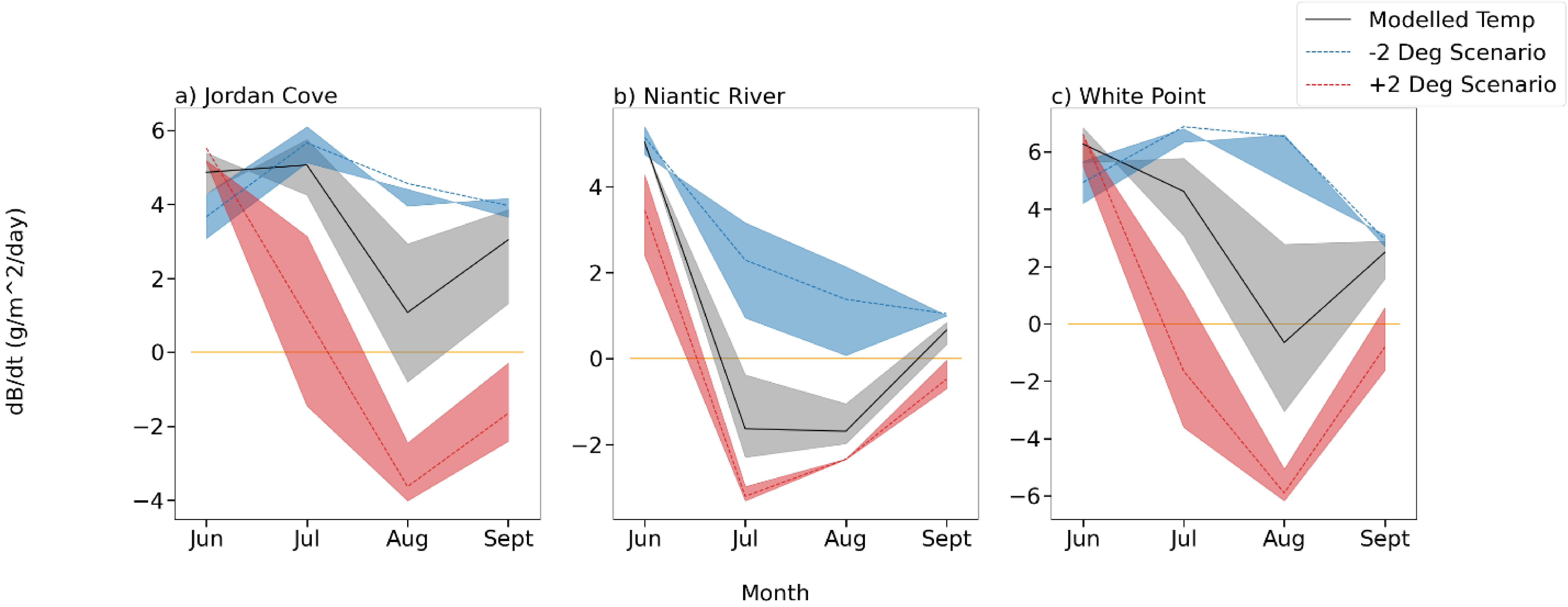
- A model was chosen from literature based on data availability and previous result (Zharova et al., 2001)
  - Model outputs biomass production at different levels of temperature and light availability
  - Model required an optimal temperature, so based on previous slides 20-22 °C were trialed
- Three sites included, w/ data from Millstone Environmental Lab. at Dominion Power Plant (thanks to Stephen Dwyer)

Results

- Areas not currently at risk for a mid-summer decrease in biomass/die-off will be in the 2° warming scenario
- 2° difference between embayment temperature model and previous model can have a substantial diff. in predicting growth

Predicted effect of Updated Temp. on Eelgrass

Long Island Sound Study
A Partnership to Restore and Protect the Sound



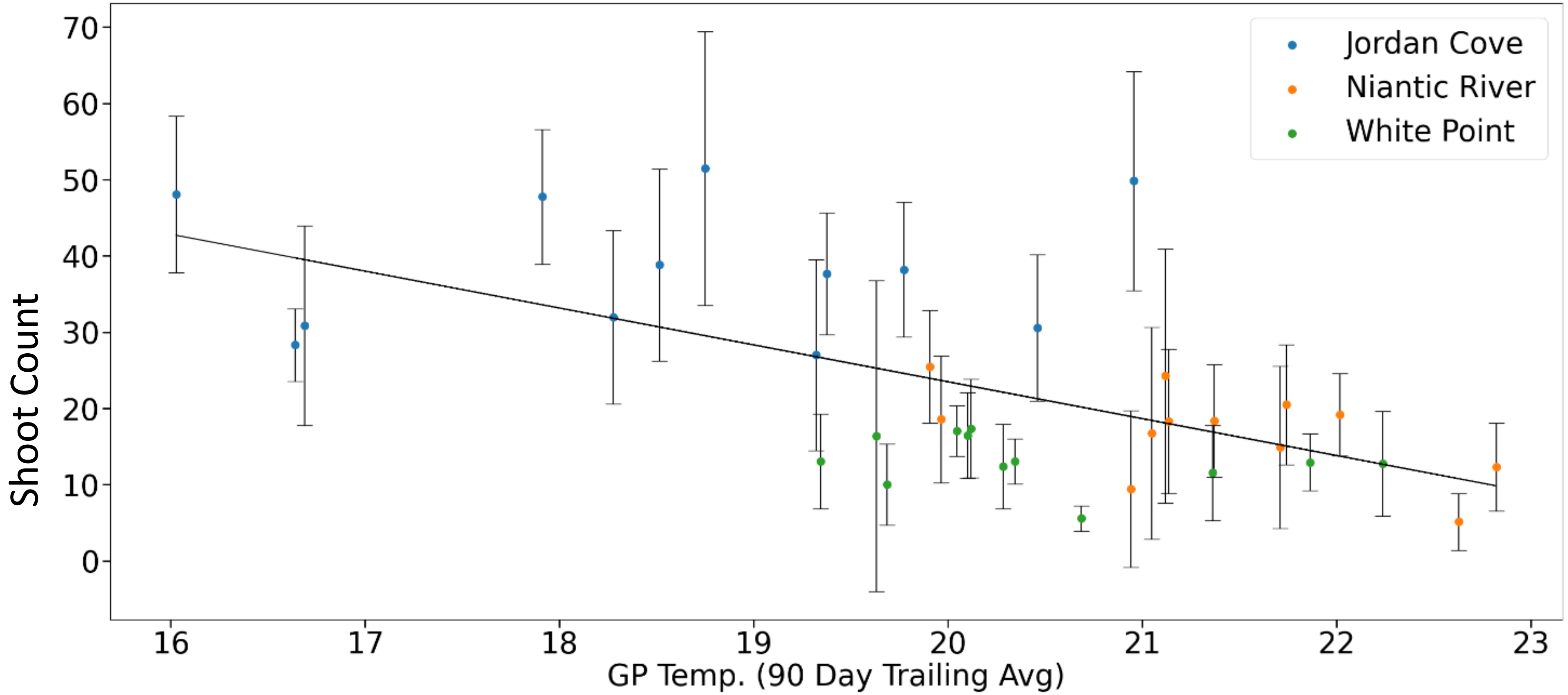
Eelgrass biomass growth rates modelled for three sites in the Niantic River/Jordan Cove area. Shaded regions represent the range of growth rates predicted for temperature optima between 20 and 22°C

Using Ground Truth-ed Data

Models for eelgrass growth can only go so far

- Data for the period 2019-2021 is limited (only includes the three Millstone Env. Lab. sites)
  - EPA monitoring of Mumford and Beebe Cove did not begin until 2022
- Temperature data not taken exactly at sites → Opportunity to use Gaussian Process to analyze impact of temp. diff. between sites
- Trailing 90 day average used to reflect the influence of temperatures on Eelgrass over long periods of time

Using Ground Truth-ed Data (cont.)



Results of regression between 90-day trailing average of Gaussian Process temp. prediction and shoot density measured by Millstone Env. Lab

- Decrease in shoot density at Niantic River site corroborated by higher temperatures
- Niantic River site is also **furthest from open sound**

Data Visualization and Analytics Challenge Storymap

Long Island Sound Study
A Partnership to Restore and Protect the Sound

Question:

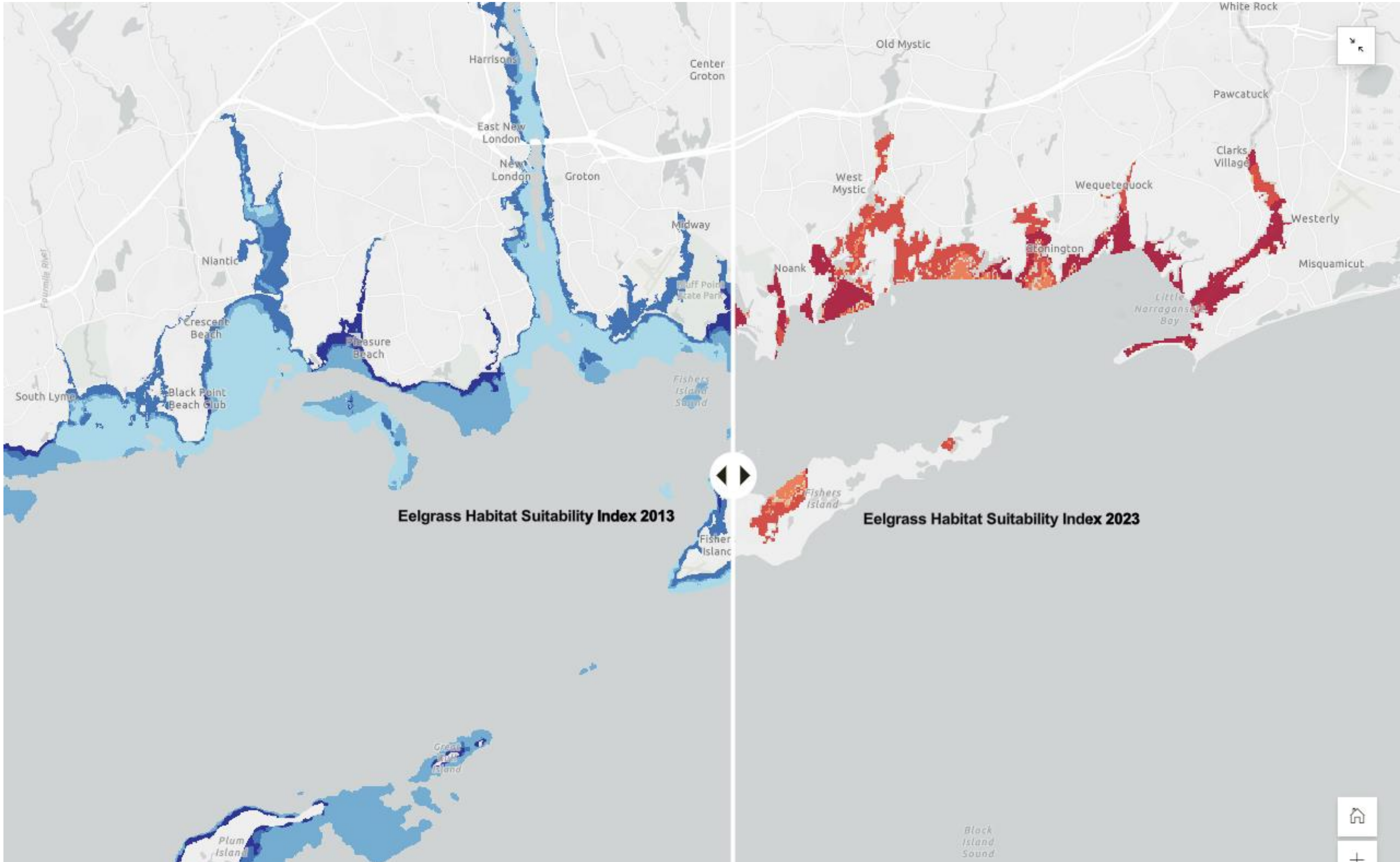
How do we present these results to the public in an engaging and educational way?

Answer:

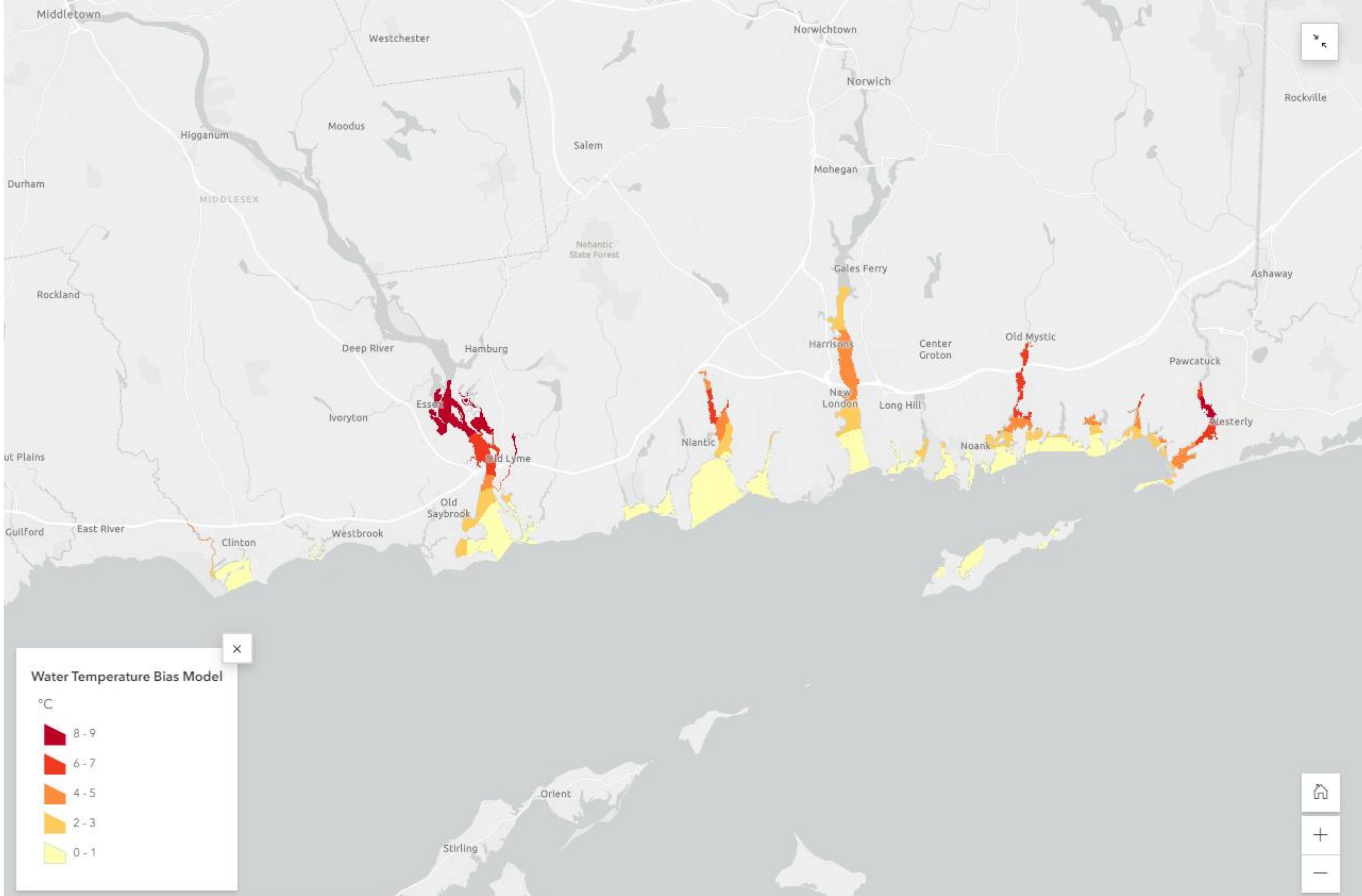
- An ArcGIS Stroymap created by EPA R2 Scientist Cayla Sullivan
- Public-facing tool to educate on advances in **water quality data** and **eelgrass habitat**
- In final rounds of review, will be released online soon and can be found at https://longislandsoundstudy.net/category/media-center/

Data Visualization and Analytics Challenge Storymap

Long Island Sound Study
A Partnership to Restore and Protect the Sound



Data Visualization and Analytics Challenge Storymap



A hearty thanks to:

- Melissa Duvall and Jim Ammerman in the Long Island Sound Study for helping structure research/write a paper
- Cayla Sullivan for putting together an amazing StoryMap to show our data
- Jamie Vaudrey at UCONN for helping to understand the 2013 EHSI construction
- Everyone who helped aggregate the data used in this work, especially Stephen Dwyer (Millstone Env. Lab), Katie O’Brien Clayton (CTDEEP), Matt Lyman (CTDEEP), Peter Linderoth (STS), and Hannah Vagts (FISM)

Citations:

Greve, T.M., Borum, J., Pedersen, O. (2003). Meristematic Oxygen Variability in Eelgrass (Zostera marina). *Limnol. Oceanogr. 48(1),* 210-216.

Magel, C., Chan, F., Hessing-Lewis, M., Hacker, S. (2022). Differential Respondes of Eelgrass and Macroalgae in Pacific Northwest Estuaries Following an Unprecedented NE Pacific Ocean Marine Heatwave. *Front. Mar. Sci., 9.*

Moore, K., Shields, E., & Parrish, D. (2013). Impacts of Varying Estuarine Temperature and Light Conditions on Zostera marina (Eelgrass) and its Interactions With *Ruppia maritima* (Widgeongrass). *Estuaries and Coasts, 31*, 20-30.

Reusch, T. B., Ehlers, A., Hämmerli, & Worm, B. (2005). Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings Of The National Academy Of Sciences, 102(8),* 2826-2831.

Zharova, N., Sfriso , A., Voinov, A., Pavoni, B. (2001). A simulation model for the annual fluctuation of Zostera marina biomass in the Venice lagoon. *Aquatic Botany*, *70*, 135-150.