Contract No: 21-060-025-D274
Integrative Assessment of Bacterial Pollution

Final report

Submitted to General Land Office (GLO)

October 10, 2023

By

Gabriele Bonaiti (gbonaiti@tamu.edu)
Anish Jantrania
Terry Gentry

Biological and Agricultural Engineering
Texas A&M AgriLife Extension Service

"This report was funded in part by a Texas Coastal Management Program grant approved by the Texas Land Commissioner, providing financial assistance under the Coastal Zone Management Act of 1972, as amended, awarded by the National Oceanic and Atmospheric Administration (NOAA), Office for Coastal Management, pursuant to NOAA Award No. NA20NOS4190184. The views expressed herein are those of the author(s) and do not necessarily reflect the views of NOAA, the U.S. Department of Commerce, or any of their subagencies."

Executive Summary

The General Land Office (GLO) contracted with Texas A&M AgriLife to conduct the analysis titled "Integrative Assessment of Bacterial Pollution", Contract No: 21-060-025-D274. Texas A&M AgriLife used these CMP Cycle 25 funds to identify hotspots and potential drivers of coastal fecal bacterial pollution. This increased resolution as well as new data linking bacterial pollution with on-site sewage facilities (OSSF), sanitary sewer overflows (SSO), stormwater runoff, wastewater treatment plant (WWTP) effluent, and beach attendance is expected to inform retrofit planning. This final report provides a detailed summary of various tasks (T) completed, and deliverables (D) submitted to GLO related to water quality dataset cleaning and analysis, and Enterococci Data and Human-Specific Fecal Pollution Analysis for Galveston Island, Texas. Main output from each task was summarized in an Infographic which is included in the Appendix-A. Dataset and source files data analyses, along with all deliverables were submitted electronically to the GLO during the project and with this final report.

Texas Beach Watch Enterococci Data dataset had 31,225 records from 1/15/2009 to 2/23/2022. Anomalies, duplicate samples, and "field duplicates" (Required for quality assurance) were flagged and edited, resulting in the creation of a new database (**BW Data _2009-Feb2022_Final_Island.XLSX**). Summary statistics were calculated, including maximum, minimum, average, median, geometric mean, and percentage of exceedance (104 MPN/100mL). Time Analysis showed slight positive correlation with time (Kendall correlation coefficient) and three peak periods (summer, spring, and fall), particularly evident for the Sea Wall stations. Space-Time Analysis showed that sampling stations in close geographic proximity shared trends and characteristics. Beaches and sampling stations were ranked based on the Enterococci exceedance percentage.: High (> 10%), Medium (5 – 10%), and Low (< 5%). Out of all the 36 stations, 11% fell in the High category, 69% in the Medium category, and 19% in the Low category. Most stations and beaches in the Seawall were classified as High or Medium, while those in the West End were mostly classified as Low category.

Environmental metadata were collected from various sources, including TexMesonet, GCOOS, and NOAA and analyzed (one station for rainfall and four stations for sea level). Analysis indicated that correlation between rainfall and Enterococci was higher compared to correlation between water sea level and Enterococci. Coefficients were low, indicating that there might be other drivers.

A micro-watershed map of Galveston Island was created using LIDAR data to identify potential sources of pollution (OSSFs, stormwater and WWTP outfalls, sewer infrastructure, and leaks/spills in sewage systems). Analysis showed that most OSSF are located on the West End and that most sources of pollution discharge on the bay side. No conclusive evidence was found that OSSFs or flow violations have impact on Enterococci sampling results. A significant correlation between Violation of E. coli from WWTP and Enterococci was found for Station #21 (sum of count in the 7-15 days following violation).

Direct and indirect estimates of recreational beach attendance on Galveston Island were determined using foot traffic data from various sources. Direct estimates, including Texas Beach Watch and field observations, helped identify trends both spatially and in time. Indirect

estimates, including Park Board Hotel occupancy records (HOT) and parking monthly sums, confirmed consistent peaks in March, June, and July, and higher totals in the Seawall zone. Statistical Clustering and Space-Time Pattern Analyses showed hot spots for HOT data in the Western portion of the Seawall (stations #34-39) Correlation analyses (Kendall coefficient) between monthly HOT (all structures) and monthly Enterococci geomean for the period 2015-2021, stations by station, indicated best correlation for sampling stations #22, #30-36, #45-47, and #49-50.

Microbial Source Tracking (MST) analysis was conducted for selected water samples collected from Galveston. A total of 114 samples that exceeded the enterococci recreational water quality limit (104 MPN/100 mL) were collected from the period of March 2022 - May 2023. Samples were analyzed using qPCR markers and DNA sequencing-based source tracking for human, dog, and seagull sources. Gull was the most common and most abundant source detected, while human marker was only detected at low levels at a small number of stations. Correlation with Enterococci and with environmental metadata (rainfall and water level) were not significant.

Table of Contents

APPENDICES

APPENDIX A - Infographics from each Task

List of Tables

List of Figures

Introduction

The General Land Office (GLO) contracted with Texas A&M AgriLife to conduct the analysis titled "Integrative Assessment of Bacterial Pollution", Contract No: 21-060-025-D274. This final report summarizes the data and findings in written narrative, graphs, charts, and tables from the project. Copy of data and source files for all analyses are submitted electronically to the GLO Project Manager. The report provides a detailed account of various tasks related to water quality dataset cleaning and analysis for Galveston Island, Texas. Each task focuses on specific aspects of the analysis.

Studies have shown that fecal pollution is associated with a decrease in the resilience and diversity of marine coastal systems. A meta-analysis of 216 studies clearly demonstrated that anthropogenic contamination, including sewage contamination, reduces diversity and resilience in coastal marine systems (Johnston and Roberts, 2009). Threats to diversity and resilience disrupt ecosystem services and endanger the sustainability of marine and socioeconomic systems (Levin and Lubchenco, 2008). For example, the presence of human pathogens associated with sewage contamination can negatively impact recreational bathing and shellfish hygiene (Malham et al., 2014).

A long-term analysis of Texas Beach Watch (TBW) bacterial data by Texas A&M University-Corpus Christi (TAMU-CC) revealed that 25 Texas beaches are hotspots of bacterial pollution. Results also revealed that bacterial pollution is increasing with time, population growth, and sea level rise. Texas A&M AgriLife used these CMP Cycle 25 funds to identify hotspots and potential drivers of coastal fecal bacterial pollution. Data were re-analyzed to pinpoint individual sampling stations that exhibit a history of bacterial pollution. Potential drivers of coastal bacterial pollution were evaluated by assessing 1) the density and integrity of On-Site Sewage Facilities (OSSF), 2) the occurrence of leaks, spills, and sanitary sewer overflows (SSO), 3) the potential connectivity between wastewater infrastructure and surface water pollution, 4) the inflow of stormwater runoff and Wastewater Treatment Plant (WWTP) effluent, and 5) changes in recreational beach attendance. Additionally, the presence of human, canine, and gull fecal waste was confirmed by collecting water samples and testing for the abundance of host-specific molecular markers of fecal pollution.

This increased resolution as well as new data linking bacterial pollution with OSSF, SSO, stormwater runoff, WWTP effluent, and beach attendance is expected to inform retrofit planning. Data derived from this project will inform retrofit planning, primarily through engagement with the local jurisdiction, with a goal of improving coastal water quality, which is essential to the sustainability of coastal ecosystems and coastal economies.

Task 1: Analyze Texas Beach Watch Enterococci Data

T1D1 - Data Cleaning:

- This Task begins with an explanation of the data cleaning process for the *Enterococci* dataset provided by GLO's Texas Beach Watch (TWB) (**BW Data _2009-Feb2022.XLSX**), which includes stations in Galveston County.
- This dataset included 31,225 records, from 1/15/2009 to 2/23/2022, and each record is supposed to correspond to an individual sample.
- Anomalies, duplicate samples, and "field duplicates" were flagged and addressed during cleaning. As a result of cleaning the Enterococci dataset, a total of 8 records were deleted, 7,450 records were corrected for the Enterococci result, and 18 records were corrected for the analysis method. Identified flags include:
  - Anomalies: Enterococci result = 0 or under the limit of detection; change of analysis method; or assignment to the wrong analysis method.
  - Duplicate samples: Sample results entered in the database by mistake.
  - Field duplicates: Required for quality assurance sample taken on the same day at the same station with the same event tag.
- The cleaning process resulted in the creation of a new, cleaned database (**BW Data _2009-Feb2022_Final.XLSX**). Additional columns were created to identify records' unique IDs and three zones as indicated by GLO. Zones include West End (Stations ID 1-33), Seawall (34-47), and East End (48-55).
- As the project area includes only stations falling inside the Galveston Island (Site IDs GAL001-GAL055), a second file was created with only these stations, and is the one used for analysis conducted in the other Tasks (**BW Data _2009-Feb2022_Final_Island.XLSX**) (Figure 1).

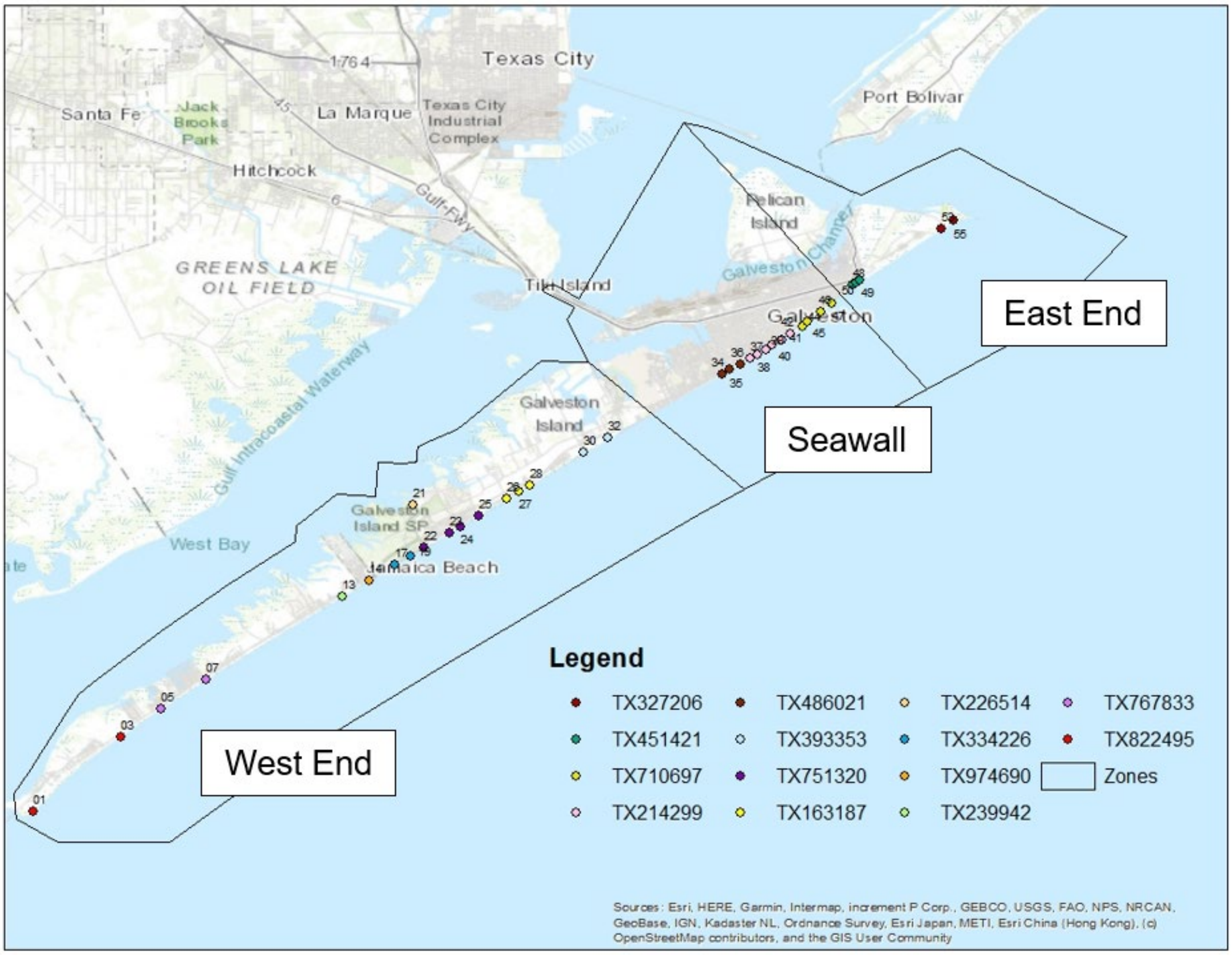


Figure 1. Location map showing Stations, Beaches, and Zones in the project area. Beaches' names are listed in the legend (stations in the same beach have same color)

T1D2 - Summary Statistics:

- This Task outlines the generation of summary statistics for the cleaned Enterococci dataset (maximum, minimum, average, median, geometric mean, and percentage of exceedance).
- For each summary statistic, we created a universal key to identify stations and beaches to simplify the look of tables and figures (Table 1).
- Trends and changes over time and space were explored, revealing correlations between specific stations and seasonal variations.
- Exceedance is known in this context as the percentage of *Enterococci* above the coastal water quality standard of 104 MPN/100mL, established by the Beaches Environmental Assessment and Coastal Health (BEACH) Act with the goal of protecting human health.
- Peaks in maximum value were found in station #21 (the only station in Galveston Bay), Dellanera Park Beach (Stations #30 and #32), West part of Sea Wall (station #37 in particular), and Stewart Beach (station #48-50).. Average values had a similar behavior. Median values showed a similar overall trend, but with lower peaks, and with highest values for the very West of the Sea Wall (Stations #34 and #35). The trend for median values was shared quite closely also by geomean and exceedance values (Figure 2).
- Time Analysis:
    - Slight positive correlation with time (Kendall correlation coefficient), with peaks in the years 2014 and 2015 (water temperature and algae blooms), and relative lower values in 2011 (exceptional drought year) and 2020 (beach closures due to COVID pandemic (Figure 3).
    - Most stations had higher values in the summer, while some had peaks also in spring and fall. These three peak periods are particularly evident for the Sea Wall stations (Figure 4).
- Space-Time Analysis:
    - Space patterns in the project area showed that sampling stations in close geographic proximity shared trends and characteristics.
    - The general spatial pattern is similar in the years, but some stations have peaks in different years (e.g., Sea Wall stations); this is consistent for both variables analyzed (geometric mean, and percentage of exceedance) (Figure 5).

Table 1. Keys for each Station and Beach, and their relation

Site Simplified ID	Site ID	Beach Simplified ID	Beach Name
1	GAL001	1	TX822495
3	GAL003		
5	GAL005	2	TX767833
7	GAL007		
13	GAL013	3	TX239942
14	GAL014	4	TX974690
17	GAL017	5	TX334226
19	GAL019		
21	GAL021	6	TX226514
22	GAL022	7	TX751320
23	GAL023		
24	GAL024		
25	GAL025		
26	GAL026	8	TX163187
27	GAL027		
28	GAL028		
30	GAL030	9	TX393353
32	GAL032		
34	GAL034	10	TX486021
35	GAL035		
36	GAL036		
37	GAL037	11	TX214299
38	GAL038		
39	GAL039		
40	GAL040		
41	GAL041		
42	GAL042		
44	GAL044	12	TX710697
45	GAL045		
46	GAL046		
47	GAL047		
48	GAL048	13	TX451421
49	GAL049		
50	GAL050		
53	GAL053	14	TX327206
55	GAL055		

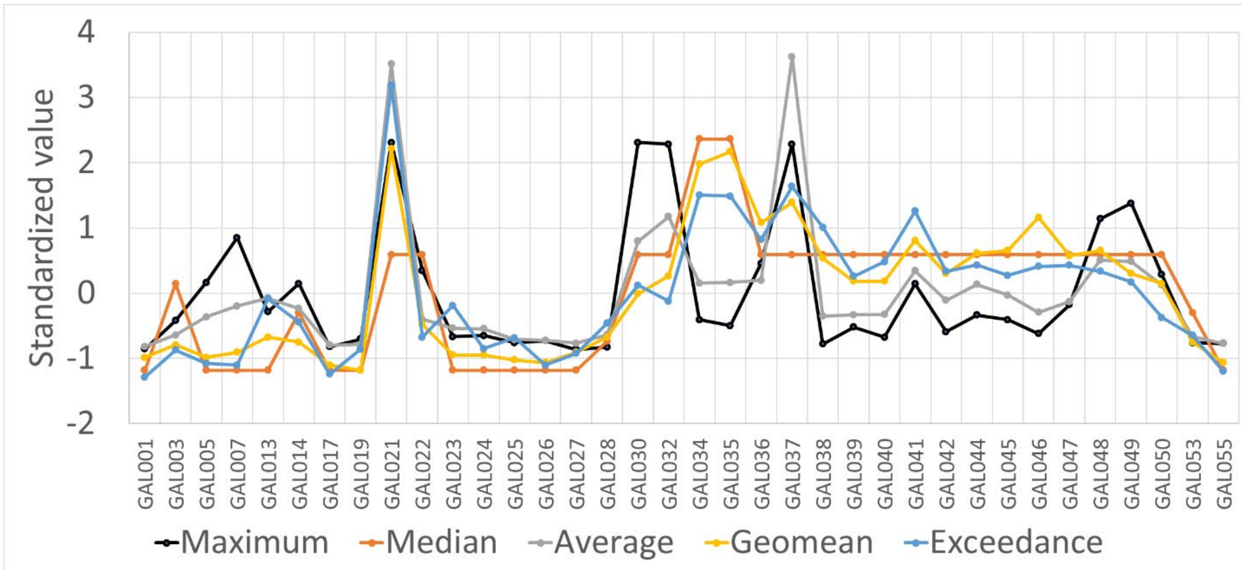


Figure 2. Standardized values for all statistics, so trends can be compared.

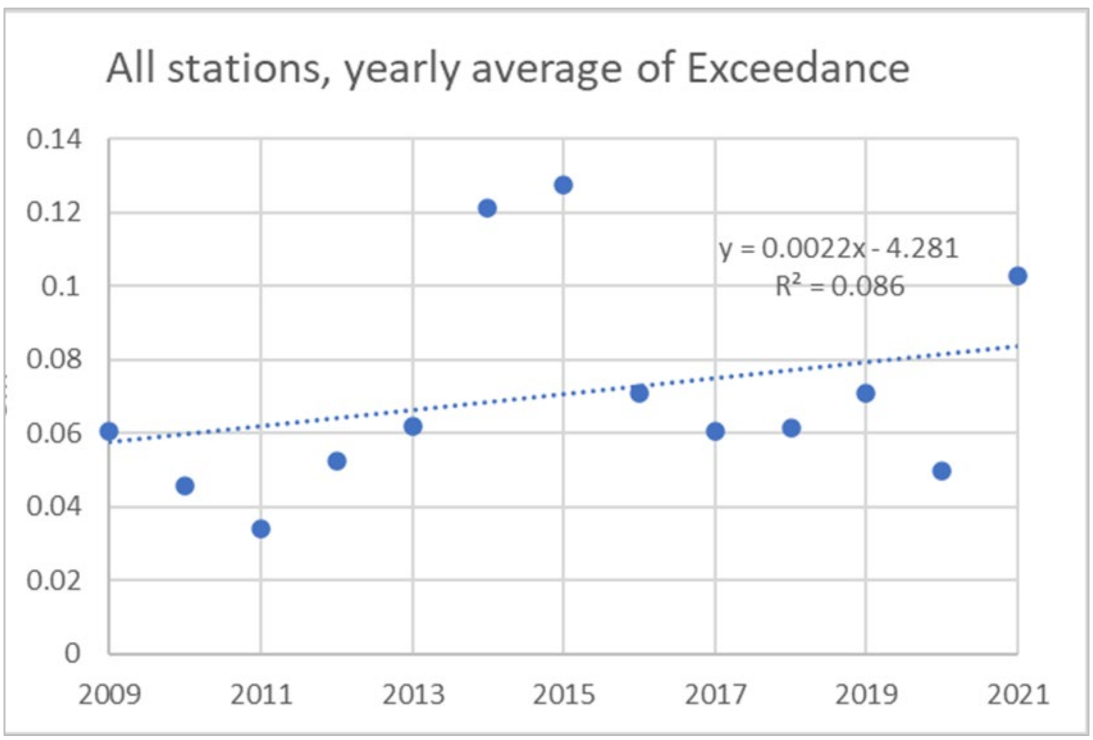


Figure 3. Yearly average exceedance for all stations combined, and linear regression interpolating lines.

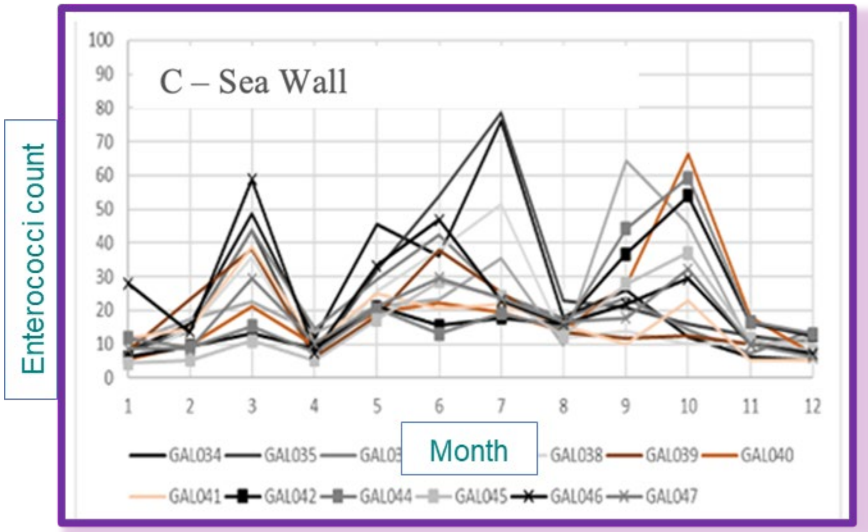


Figure 4. Seasonal predictions for Enterococci count for the Sea Wall stations.

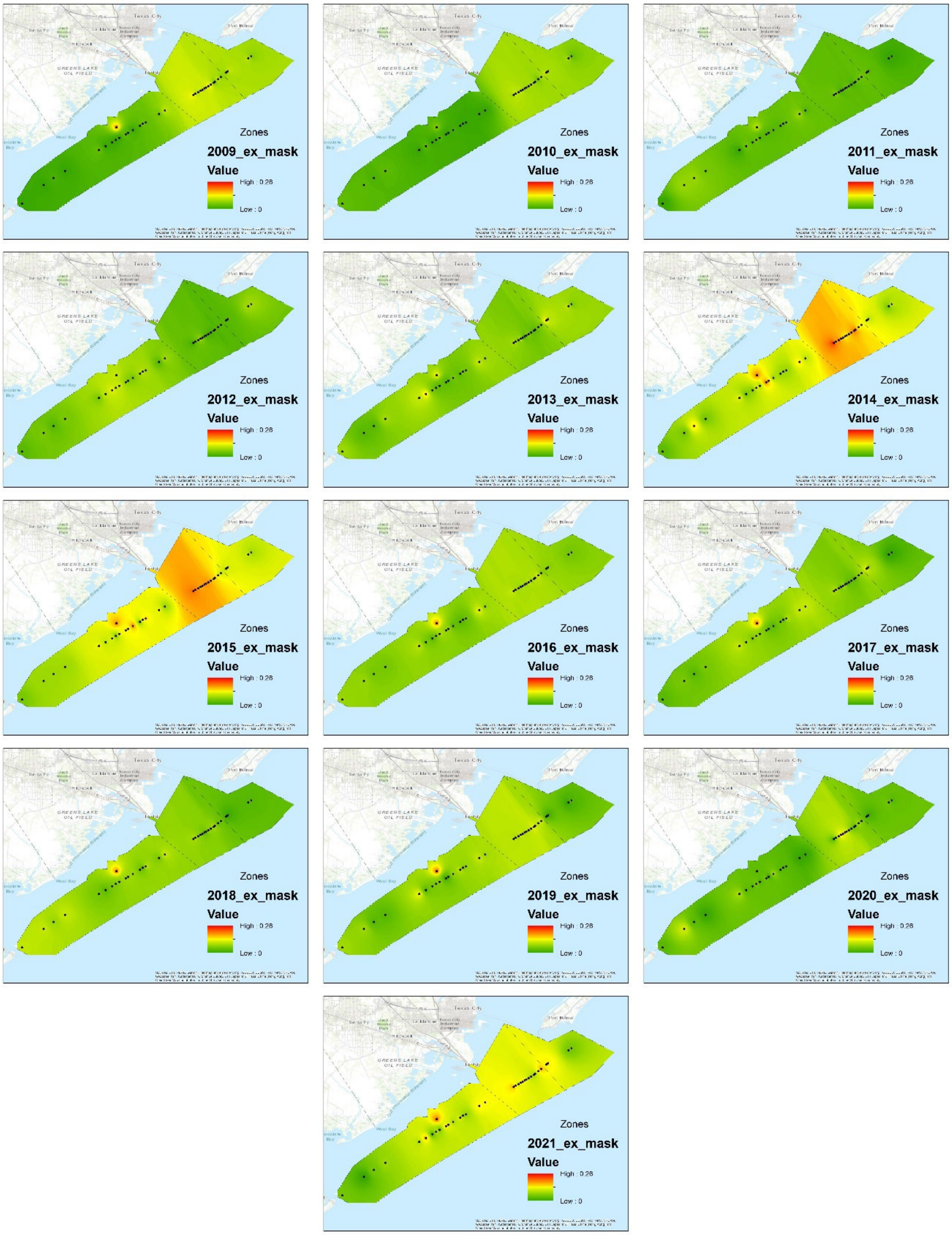


Figure 5. Inverse Distance Weighting tool (IDW) analysis on yearly average exceedance by station.

T1D3 - Beach Ranking:

- The report details the process of ranking beaches and sampling stations based on bacterial pollution levels (exceedance percentage).
- Three categories were established: low (< 5%), medium (5 – 10%), and high (> 10%) based on previous studies.
- Out of all the 36 stations, 11% fell in the High category, 69% in the Medium category, and 19% in the Low category (Table 2).
- Most stations and beaches in the Seawall were classified as high or medium, while those in the West End were predominantly low, with some exceptions (Figures 6 and 7).

Table 2. Ranking of stations and beaches based on % Exceedance of water quality limit of 104MPN/100mL: Low (Green < 5%), Medium (Yellow 5-10%), High (Red > 10%)

ID		% Exceedance	
Site	Beach	Site	Beach
1	1	4.13	4.61
3		5.07	
5	2	4.60	4.58
7		4.55	
13	3	6.86	6.86
14	4	6.04	6.06
17	5	4.23	4.66
19		5.10	
21	6	14.31	14.31
22	7	5.52	5.67
23		6.63	
24		5.12	
25		5.50	
26	8	4.55	5.18
27		4.96	
28		6.01	
30	9	7.33	7.06
32		6.78	
34	10	10.48	10.01
35		10.44	
36		8.92	
37	11	10.77	8.96
38		9.35	
39		7.65	
40		8.15	
41		9.92	
42		7.82	
44	12	8.04	7.94
45		7.67	
46		7.99	
47		8.03	
48	13	7.82	7.17
49		7.45	
50		6.21	
53	14	5.59	4.97
55		4.33	

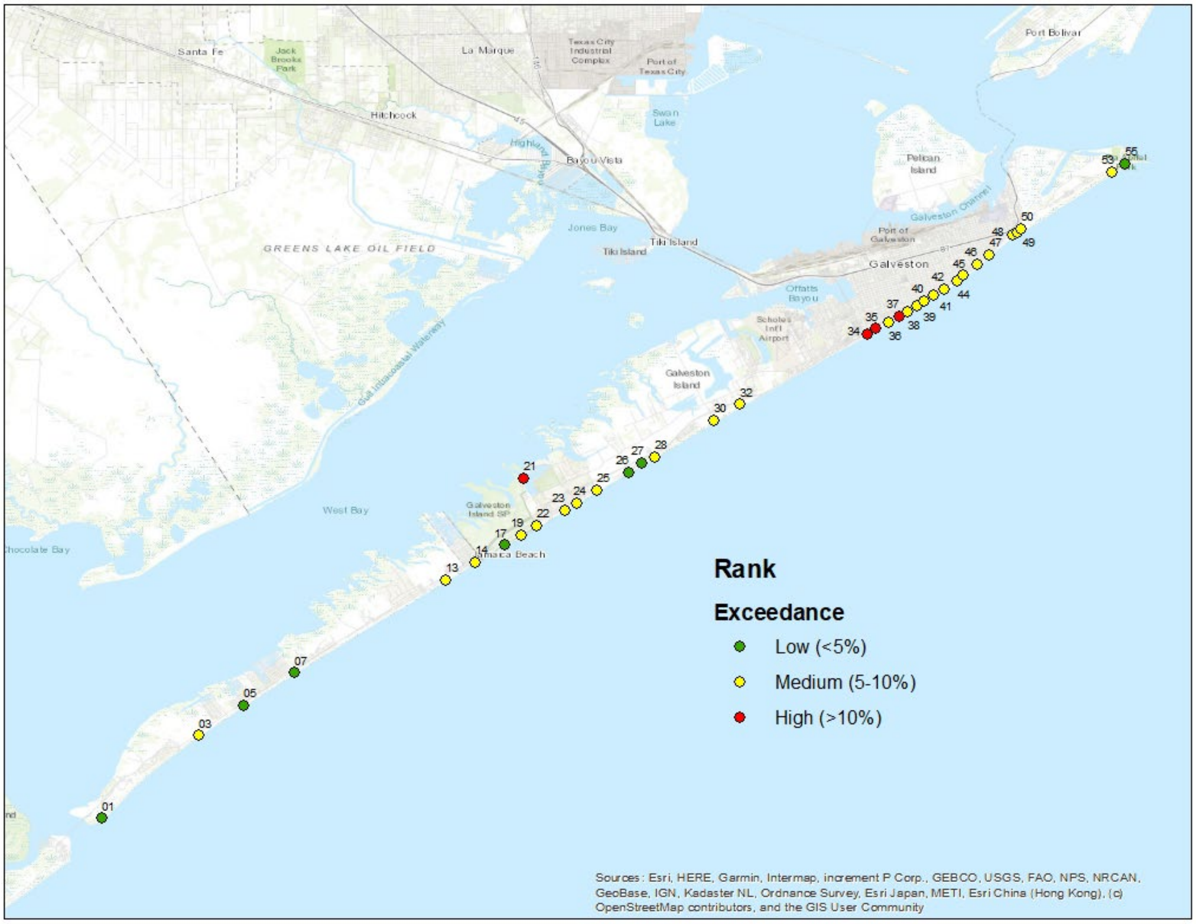


Figure 6. Ranking of stations based on Exceedance of water quality limit of 104 MPN/100mL: Low (Green < 5%), Medium (Yellow 5-10%), High (Red > 10%).

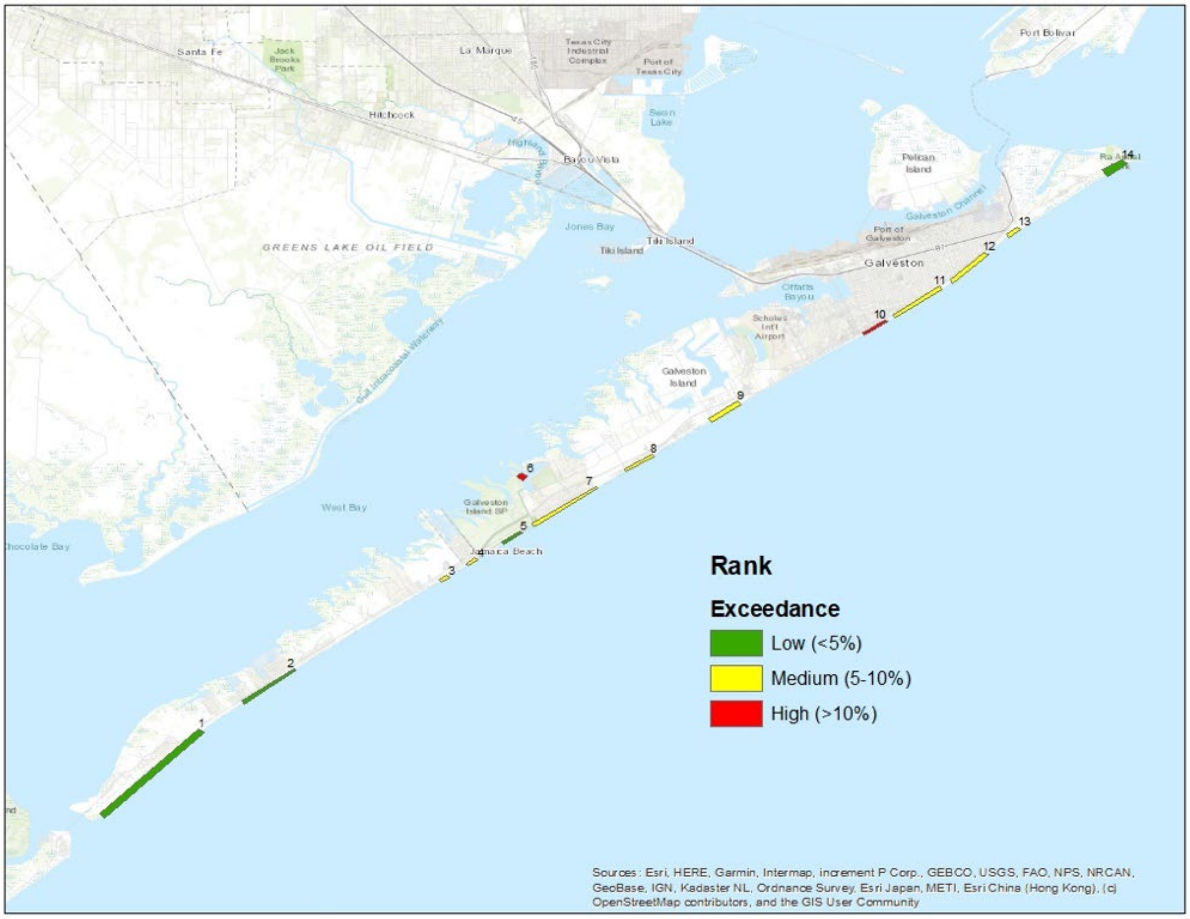


Figure 7. Ranking of beaches based on Exceedance of water quality limit of 104 MPN/100mL: Low (Green < 5%), Medium (Yellow 5-10%), High (Red > 10%).

Task 2: Compare Enterococci Data to Environmental Data

T2D1 - Environmental Data:

- This report discusses the collection and processing of rainfall and sea level data from various sources, including TexMesonet, GCOOS, and NOAA.
- Details about data format, sources, and quality control are provided for each dataset.
- A total of 13 datasets were collected, of which one for rainfall (Galveston Airport) and 12 for sea level (four stations: San Luis Pass, Galveston Railroad Bridge, Galveston Pier 21, and Galveston Bay Entrance) (Figure 8).

Figure 8. Locations of rainfall and sea level measuring stations (TexMesonet, GCOOS, and NOAA)

T2D2 - Enterococci Dataset and Environmental Metadata Comparisons:

- The report covers statistical methods and outputs for comparing Enterococci concentrations with environmental datasets prepared in Task 2.
- Statistical tests, including T-tests and correlation analyses, were used to assess relationships between environmental data and Enterococci concentrations.
- Largest rain events (about > 2 inches) always correlated with an Enterococci result higher than the minimum level of detection. Correlation did not improve using multiple days of rainfall sums (Figure 9).
- Analysis indicated a higher correlation between rainfall and Enterococci compared to water sea level and Enterococci. Correlation coefficients were low, indicating that there might be other drivers.

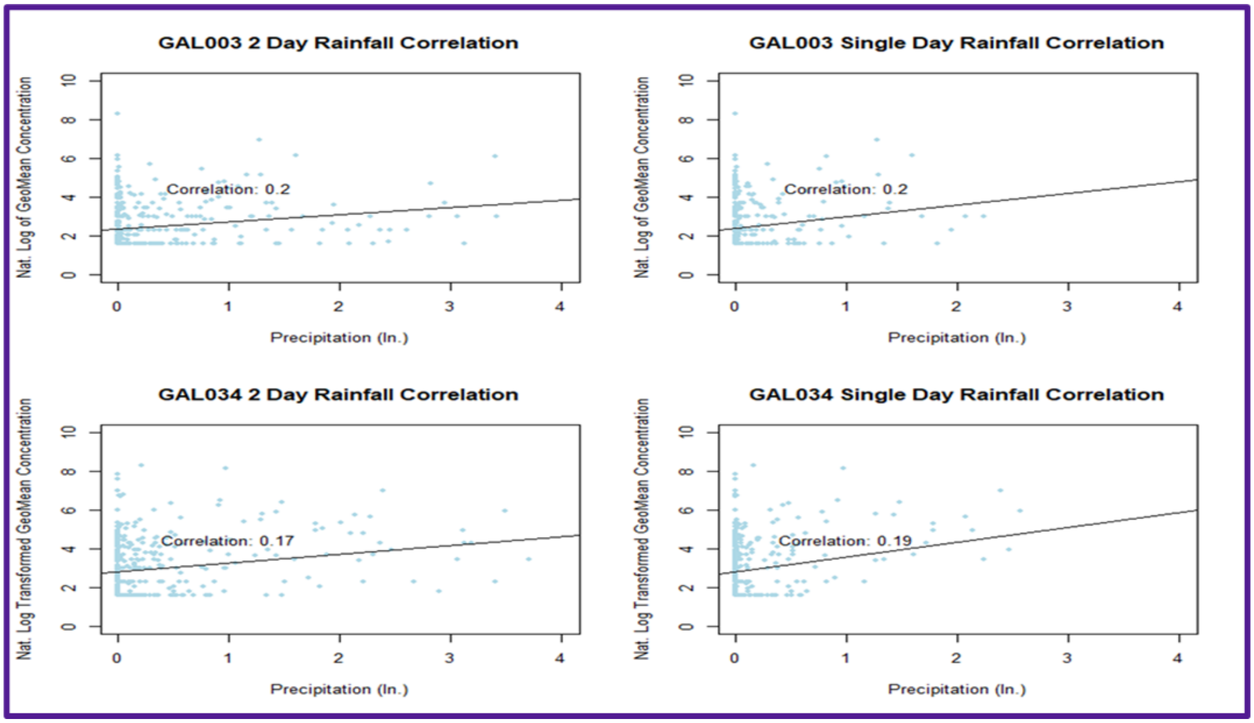


Figure 9. Two examples of Rainfall 2 Day and Single Day Kendall correlation (in inches) compared to natural log of *Enterococci* concentrations. Sampling station #3 (top), and Sampling station #34 (bottom); 2-days rainfall sum (left) and 1-day rainfall sum (right).

Task 3. Compare Enterococci Data to Bacterial Pollution

T3D1 - Micro-Watershed Analysis:

- This report focuses on the creation of a micro-watershed map of Galveston Island using LIDAR data to identify potential sources of pollution.
- The estimated flow direction indicated that drainage is mostly toward the bay (Figure 10).
- The analysis includes the identification of coastal OSSFs, stormwater and WWTP outfalls, sewer infrastructure, and leaks/spills in sewage systems.
- Most OSSF are located on the West End, with older systems are in the "far West", "far East", and coastal portions of West End. Most Stormwater outfalls and all WWTP outfalls discharge on the bay side.
- Most flow violations are located along the sea wall (Sanitary Sewer Overflows database, SSO), 3% of which are inside a micro-watershed discharging to the ocean.
- Most E. Coli violations are located along the sea wall and in Jamaica Beach (Enforcement and Compliance History Online database, ECHO), all inside a micro-watershed discharging to the bay.

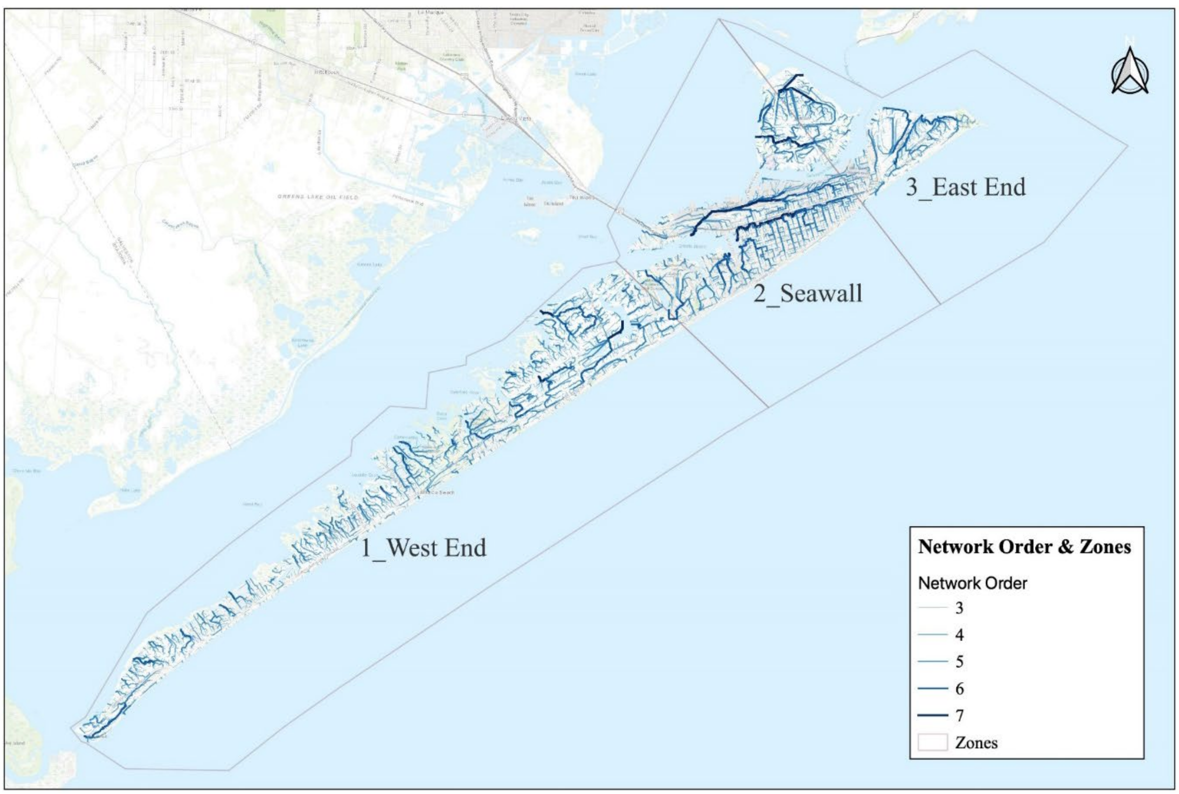


Figure 10. Zones and estimated network relative to Galveston Island. Analysis was conducted using ArcGIS Software tools starting from the 2018 Digital Elevation Model (DEM)

T3D2 - Sewage Contamination Analysis:

- This report examines potential sewage contamination sources in micro-watersheds and their relationship with Enterococci concentrations.
- Statistical analyses were performed to assess the impact of OSSFs, flow violations, and E. Coli violations on water quality.
- No conclusive evidence was found that OSSFs or flow violations have impact on Enterococci sampling results.
- A significant correlation between Violation of E. coli from WWTP and Enterococci was found for Station #21 (sum of count in the 7-15 days following violation).

Task 4: Compare Enterococci Data to Beach Attendance

T4D1 - Recreational Beach Attendance Estimates:

- The report provides direct and indirect estimates of recreational beach attendance on Galveston Island using foot traffic data from various sources.
- Direct estimates:
  - Texas Beach Watch data were collected early morning and identified peak usage in June-July.
  - Field observations were conducted in September and October 2022, including interviews, and helped identifying trends both spatially (Jamaica Beach, Sea Wall and ones with open access near Sea Wall, Stewart Beach) and in time (peaks in early and late afternoon, weekends, holidays such as July 4th, July, June, May, March for spring break; beaches closed in March-May 2020).
- Indirect estimates:
  - Hotel occupancy records (Hotels/Full Service, Hotels/Limited Service, and Independents) confirmed consistent peaks in March, June, and July (Figure 11).
  - HOT 911 addresses locations, aggregated into hexagons (13 sq mi), showed higher totals in the Seawall zone (Figure 12).
  - Parking data monthly sums (Payment Amount Total, available only for the Sea Wall area) were consistent with Park Board HOT monthly and yearly patterns.

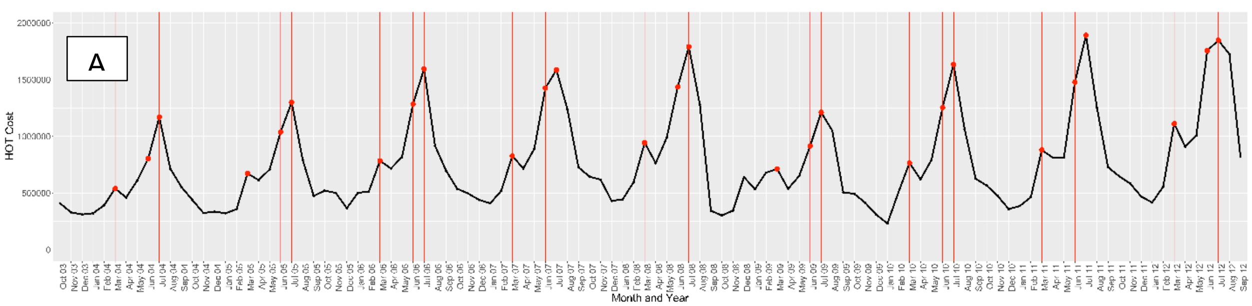


Figure 11. Time series of monthly City HOT data for the entire available dataset (2003-2012); three peaks are consistently showing (March, June, July)

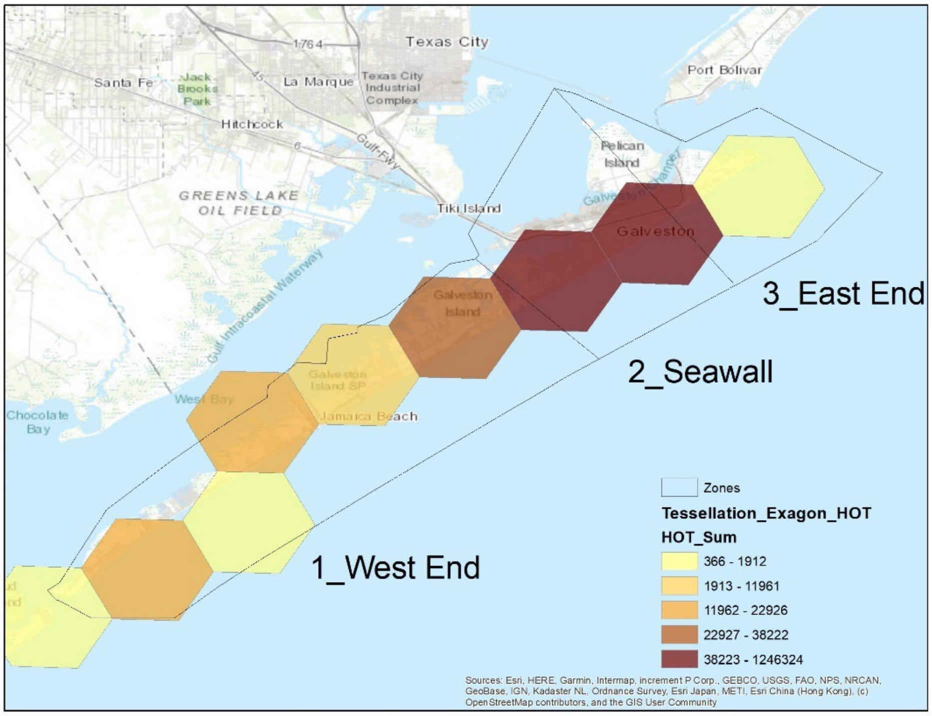


Figure 12. Park Board HOT for Hotels and Independents structures. Exagon Tessellation covering all hotels, and assignment of the sum of payment amounts of all hotels falling inside the same hexagon (categories were automatically generated using the Natural Breaks ArcMap function)

T4D2 - Statistical Clustering and Space-Time Pattern Analyses:

- This report explores spatial and temporal patterns of recreational beach attendance using statistical clustering and space-time pattern analyses.
- Clustering analysis was conducted on hotel data and Park Board HOT data for June 2022, revealing certain clusters and hotspots (only data for which location and spatial variability were available).
- Hot spots were found in the Western portion of the Seawall (stations # 34-39) and several clusters and few outliers in West End and Seawall (Figure 13).
- Space-Time Pattern Analysis was not possible due to limitations in all datasets: Texas Beach Watch (8 AM only); Field truth (one-time); City HOT (location not available); Park Board HOT (data format time consuming to download); Parking (location not available).

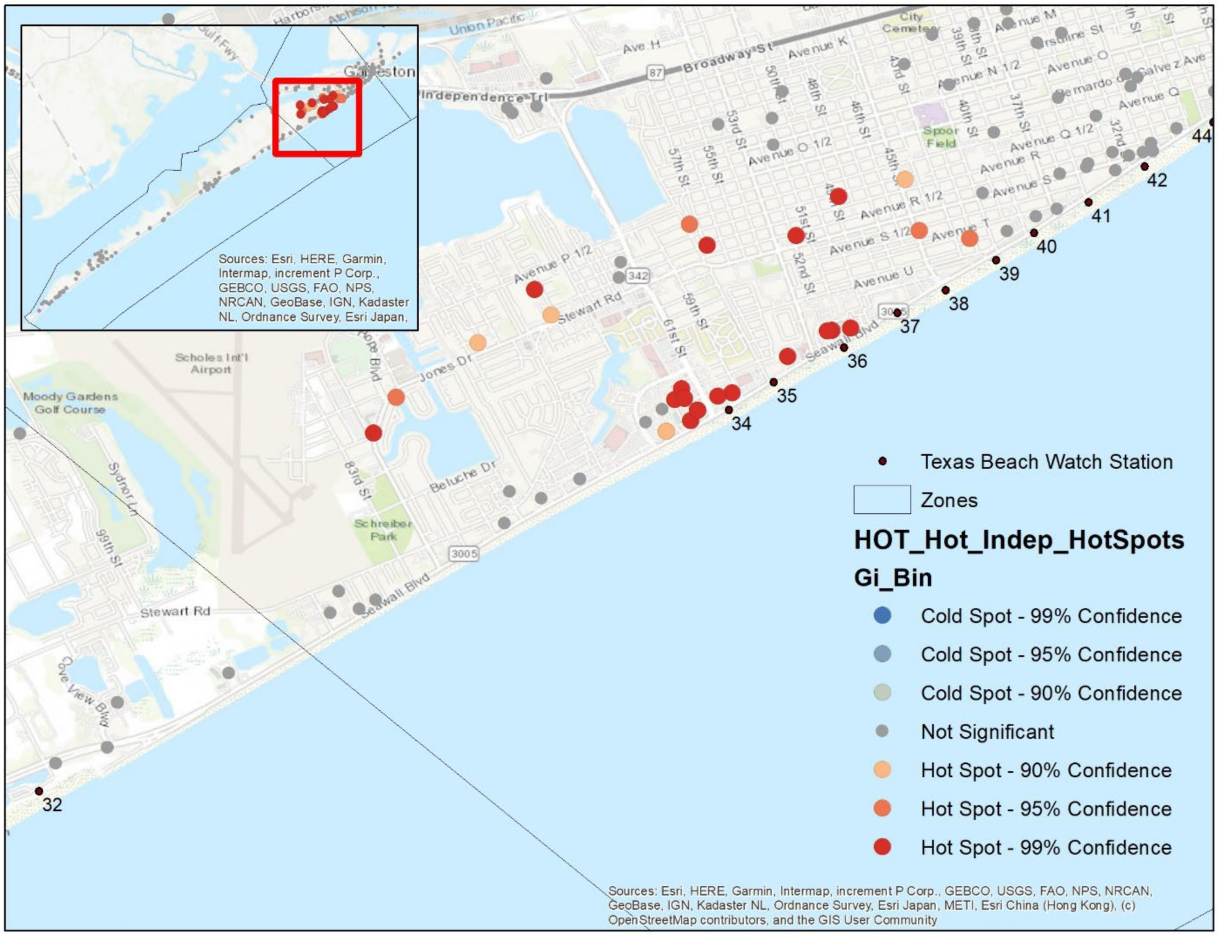


Figure 13. Park Board HOT Hotels and Independent structures: Hot Spot Analysis (Getis-Ord Gi*) for the month of June 2022, detecting a hot spot area in the Western portion of the Seawall

<u>T4D3 - Statistical Outputs from Enterococci Dataset and Estimated Recreational Beach Attendance:</u>

- This report investigates the relationship between estimated recreational beach attendance and Enterococci concentrations using correlation tests and spatial regression.
- Correlation analyses (Kendall coefficient) between Park Board monthly HOT (all structures) from 2015 to 2022 and monthly Enterococci geomean for the period 2015-2021, stations by station, indicated best correlation for sampling stations #22, 30-36, 45-47, and 49-50.
- Spatial Regression (Geographical Weighted Regression, GWR) was conducted to compare hexagon tiles of HOT (Hotels and Independents structures and June 2022 only) (Figure 12) and monthly Enterococci geomean (Year-round, 2015-2021) (Figure 14).
- GWR indicated that predictions' confidence was always higher than 95% (standardized residuals < ±2.5), with best agreement in the Eastern part of the island (Figures 15 and 16).

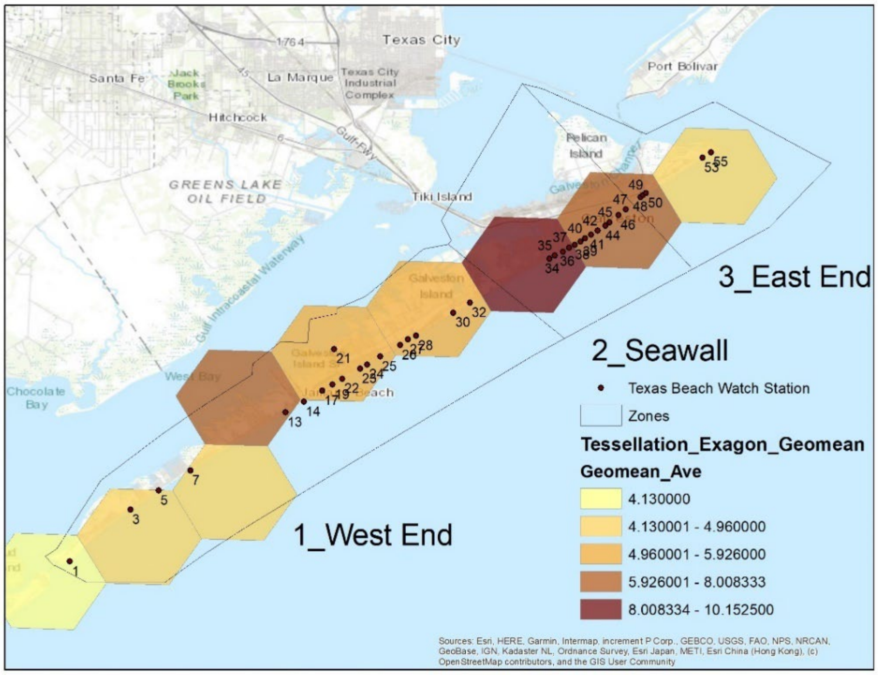


Figure 14. Enterococci geomean overall data (Year-round, 2015-2021) after using the Tessellation tool.

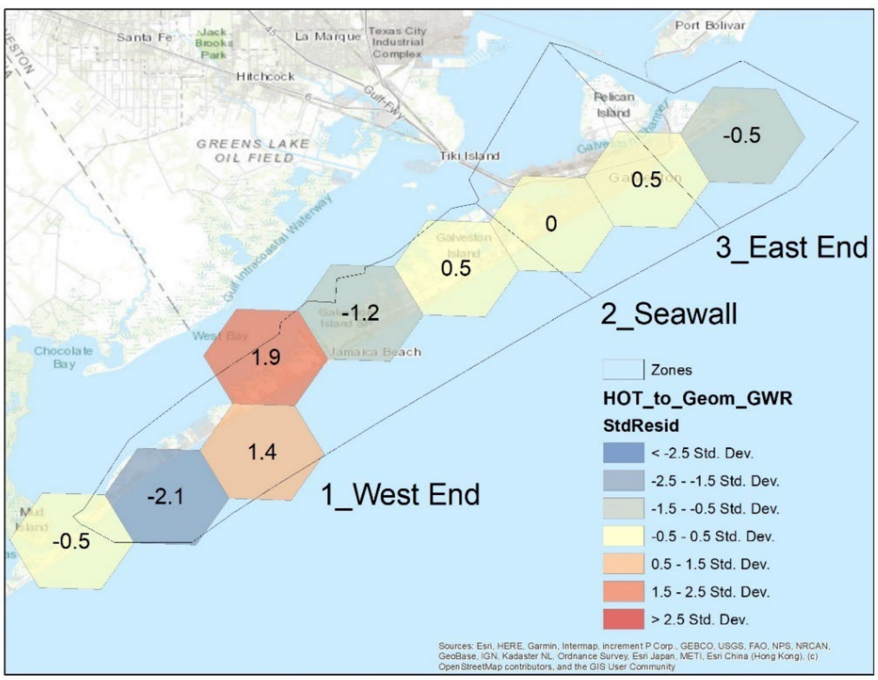


Figure 15. Geographically Weighted Regression (GWR) results as standardized residuals, for Enterococci overall geomean (Year-round, 2015-2021) and June 2022 Park Board HOT data. Labels report each tile's value. Prediction confidence is higher than 95%, especially in the Eastern part of the island.

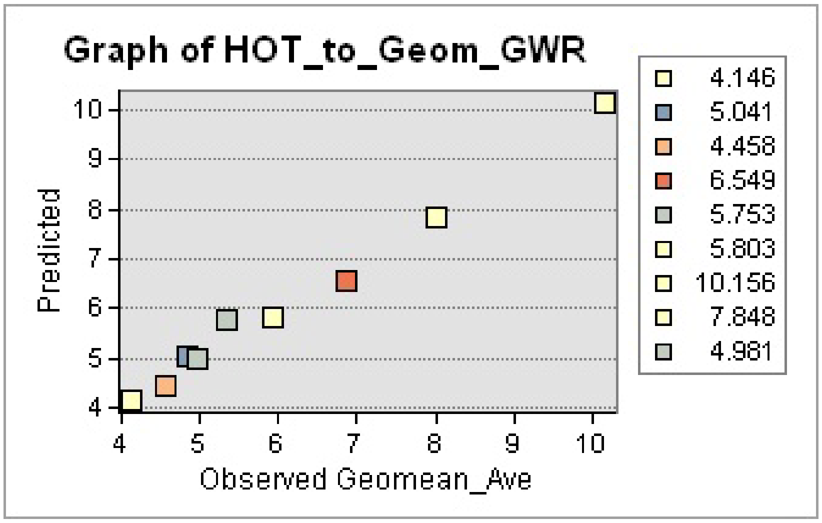


Figure 16. Geographically Weighted Regression (GWR) results as scatter plot of observed vs predicted geomean values, for Enterococci overall geomean (Year-round, 2015-2021) and June 2022 Park Board HOT data. Predicted and observed geomean values are in good agreement.

Task 5: Enterococci Data and Human-Specific Fecal Pollution Analysis

- This report describes microbial source tracking analysis for selected water samples collected from Galveston, TX during 2022-2023.
- A total of 114 samples that exceeded the enterococci recreational water quality limit (104 MPN/100 mL) were collected from the period of March 2022 - May 2023.
- Samples were analyzed using qPCR markers for human, dog, and seagull sources. Additionally, samples from July and August 2022 were analyzed using DNA sequencing-based source tracking.
- Of all the samples tested, gull was the most common and most abundant source detected using both qPCR and DNA sequencing. Human markers were detected at low levels below the limit of quantification except for GAL032 in West End (Figures 17-19).
- DNA sequencing source tracking results indicated a greater impact on the water bacterial community from the treated WWTP effluent than from WWTP or septic sources.
- No correlation was found between the qPCR markers, enterococci levels, or environmental data (rainfall and water level). This could be due to variation in the persistence of the various markers, microorganisms detected by markers being different from enterococci, and the relatively small number of samples examined.

Figure 17. TBW stations where human markers were detected.

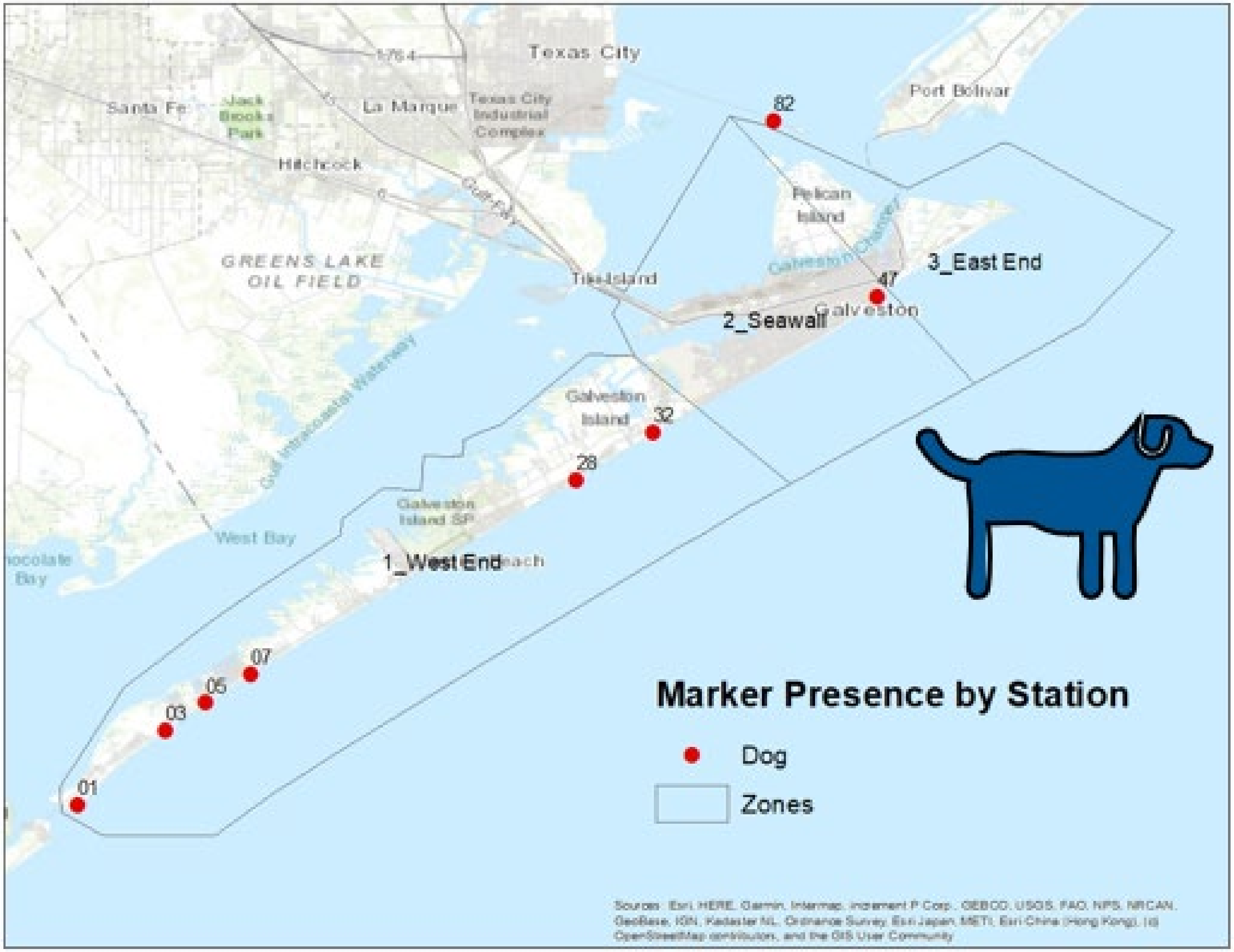


Figure 18. TBW stations where dogs markers were detected.

Figure 19. TBW stations where gull markers were detected.

APPENDIX-A

Infographics from each Task

T1D1

Cleaned up Texas Beach Watch Enterococci Dataset

Report contains:

- Received dataset (**BW Data _2009-Feb2022.XLSX**)
- Flags: anomalies, duplicate samples, and "field duplicates."
  - Anomalies: Entero result = 0 or under the limit of detection; change of analysis method; or assignment to the wrong analysis method
  - Duplicate samples: Sample results entered in the database by mistake
  - Field duplicates: Required for quality assurance, and is a sample taken on the same day at the same station with the same vent tag (two or three samples)

Datasets

1st: Received from the General Land Office (GLO) included 51,701 records (identified extra not needed information)

2nd (corrected): 31,225 records, from 1/15/2009 to 2/23/2022, each record corresponds to an individual sample

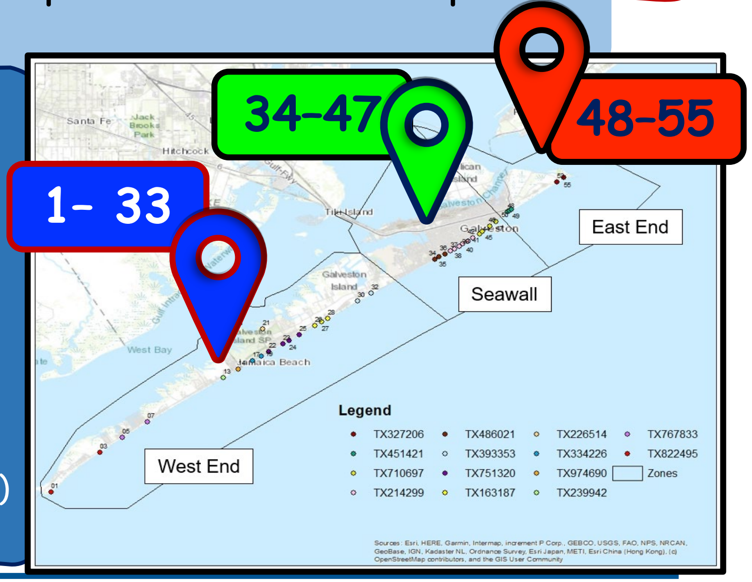
Result of cleaning

→ **BW Data _2009-Feb2022_Final.XLSX**: Total of 8 records were deleted; 7,450 records were corrected for the entero result, and 18 records were corrected for the analysis method.

→ **BW Data _2009-Feb2022_Flagged.XLSX:** This file includes all flags for anomalies, duplicates, and field duplicates (column "Flag"), and all notes for changes (column "Note")

New Columns:

- "Sample ID", was filled with a progressive unique identification number, and was introduced to facilitate conversation regarding any changes made to the dataset.
- "Zones", was added to identify three zones as indicated by GLO: West End (Stations ID 1-33), Seawall (34-47), East End (48-55)



One sample (record ID 107) was identified as a Flag 1 anomaly (Entero result = 0). This sample was removed from the dataset as instructed from GLO:

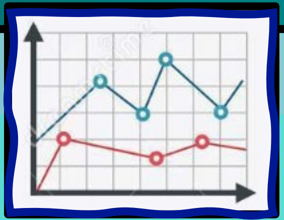
Sample ID	Beach ID	Project Name	Site ID	Station Name	Entero Result	Units	Sample Date	Sample Time	Event Tag	Flag	Note
107	TX822495	West End Galveston – San Louis Pass	GAL001	San Louis Pass Troll Bridge	0	Cfu/100 mL	08/16/2010	9:00:00 AM	11297	1	Remove

T1D2

Summary Statistics for Each Beach and Sampling Station

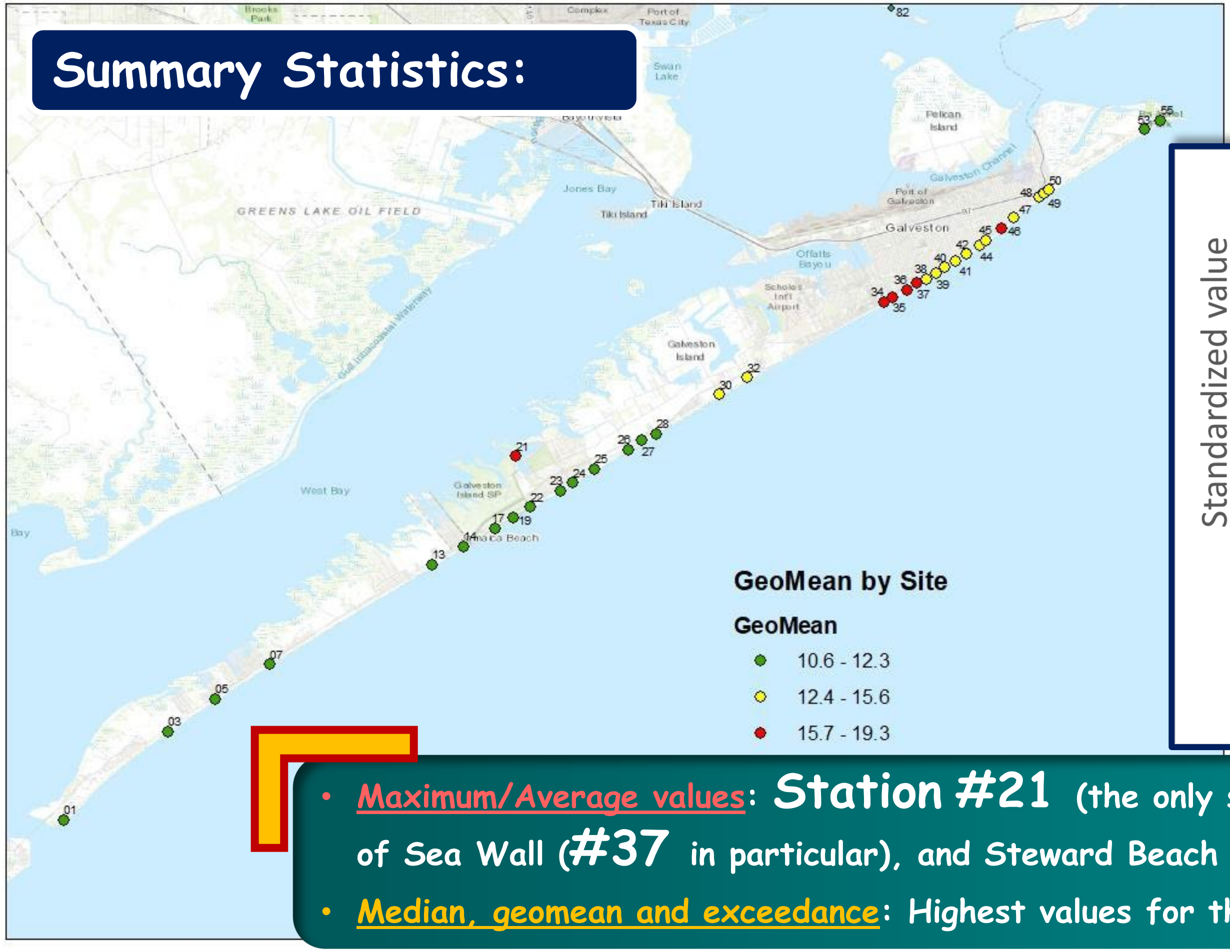
Report on:

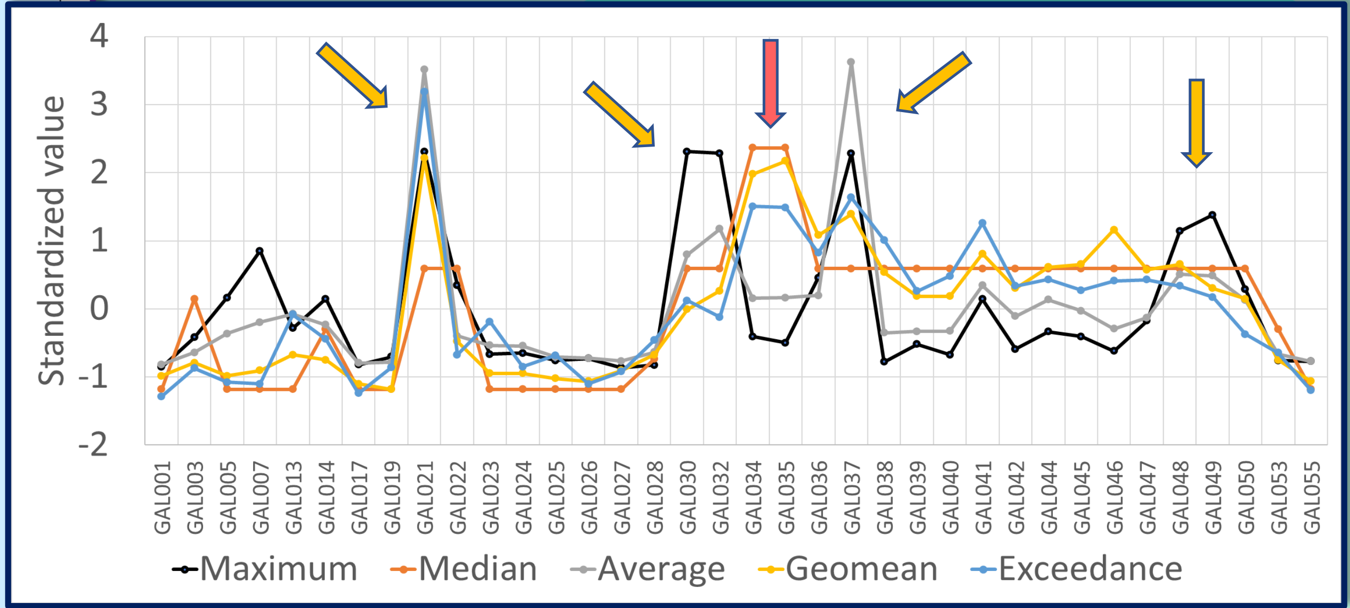
- Summary Statistics
- Time & Space



- Dedicated sections on geometric mean & percentage of exceedance calculations: **104 MPN/100mL** = Beaches Environmental Assessment and Coastal Health (BEACH) Act with the goal of protecting human health

Summary Statistics:





- Maximum/Average values: **Station #21** (the only station in the Bay), Dellanera Park Beach (**#30 and 32**), West part of Sea Wall (**#37** in particular), and Steward Beach (**#48 and 49**)
- Median, geomean and exceedance: Highest values for the very West of the Sea Wall (**Stations #34 and 35**)

T1D2

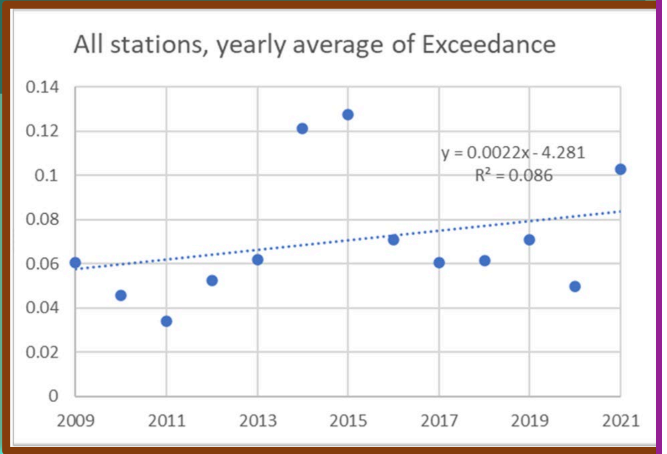
Summary Statistics for Each Beach and Sampling Station

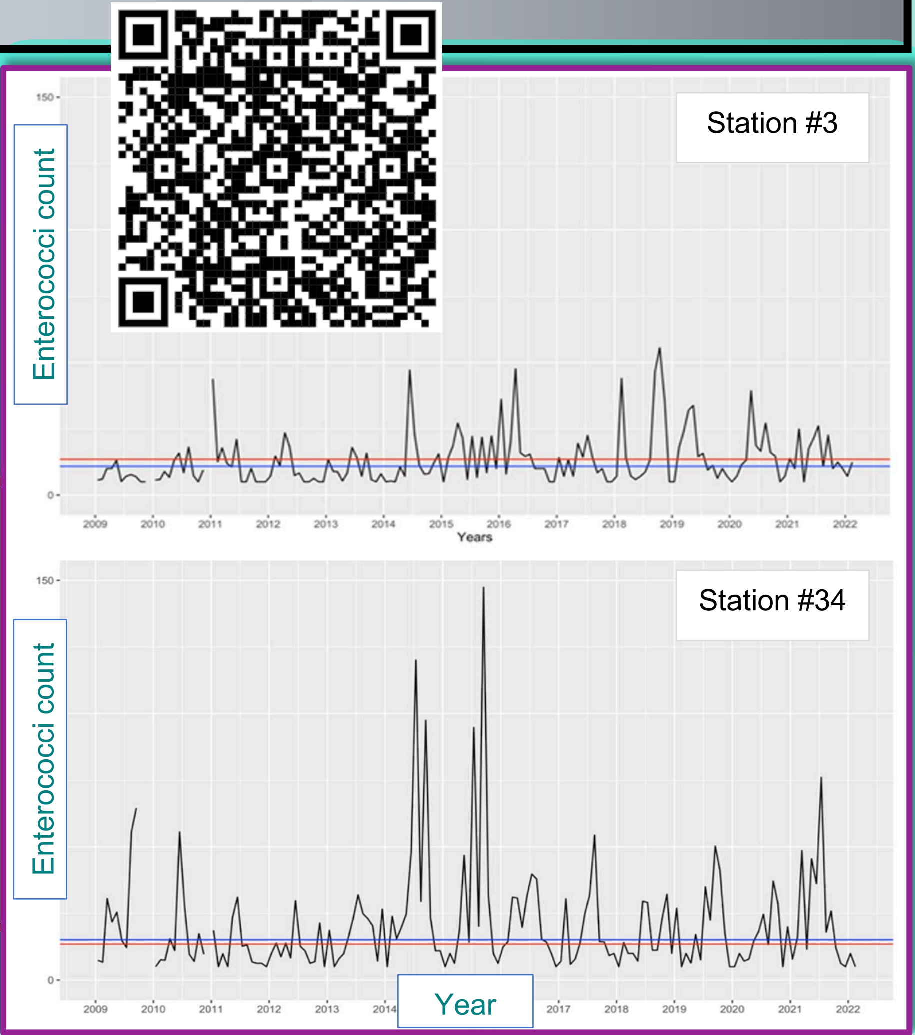
Time analysis:

Space patterns in the project area showed that sampling stations in closed geographic proximity shared trend and characteristics

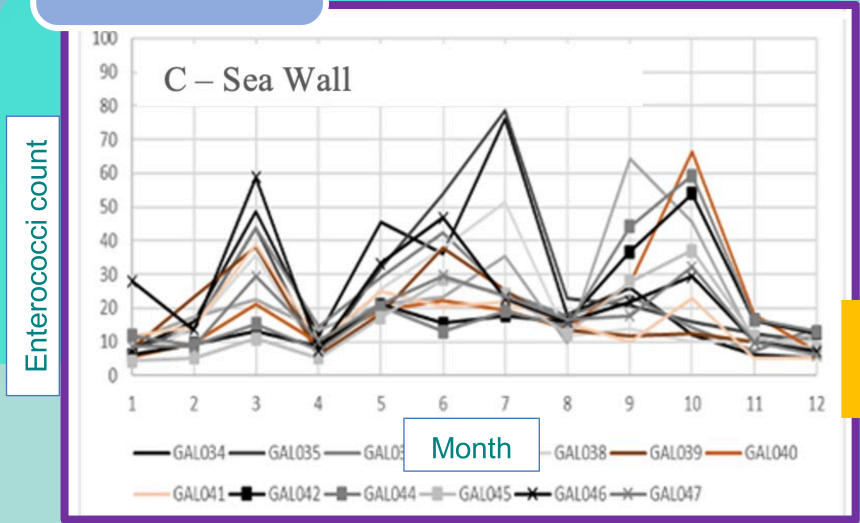
Yearly Trends

Slight positive correlation with time (Kendall correlation coefficient), with peaks in the years 2014 and 2015, and relative lower values in 2011 (exceptional drought year) and 2020 (beach closures due to COVID pandemic)





Seasonal Trends



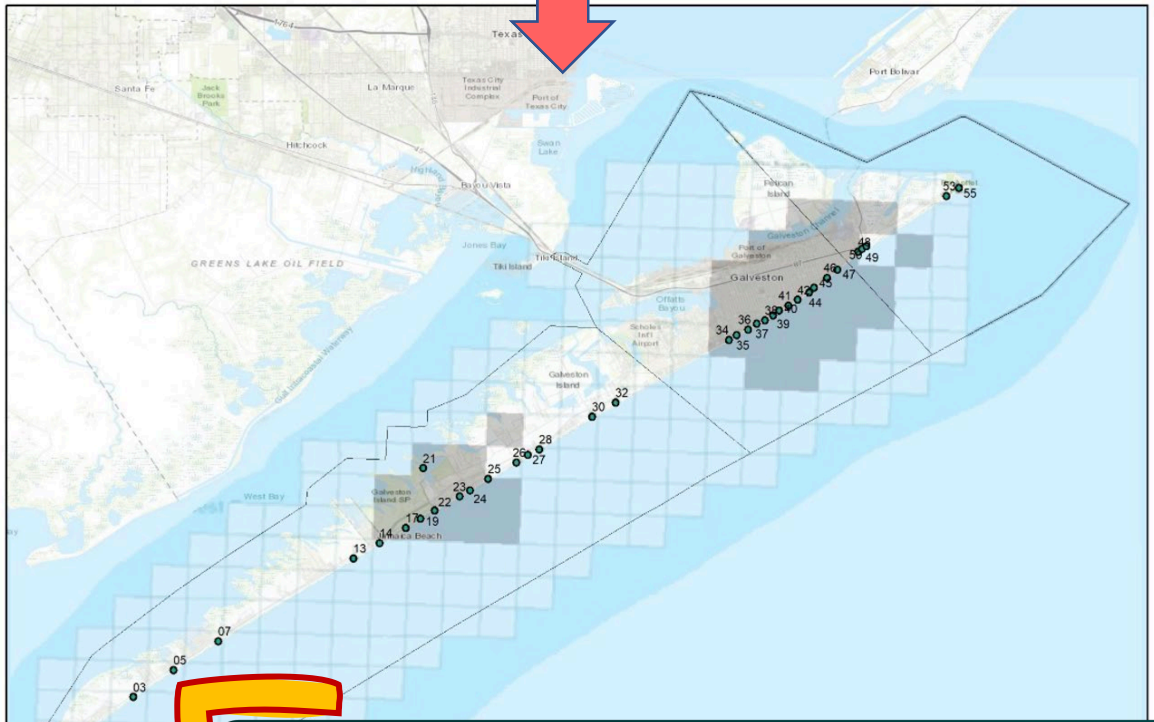
- Auto-Regressive Integrated Moving Average (ARIMA) Models
- Most stations had higher values in the summer, while some had peaks also in spring and fall. These three peak periods particularly evident for the Sea Wall stations.

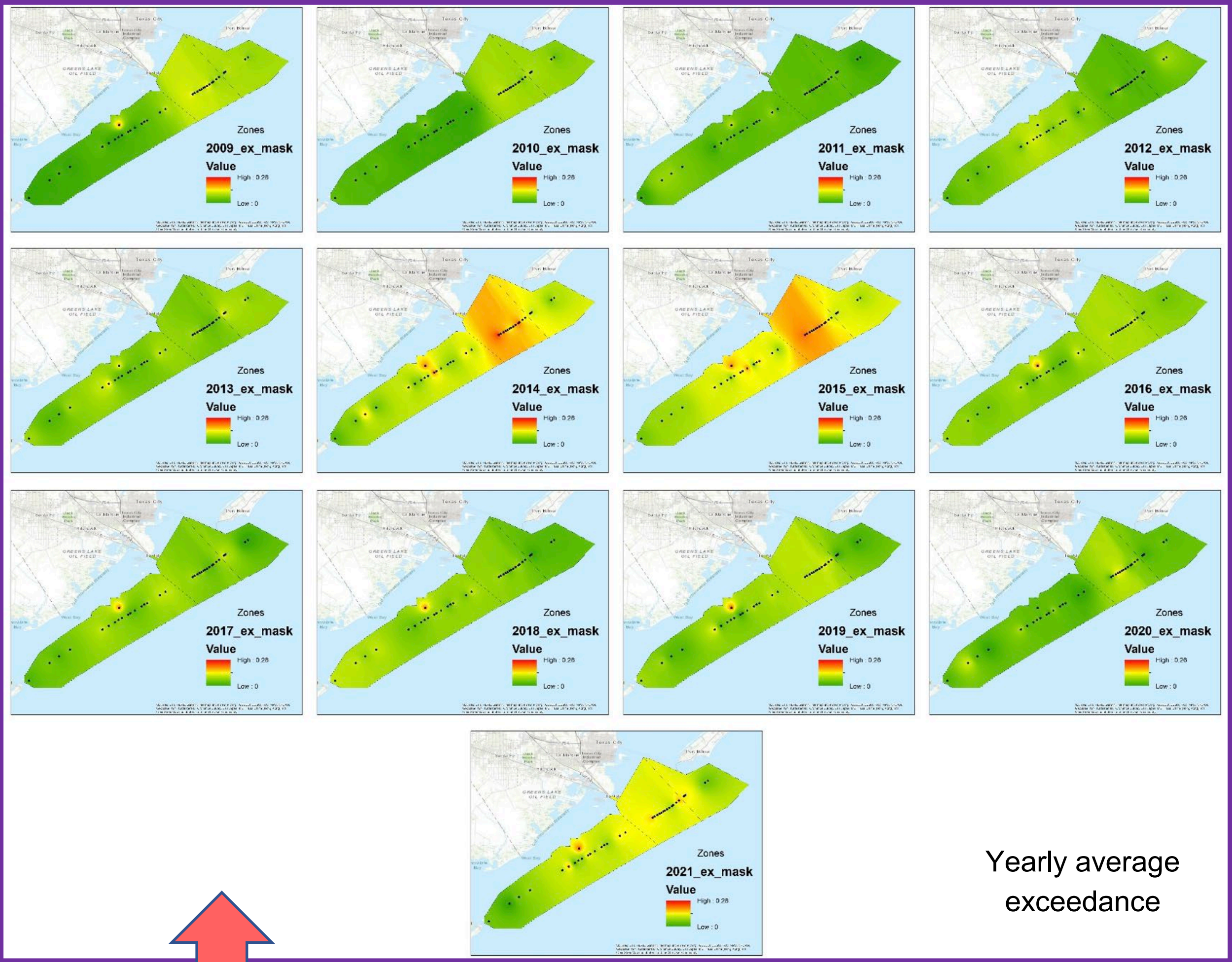
T1D2

Summary Statistics for Each Beach and Sampling Station

Space-Time Analysis:

Emerging Hotspot Analysis: Persistent high annual Enterococci GeoMean and Exceedance values (**Sea Wall and Station #21**)



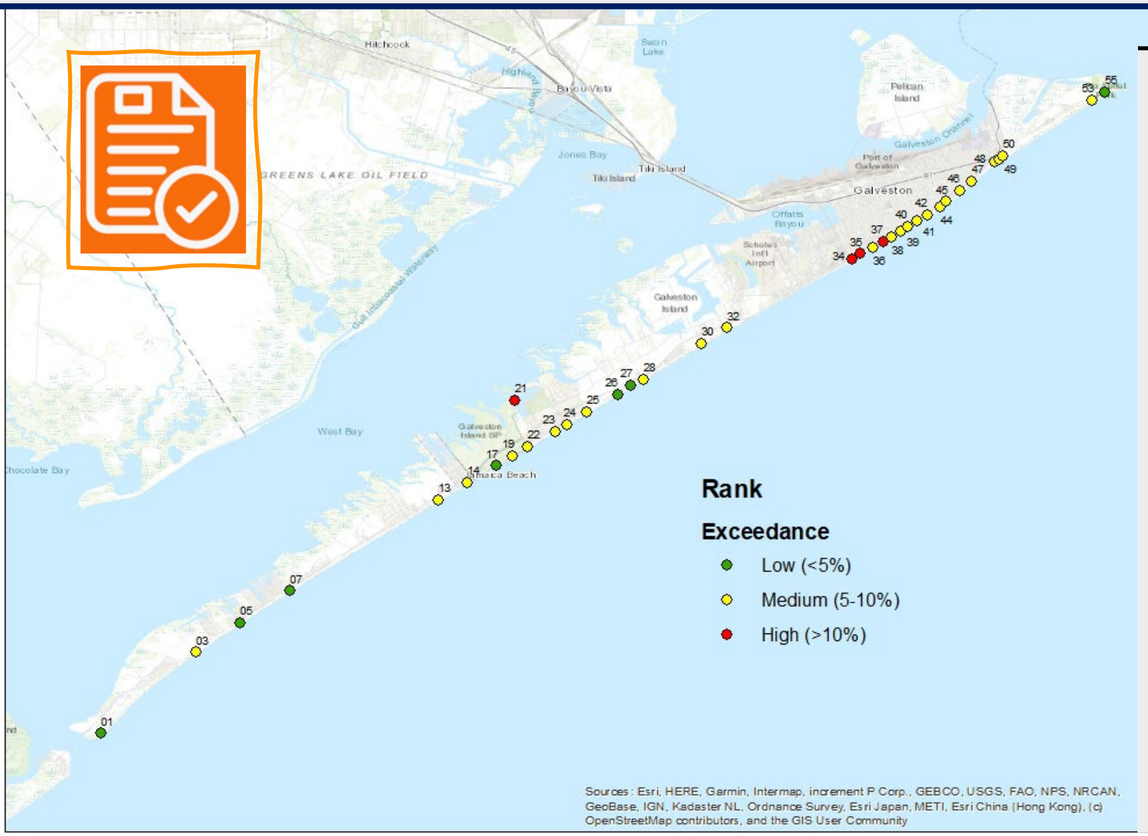


Inverse Distance Weighting tool (IDW): Similar to what was done in previous studies analyzing Beach Watch data for the entire coastal zone of Texas, i.e., **comparing side to side yearly maps.** The general spatial pattern is similar, but some stations have peaks in different years (e.g., Sea Wall stations); this is consistent for both variables.

Ranking of Beaches and Sampling Stations - T1D3

- Ranking of beaches and sampling stations based on levels of bacterial pollution, i.e., the exceedance percentage calculated in T1D2
- Based on the recreational water quality limit of 104 MPN/100mL.
- Three categories: **low (< 5%), medium (5 – 10%), and high (> 10%)**; as done by Powers et. al. (2021) for the entire Coastal Zone

$$\text{Exceedance} = \frac{\text{Count}>104}{\text{Total Count (n)}} \times 100 = \text{Percentage of Exceedance}$$



Out of the 36 stations

- 11% fell in the High category (RED)
- 69% in the Medium category (YELLOW)
- 19% in the Low category (GREEN).
- Seawall were High or Medium
- Bay was High (Station #21)
- The "far" West End (#1-7) resulted Low, except for Station #3 which was Medium
- Most stations and beaches in the "near" West End (#13-32) and the East End (#48-53) were Medium, and none were High

ID		% Exceedance	
Site	Beach	Site	Beach
1	1	4.13	4.61
3		5.07	
5	2	4.60	4.58
7		4.55	
13	3	6.86	6.86
14	4	6.04	6.06
17	5	4.23	4.66
19		5.10	
21	6	14.31	14.31
22	7	5.52	5.67
23		6.63	
24		5.12	
25		5.50	
26	8	4.55	5.18
27		4.96	
28		6.01	
30	9	7.33	7.06
32		6.78	
34	10	10.48	10.01
35		10.44	
36		8.92	
37	11	10.77	8.96
38		9.35	
39		7.65	
40		8.15	
41		9.92	
42		7.82	
44	12	8.04	7.94
45		7.67	
46		7.99	
47		8.03	
48	13	7.82	7.17
49		7.45	
50		6.21	
53	14	5.59	4.97
55		4.33	

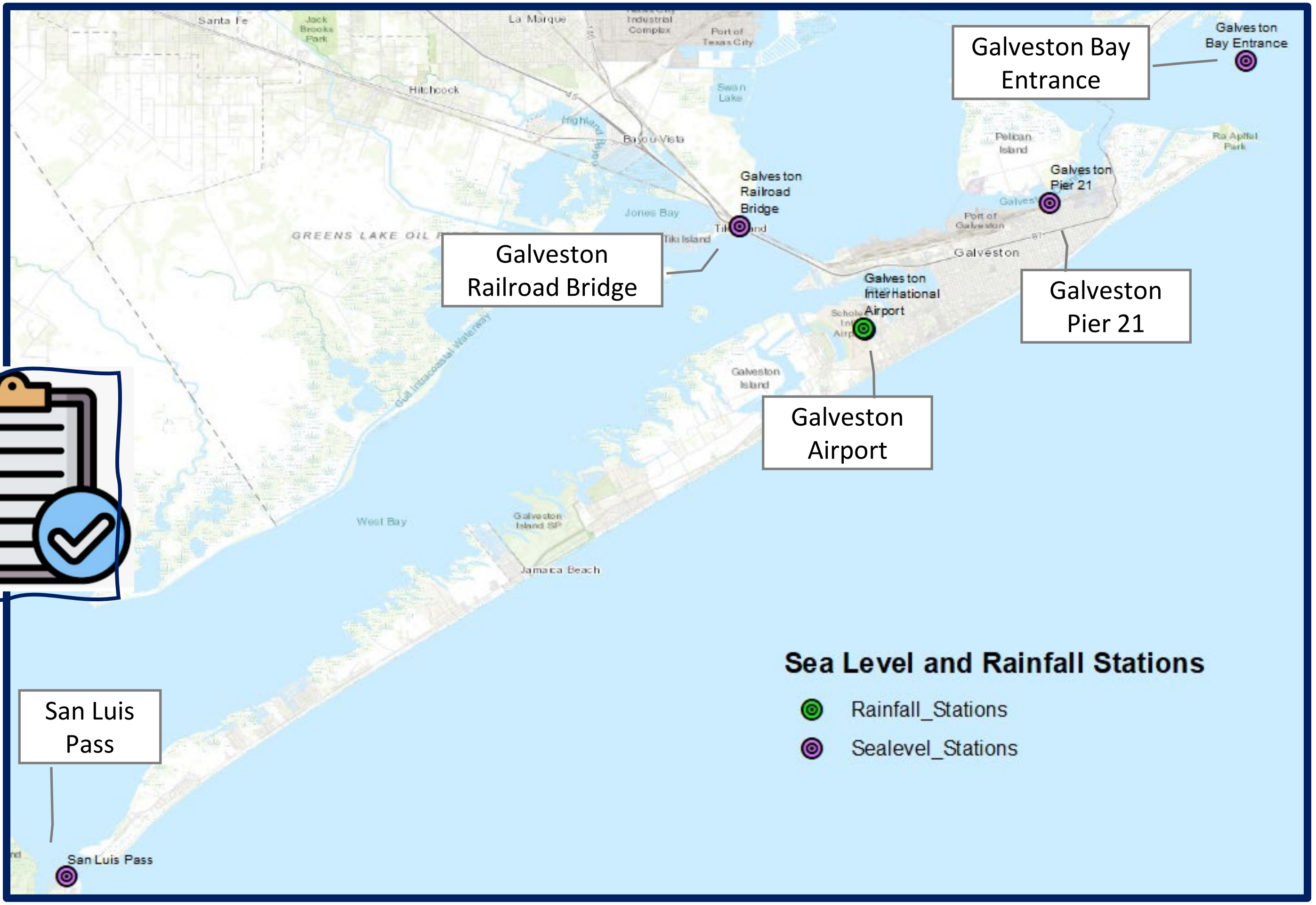
Environmental Metadata - T2D1

Report on:

- Rainfall and sea level data
- Six datasets (1 for rainfall and 12 for sea level)
- This report describes sources, format, and processing steps

RAINFALL

- TexMesonet
- Galveston Airport: Master_Rainfall_Data_2009_2022.XLSX
- 1/15/2009-3/23/2022 recorded in 24-hour daily intervals



SEA LEVEL

Extracted from the GCOOS

Recorded water level at 6-minute intervals, begin on 3/3/2014, end at different dates

1) Galveston Railroad Bridge: 8771486_Sea_Level.XLSX (ends 2/4/2022)
2) Galveston Bay Entrance: 8771341_Sea_Level.XLSX (ends 12/4/2019)
3) Galveston Pier 21: 8771450_Sea_Level.XLSX (ends 2/4/2022)
4) San Luis Pass: 8771972_Sea_Level.XLSX (ends 2/4/2022)

Data have flags for Quality control tests based on Timing/Gap, Syntax, Location, Gross range, Climatology, Spike, Rate of change, and Flatline.

Flags: Good, Untested, Suspect, Bad (*"failing one of the QC tests and deemed as inadequate"* → REMOVED: accuracy = 99.8, 99.9, 99.9, 100.0

Extracted from NOAA

- Same four stations, 6-minute intervals, in meters
- Range 01/15/2009-2/23/2022, less gaps
- MSL and STND datasets
- NOAA_MASTER_WATER_LEVEL_DATA_Station_*_MSL/STND
- **Used for this project analysis**

Statistical Outputs from Enterococci Dataset and Environmental Metadata Comparisons (T2D2)

Report on:

- Enterococci dataset from Task 1 vs Environmental dataset from Deliverable 1 of Task 2.
- Two sources of water sea level: Dataset 1 has limitations (gaps, period); Dataset 2 covers the entire project period of 2009-2022, thus it was useful for more extensive statistical analysis

Methods

Environmental Metadata

Looking at trends and comparing datasets with T-tests

Enterococci dataset vs environmental metadata

General visual pre-comparison (Station #34) to observe individual data

Correlation:

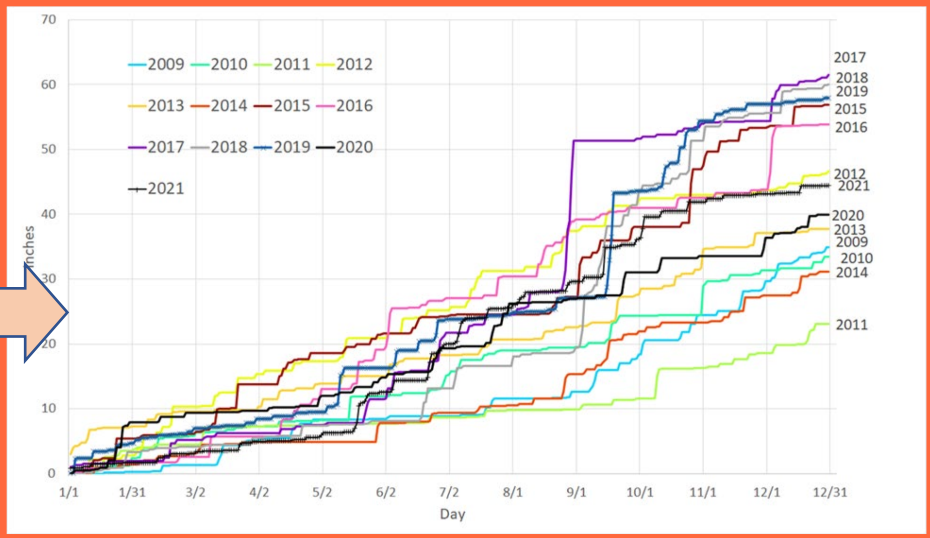
- Kendall's, Pearson's
- Correlation **strength**: ±(0 - 0.10), very weak, ±(0.10 - 0.19), weak, ±(0.20 - 0.29), moderate, and >± 0.30: strong
- Also tested 2/7-days **rainfall sums**

Results

Data Analysis

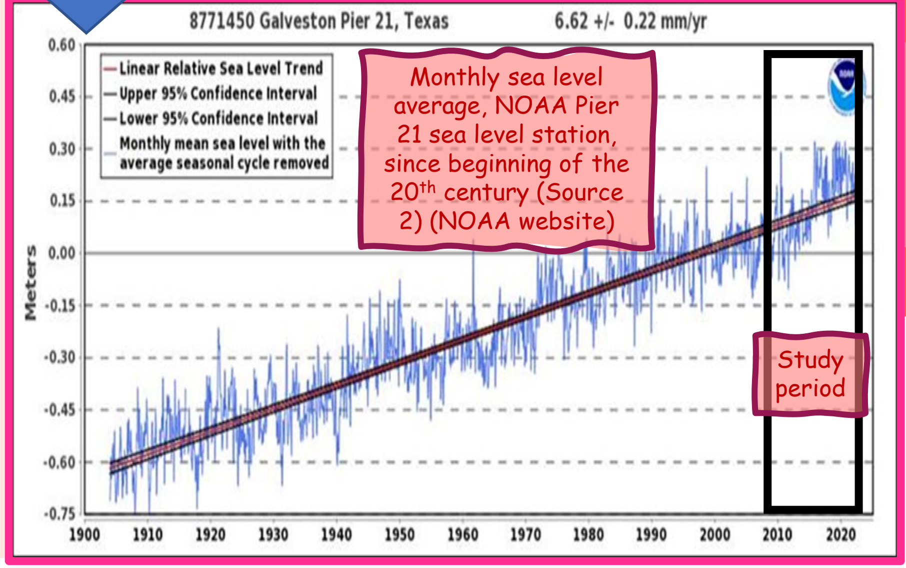
Rainfall

- **Peaks**: October 2015, December 2016, August 2017, and September 2018 and 2019
- **Driest** January to August: 2009, 2011, 2014.
- Typical **trend**: High precipitation in the second part of the year.



Sea level

- The means of the **four stations resulted different from each other using two-sided pairwise t-tests**
- Linear regressions of sea level for all project period showed an **increasing trend**, consistent with what observed on the long-term NOAA dataset



Monthly sea level average, NOAA Pier 21 sea level station, since beginning of the 20th century (Source 2) (NOAA website)

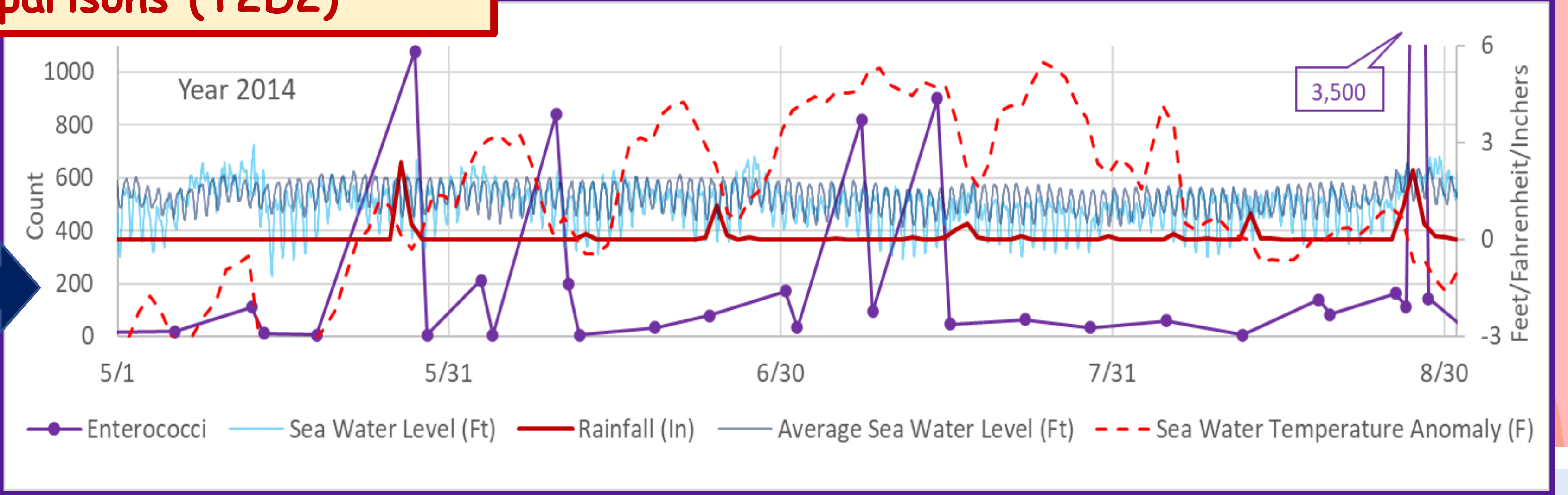
Monthly average water level shows **two picks**:

- March-April
- September-October

Statistical Outputs from Enterococci Dataset and Environmental Metadata Comparisons (T2D2)

General Visual Pre-Comparison (#34)

Rainfall and sea level could explain some but not all the observed *Enterococci* peaks. Other factors could include sea water temperature



Kendall Tau-B (Kendall) Correlation

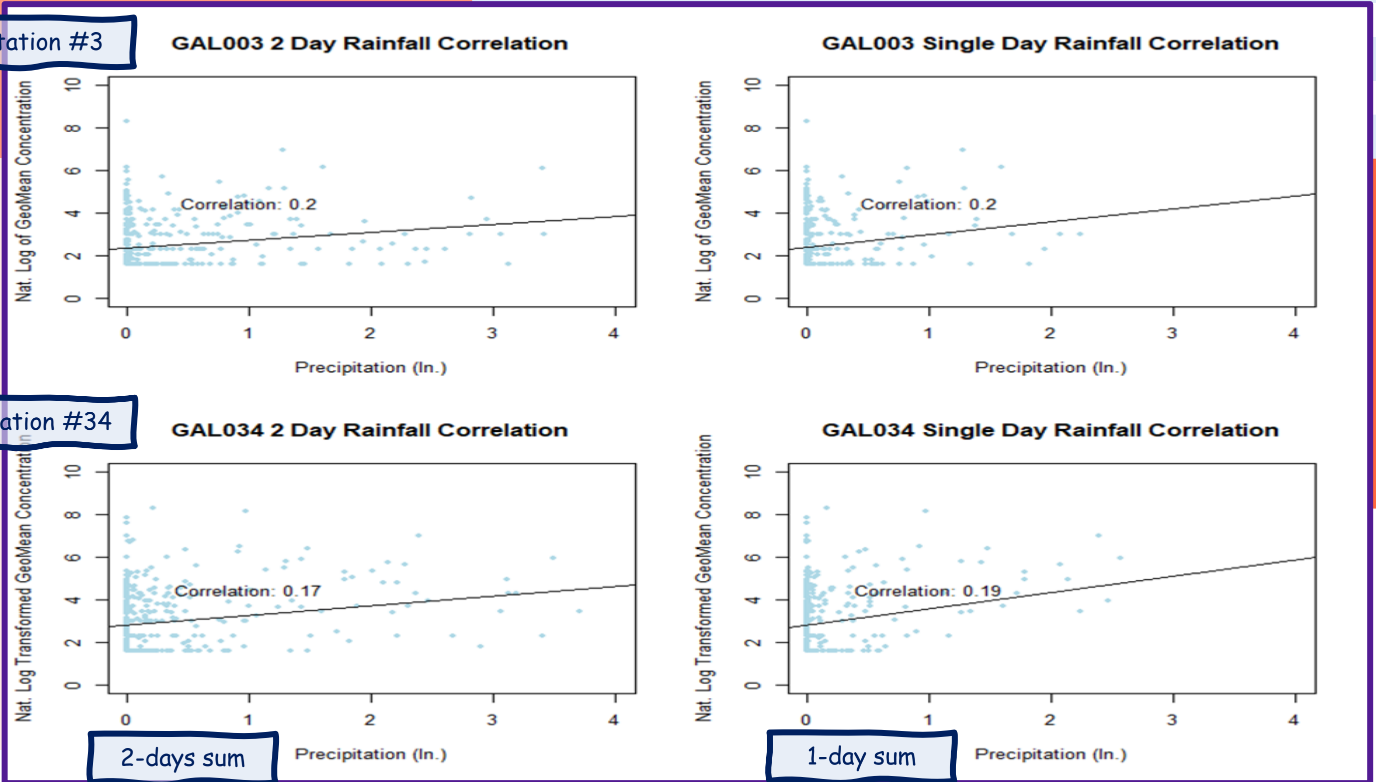
Rainfall

- Slightly better correlation with **2-days sum** on the West End and with the 1-day sum on the other Zones.
- **Largest rain events** (about > 2 inches) always correlated with an *Enterococci* result higher than the minimum level of detection.

Sea level

- Correlation is significantly different from 0, especially for sampling Station #21.

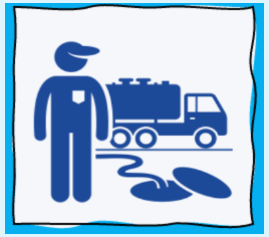
Correlation **coefficients were weak** in most cases



WATERSHED WITH MARKED POTENTIAL SEWAGE CONTAMINATION SOURCES

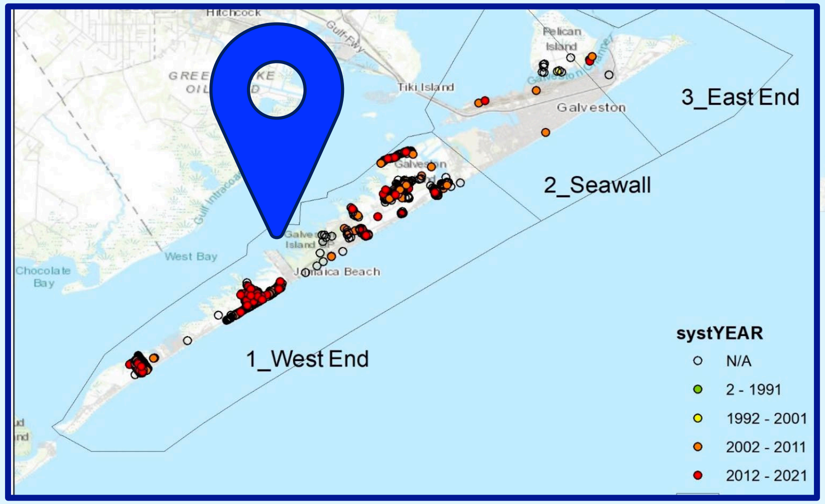
T3D1

Analysis of micro-watersheds and potential sources of pollution

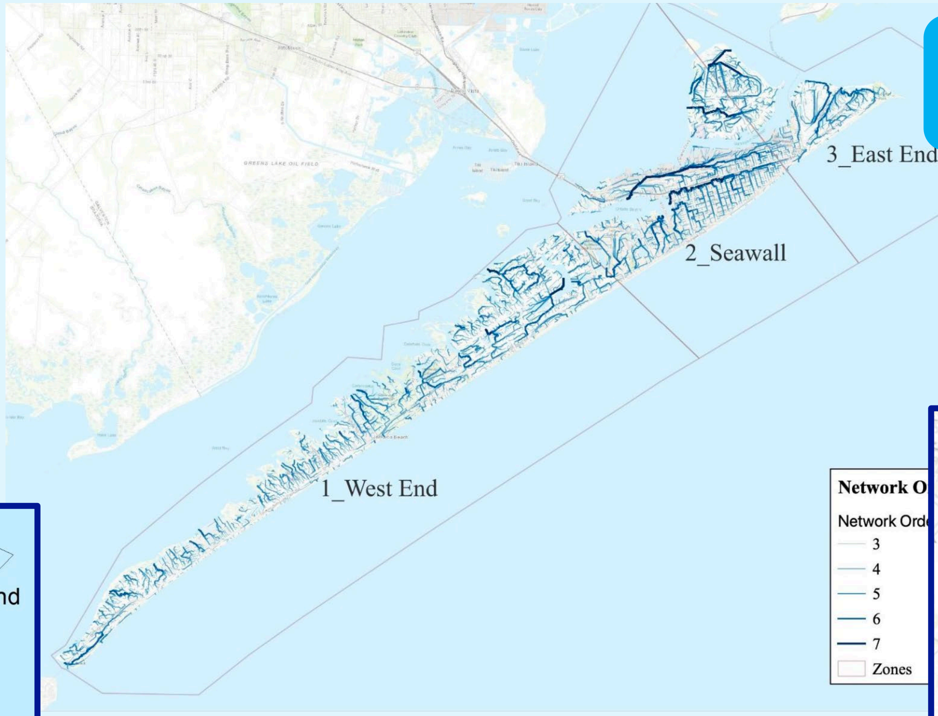


On site sewage facilities (OSSF)

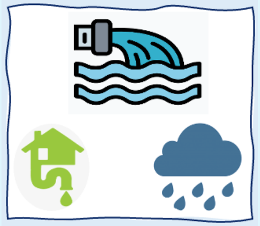
- Mostly on the West End
- Older systems are in the "far West", "far East", and coastal portions of West End



Micro watershed map with LiDAR (2018 Light Detection and Ranging)



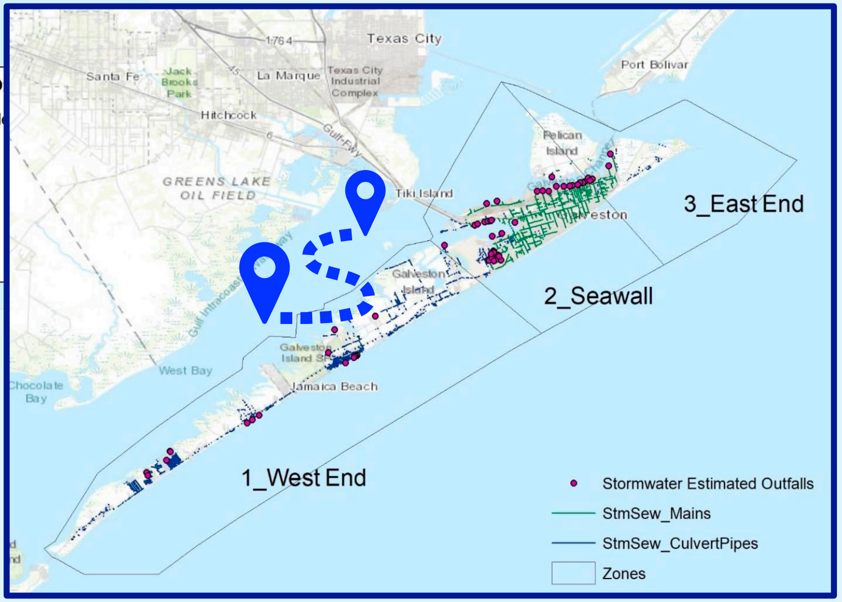
Most water during rain events likely drains in the bay



City of Galveston Stormwater & Wastewater Treatment Plants (WWTP) outfalls

Bay side:

- Most Stormwater outfalls
- All WWTP outfalls

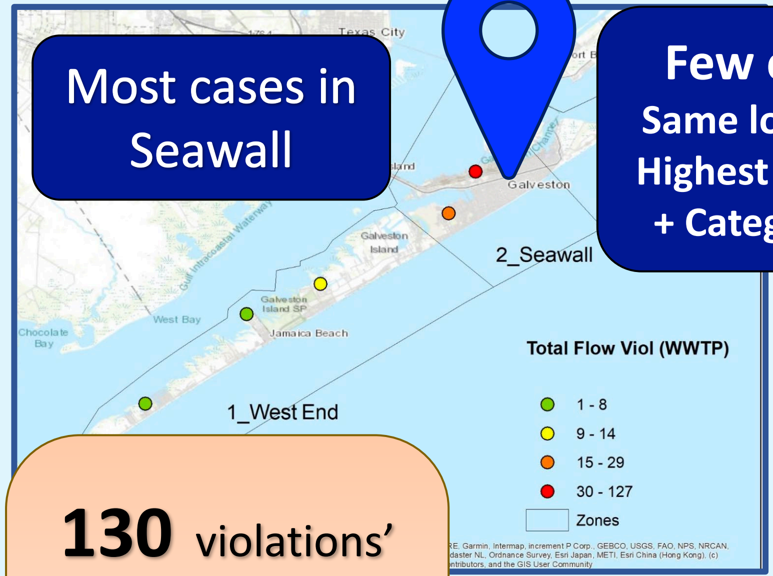


WATERSHED WITH MARKED POTENTIAL SEWAGE CONTAMINATION SOURCES

T3D1

Analysis of leaks and spills

Most cases in Seawall



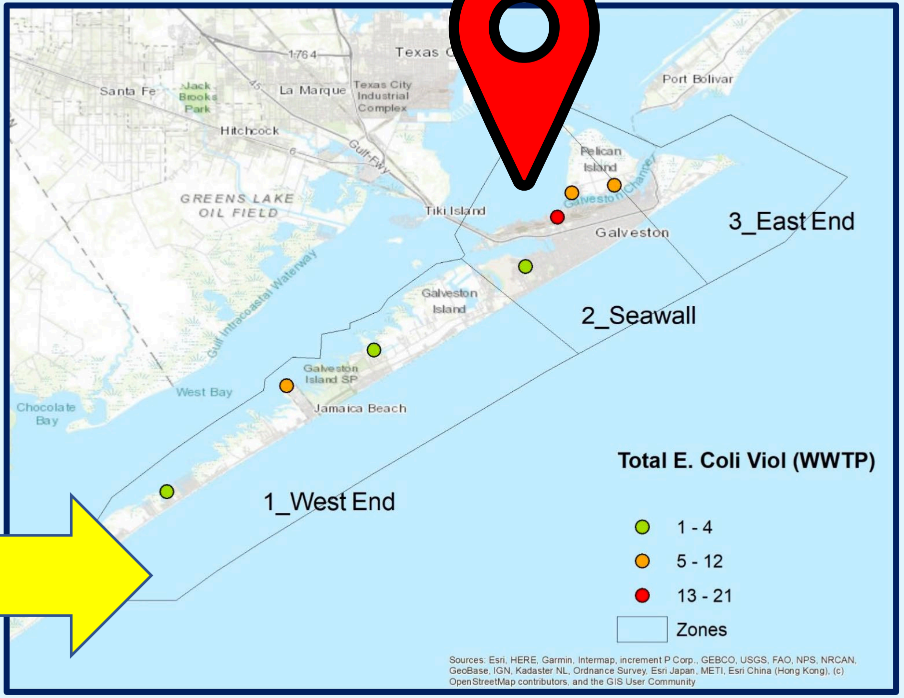
Few cases:
Same location +
Highest volumes
+ Category "1"

Flow violations

From Sanitary Sewer Overflows (SSO) database

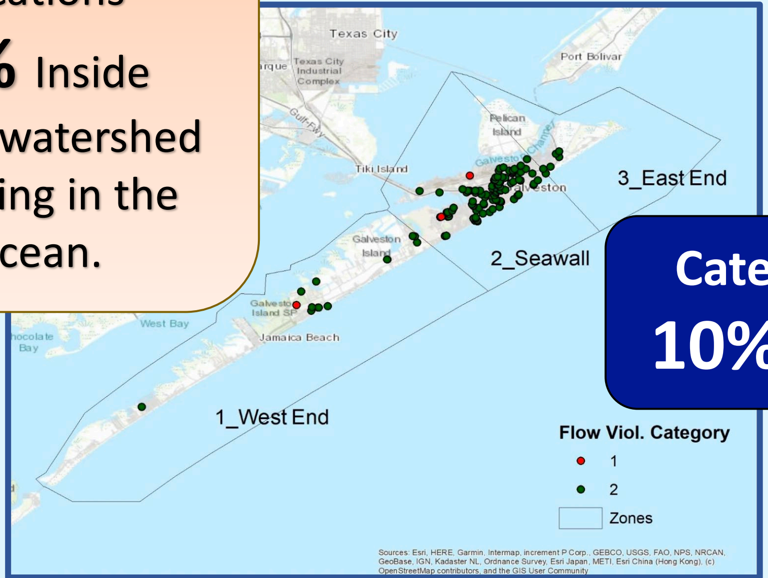
E. Coli violations

From the Enforcement and Compliance History Online (ECHO) database



130 violations' locations

3% Inside micro-watershed draining in the ocean.



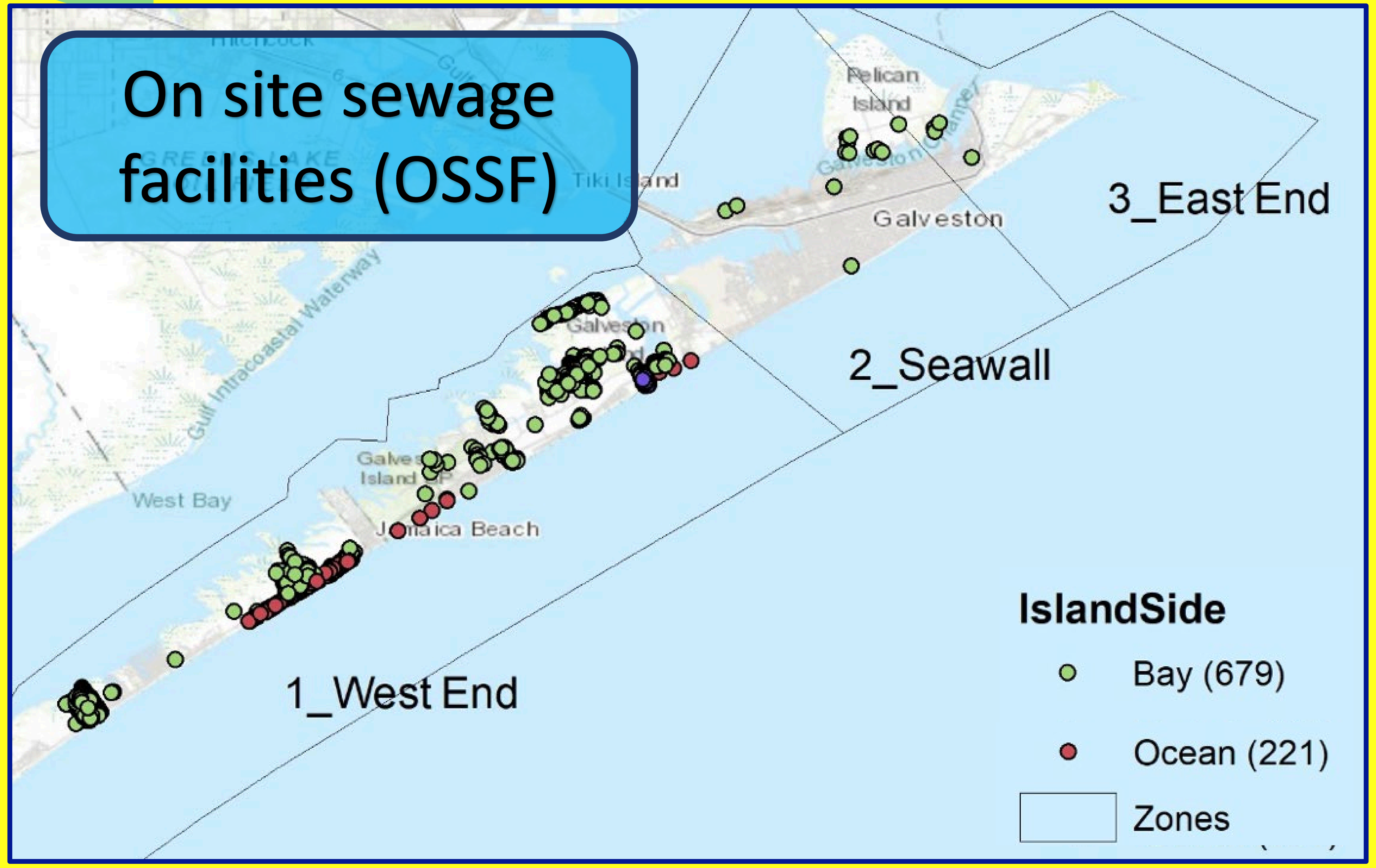
Category "1" = >
10% of max flow

Most E. Coli violations in Seawall Zone & Jamaica Beach WWTP (West End)

STATISTICAL OUTPUTS: ENTEROCOCCI DATASET & POTENTIAL SEWAGE CONTAMINATION SOURCES

T3D2

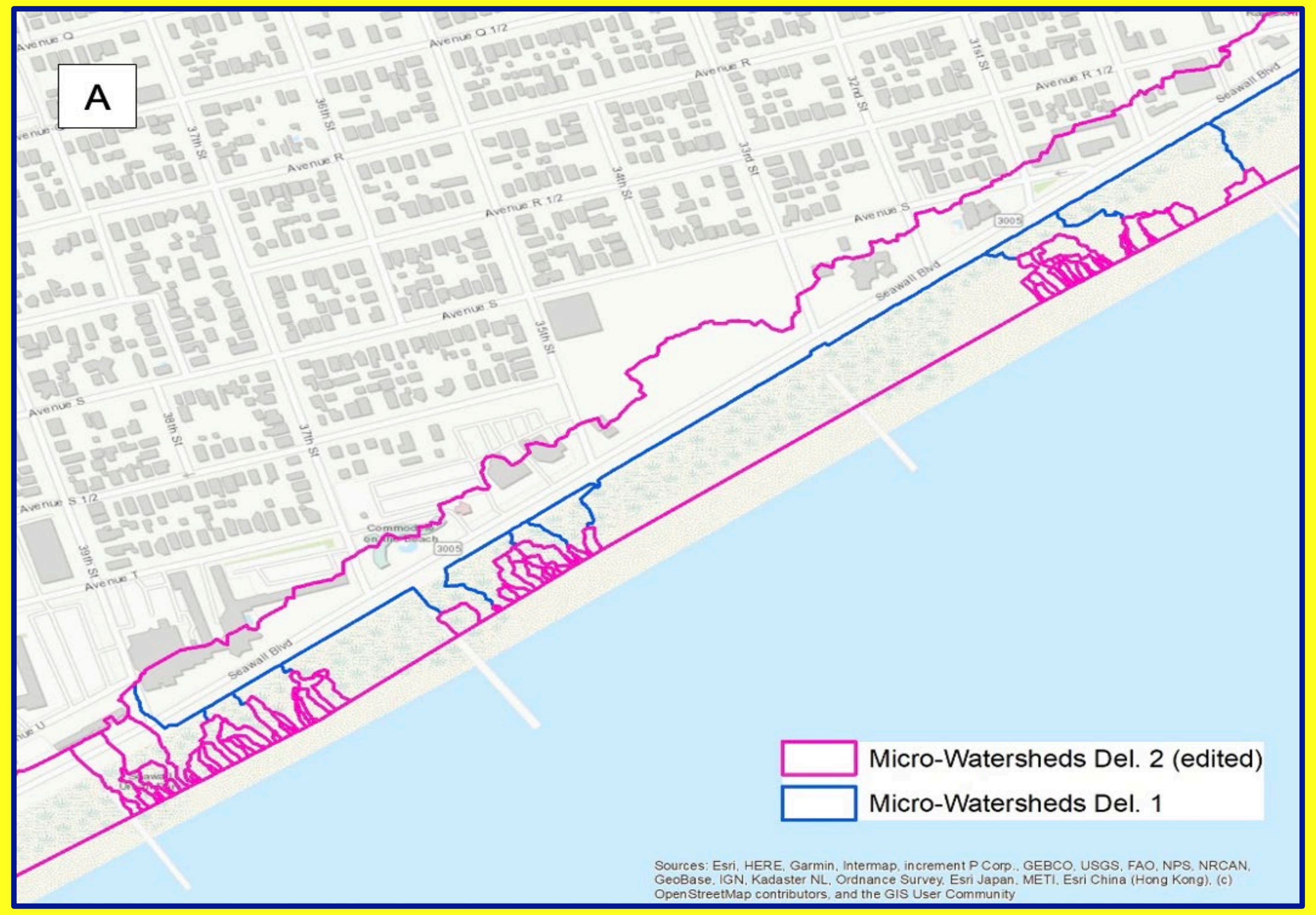
On site sewage facilities (OSSF)



Draining toward the ocean: Mostly in the West End near station **#13** (only "Average" statistic slightly higher)

Little evidence that OSSFs have a strong impact on Enterococci sampling results

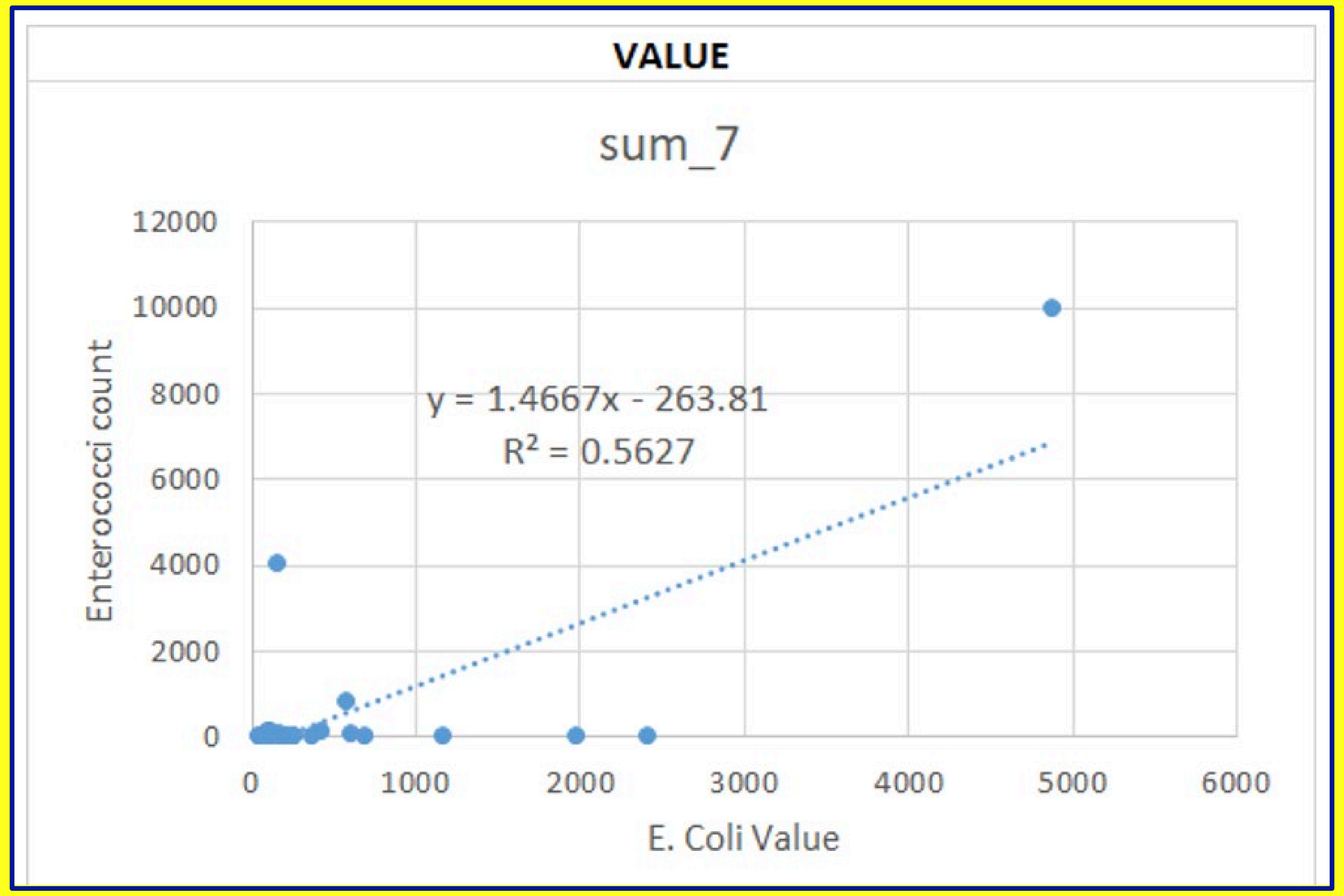
Micro-watersheds adjusted by field verifications (Sea Wall): **Culverts between 18th Street and 39th Street**



Flow violations

- From **911 address**
- Toward the ocean:
  - Sea Wall (Stations # **34**, **35**, **41**, **42**): No evidence
  - West End (Stations # **23**, **24**, **25**): No evidence
- Toward the bay: Station **#21**: No evidence

Violation of E. coli



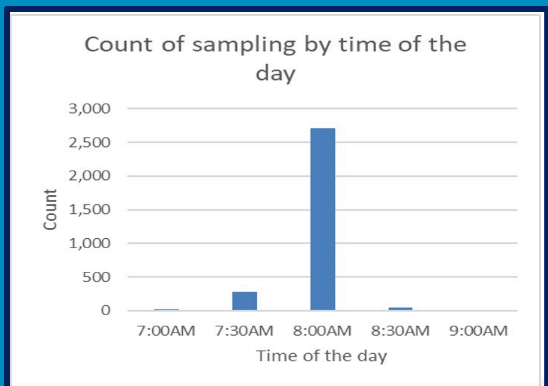
- From **WWTP**
- *All toward the bay (Station #21): **Significant correlation** with sum of Enterococci counts in the following 7-15 days*

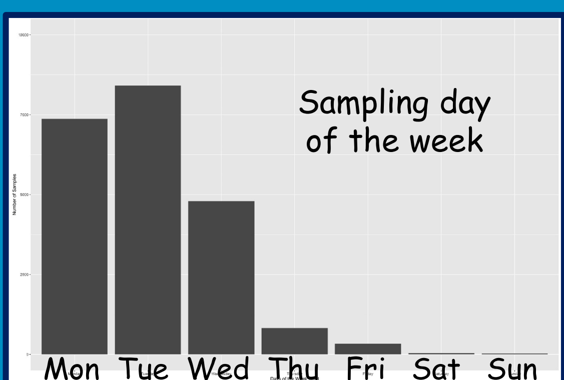
RECREATIONAL BEACH ATTENDANCE ESTIMATES

T4D1

Direct Estimate 1: Texas Beach Watch sampling:

- **8:00 am**
  - Most: "Light" usage
  - "No usage" during COVID closure
  - "Moderate" in June-July 2021



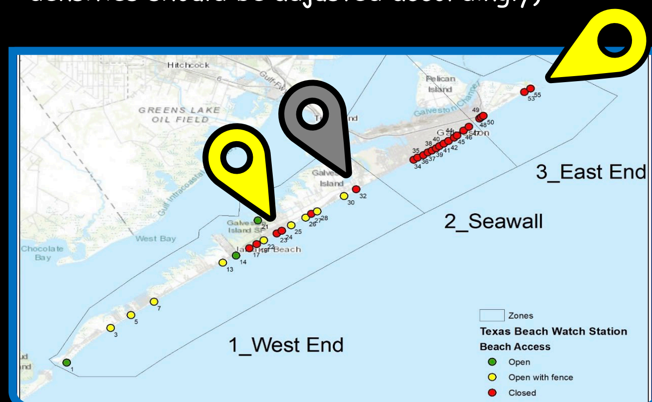


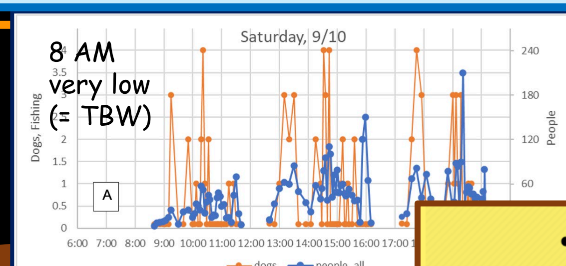
Direct Estimate 2: AgriLife field truth visits

Highest counts:

- **East End**, Steward Beach (#49 and 50)
- West End, Jamaica Beach (#14), near Sea Wall with open access (#25, 26, 28, 30)

Highest continuity: Sea Wall (actual densities should be adjusted accordingly)





- Peaks in early afternoon
- Monday/Week ends = **38%**

This report includes:

- **Direct Estimates**
- Existing: Texas Beach Watch, based on counts conducted at sampling (since October 2019). Load categories:
  - "light" <= 10 people,
  - "moderate" 10-25 people
  - "heavy" > 25 people
- Visits:
  - Counts on all sampling stations all day long (September 10 and 11, 2022) and on selected sampling stations (October 2-3, 2022)
  - Interviews

- **Indirect Estimates**
- Hotel Occupancy Tax (HOT)
  - Park Board (2015-2022)
  - City of Galveston (2009-2022)
- Parking has also been collected via the same two institutions (2015-2022)

Interviews

- Beach closed in 2020 (3/16- 5/1); once re-opened, it was packed again,
- Waves around 6-7 Pm
- Busiest periods (decreasing order): holidays (e.g., July 4th), July, June, May, and March (spring break)

RECREATIONAL BEACH ATTENDANCE ESTIMATES

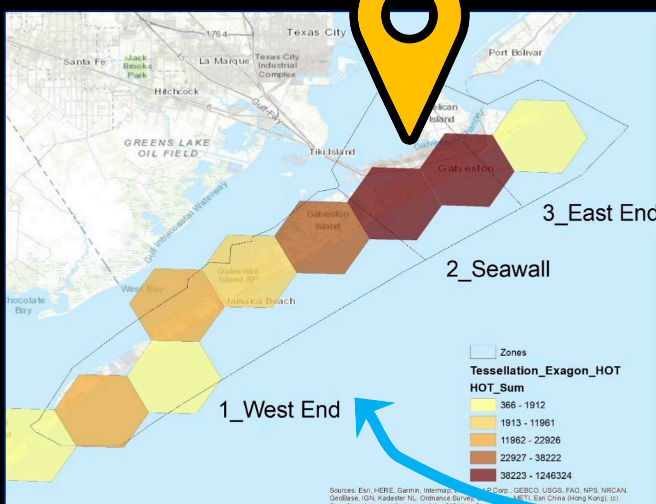
T4D1

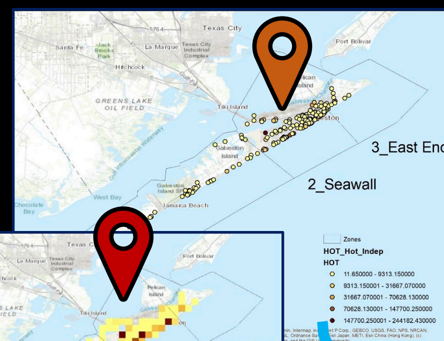
Indirect Estimate 1:

HOT

(Hotels/Full Service, Hotels/Limited Service, and Independents)

- **Peaks** are consistently showing, i.e., **March, June, July**
- HOT locations aggregated into **hexagons** (13 sq mi) **for spatial regression analysis (T4D3)**





- **Higher** totals were found inside the Seawall zone

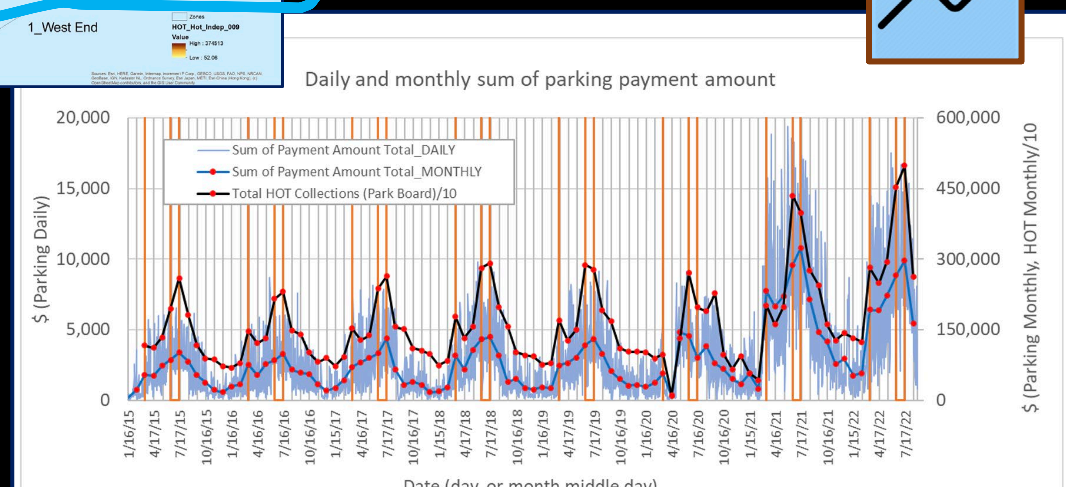
Indirect Estimate 2:

Parking datasets

(Payment Amount Total)

Peaks in March, June-July, and September

Monthly sums were **consistent with Park Board HOT** monthly and yearly patterns

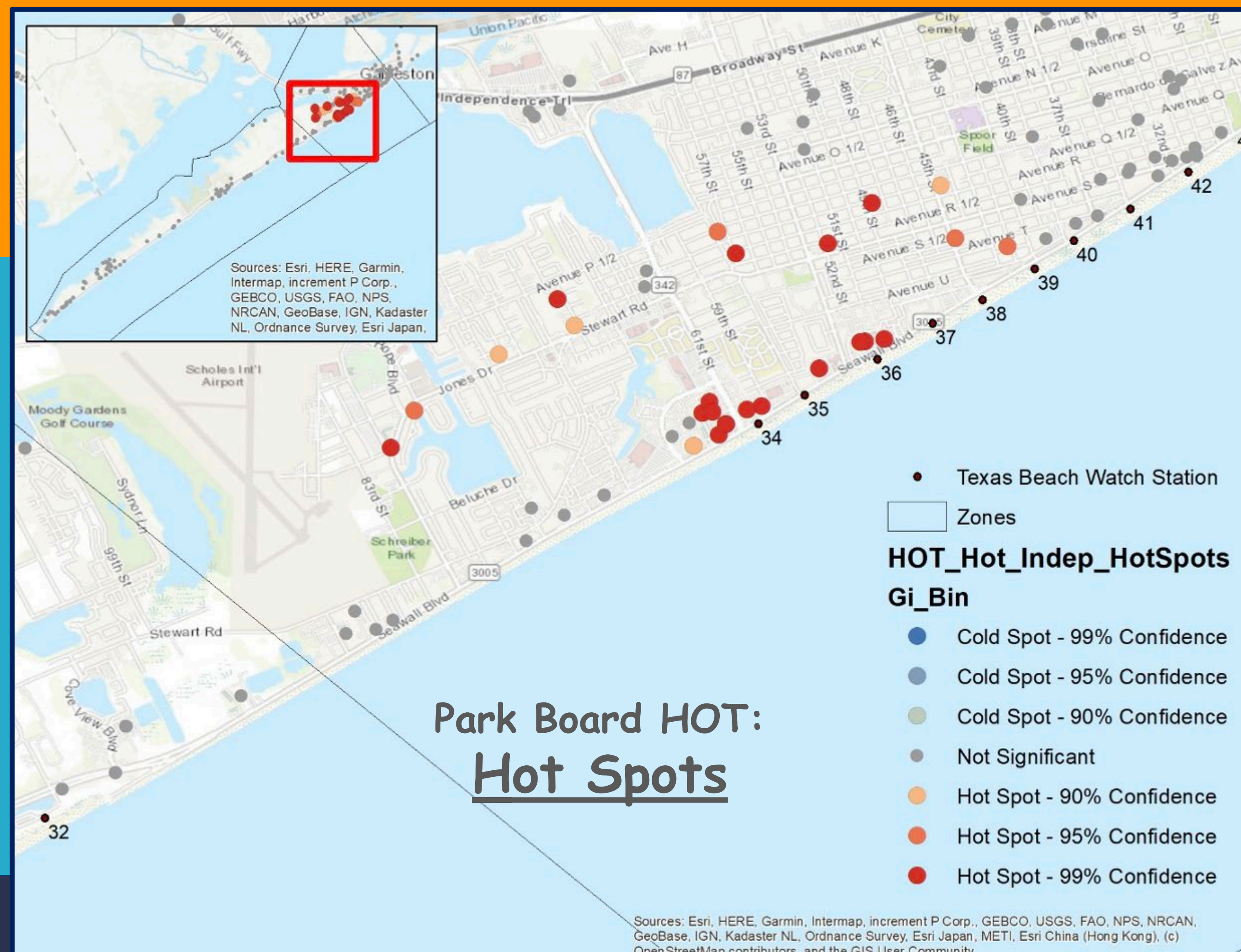


T4D2

CLUSTERING AND SPACE-TIME PATTERN ANALYSES

This report:

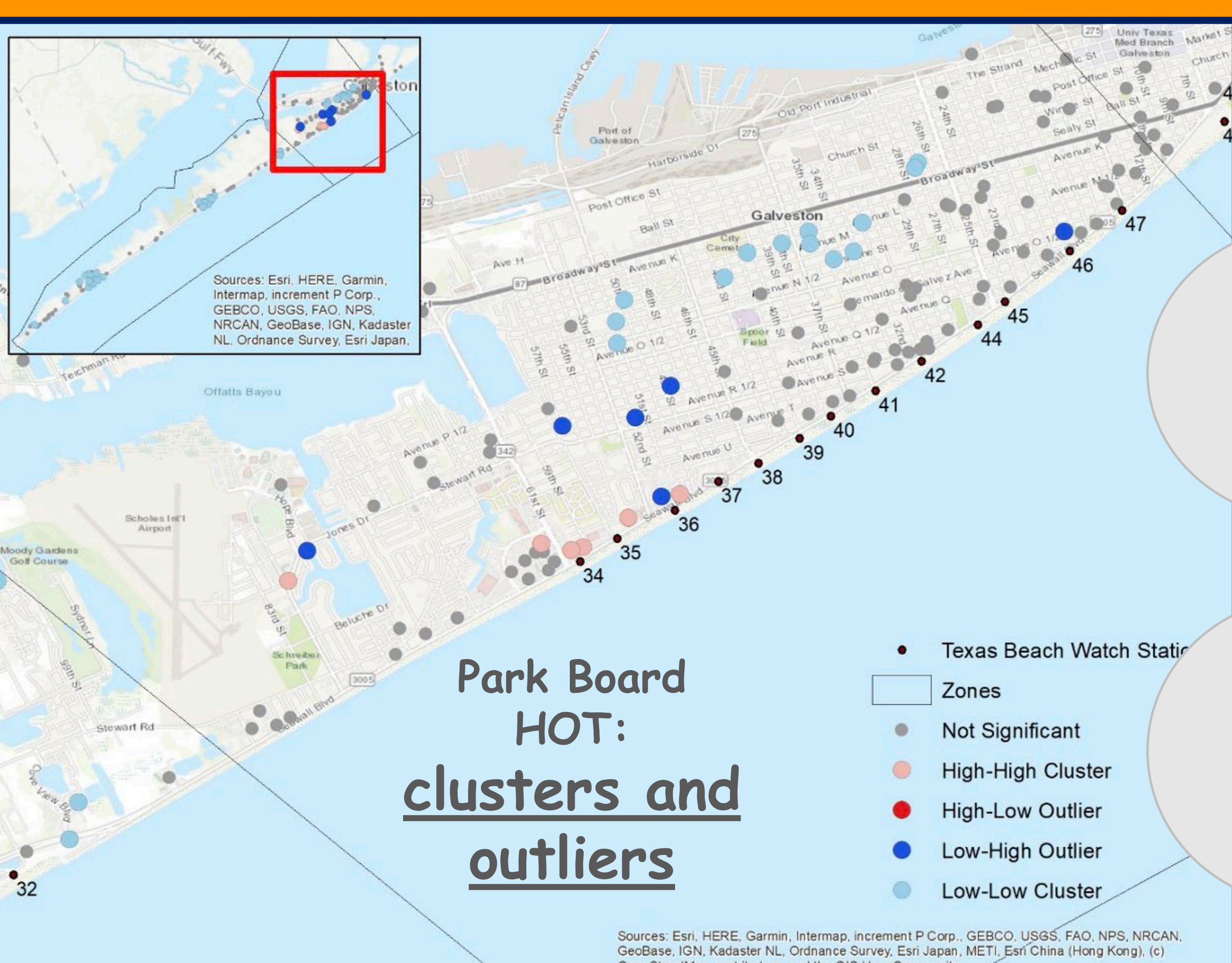
Uses available information related to foot traffic at hotels, as identified in Task 4 - Deliverable # 1 (T4D1)



Stars: Two clusters, one high near sampling station #36, and one low, near sampling station #42

Park Board HOT:

- Hot spot in the Western portion of the Seawall (stations # 34-39)
- Several clusters and few outliers in West End and Seawall.



Methods

- Hotels' Star info as downloaded from Google map
- HOT only for one month (June 2022), for Hotels and Independents structures

Tools :

- Hot Spot Analysis
- Cluster and Outlier Analysis

Statistical Clustering

- Only data for which location and spatial variability was available
- This is the case of Park Board HOT data

Space-Time Pattern

Analysis was not possible due to limitations in all datasets:

- Texas Beach Watch (8 AM only)
- Field truth (one-time)
- City HOT (location not available)
- Park Board HOT (data format)
- Parking (location not available)

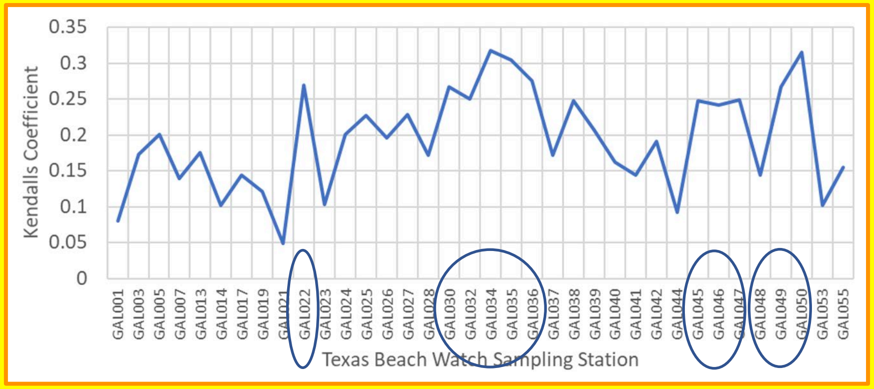
Statistical Outputs from Enterococci Dataset and Estimated Recreational Beach Attendance

This report includes Enterococci concentrations and estimated recreational beach attendance using available information related to foot traffic at hotels, as identified in Task 4 - Deliverable # 1

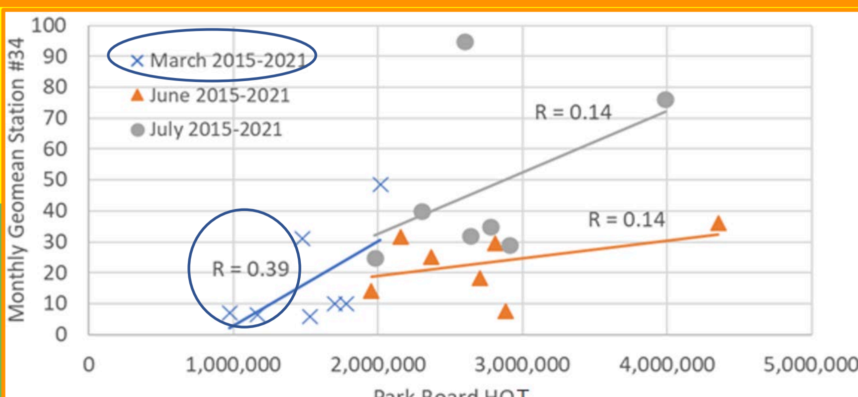
T4D3

Correlation

- Park Board monthly **HOT (all structures)** from 2015 to 2022 vs **monthly Enterococci geomean** from Task 1 for the period 2015-2021, stations by station (Kendall coefficient)
- Best correlation (Kendall) found for sampling stations **#22, 30-36, 45-47, and 49-50.**



Individual months for selected stations: best correlation for station #34 in March **(R=0.39)**

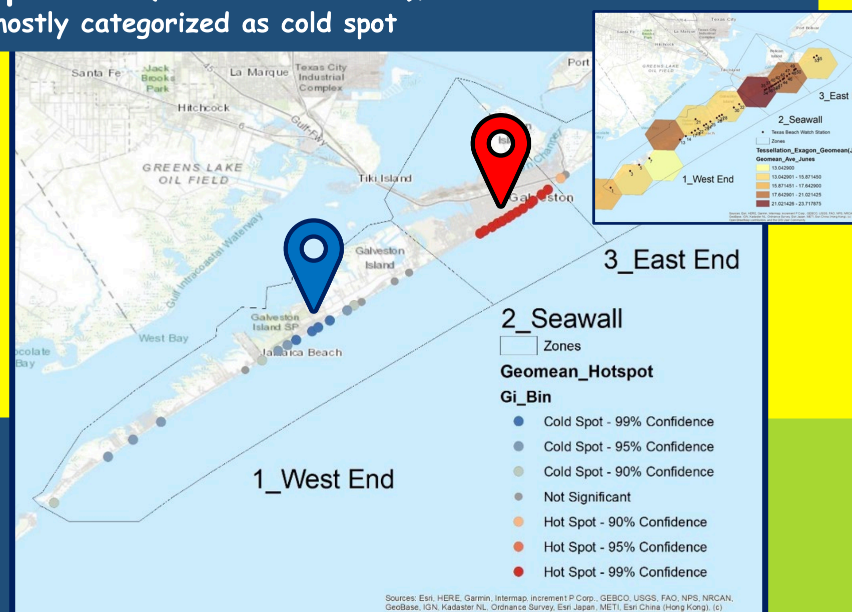


Spatial Regression

- Park Board monthly **HOT (Hotels and Independents structures and June 2022 only)** vs **monthly Enterococci geomean** from Task 1.
- ArcGIS tool: Geographical Weighted Regression (**GWR**)
- Aggregation of data into **hexagons** (13 sq mi)

Preliminary steps (Tessellation, Cluster and Outlier analysis, Hot Spot analysis):

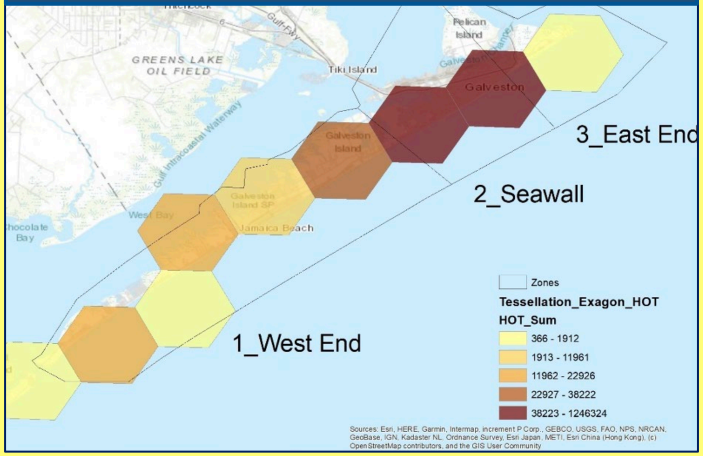
- Park Board <u>HOT</u> **Hotels and Independent structures** (see T4D1)
- <u>**Enterococci**</u> **data:** Entire Seawall is identified as **hot spot** area (**99%** confidence), while the West End is mostly categorized as cold spot



Statistical Outputs from Enterococci Dataset and Estimated Recreational Beach Attendance

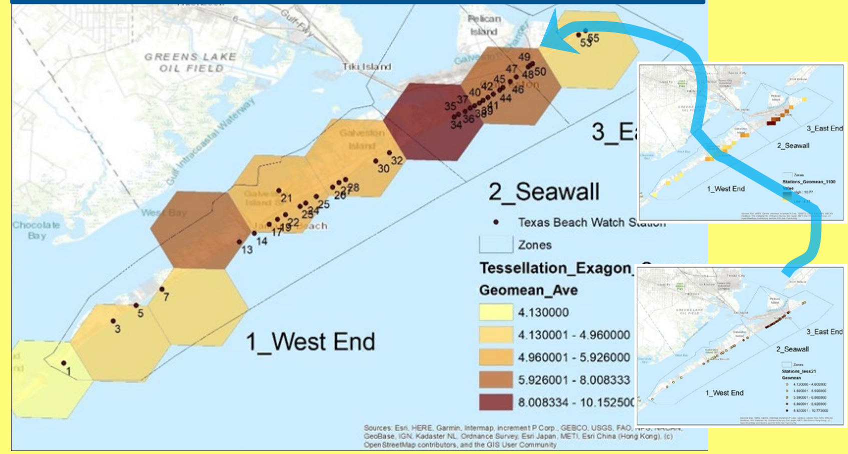
T4D3

HOT Hotels and Independent July, 2022 (T4D1)



+

Enterococci Year-round, 2015-2021

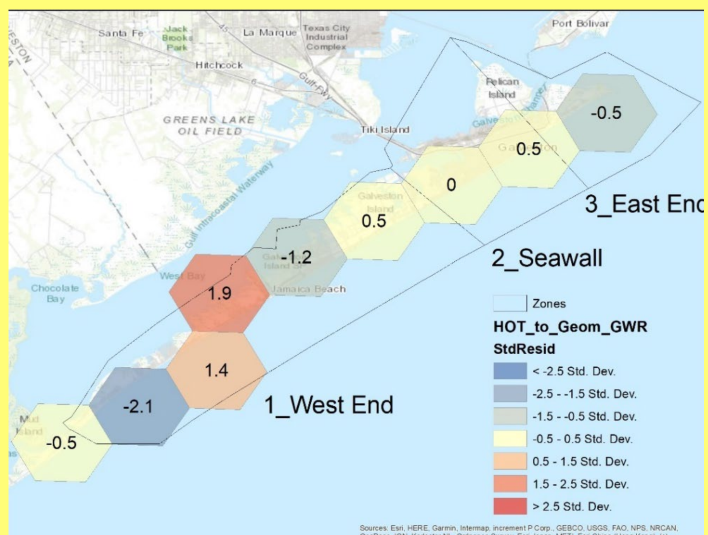


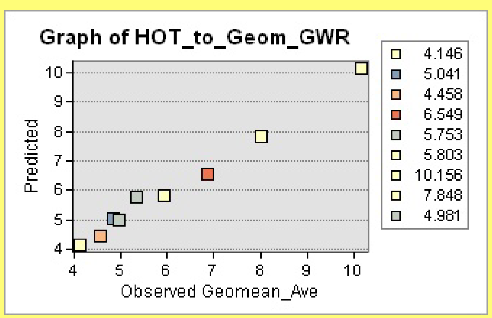
As for HOT (T4D1), **aggregation** of data to hexagon tiles is consistent with individual stations pattern

Comparison of Park Board HOT and Enterococci

- Predictions' **confidence** was always **higher than 95%** (standardized residuals < ±2.5)
- Maps of standardized residuals showed **best agreement in the Eastern part of the island**

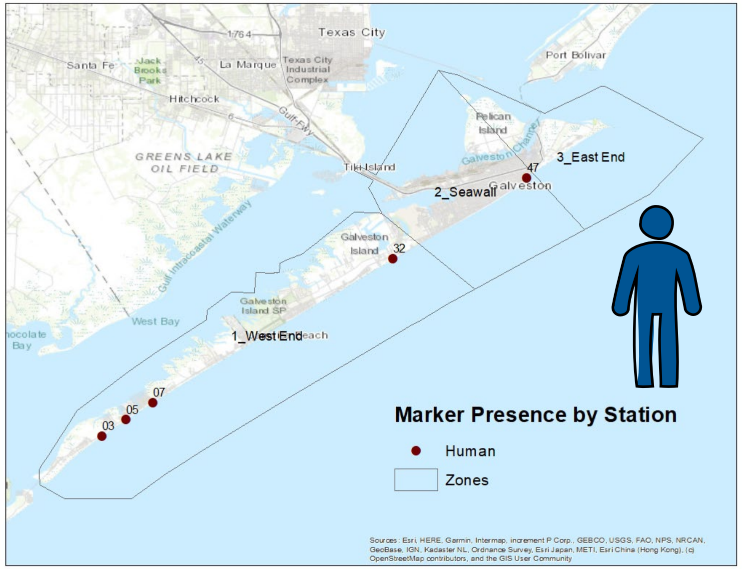
=

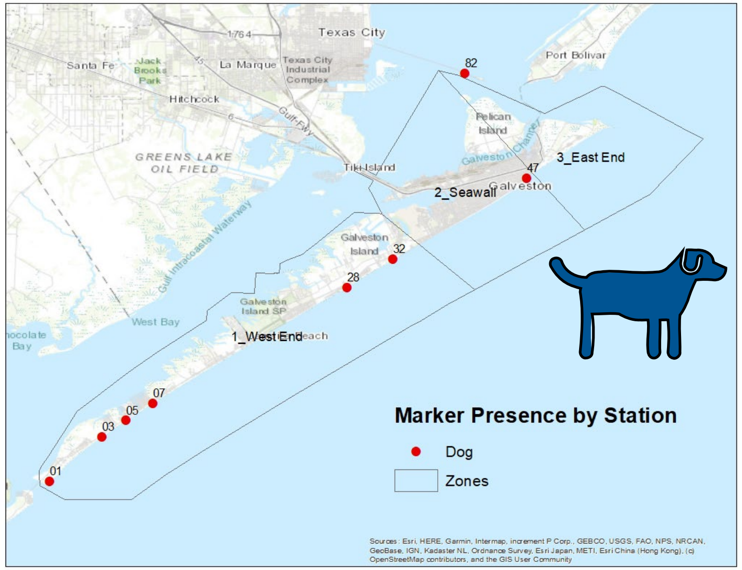


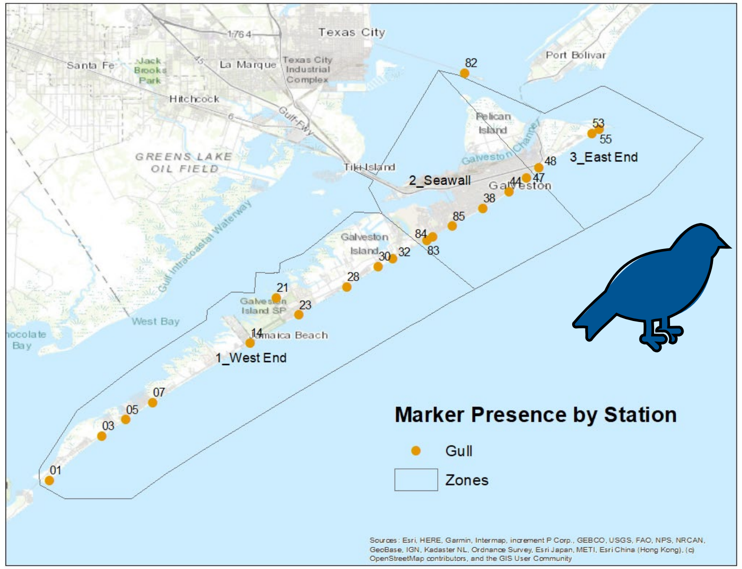


T5D1

Environmental, Enterococci, and Molecular Marker Dataset







Marker Analysis Results

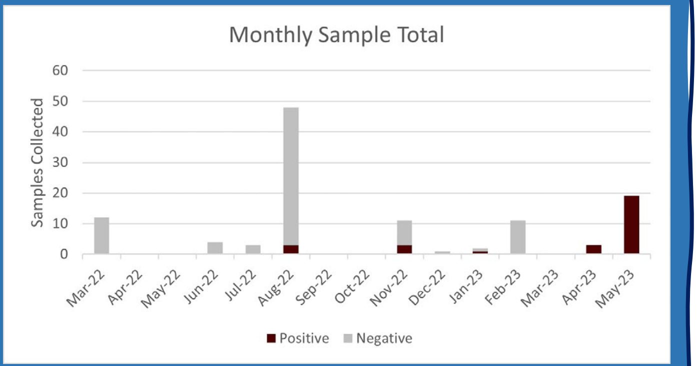
Human marker detected across the island at low levels below the limit of quantification <u>except for GAL032 in West End.</u>

Dog marker found at relatively low levels and largely <u>at same sites as human marker.</u>

Gull marker was the most frequently detected with <u>copy numbers much higher than human or dog markers.</u> Highest gull maker numbers observed at stations in the East End.

<u>Project Timeline and Sampling</u>:

- Samples collected from stations at West End, Seawall, and East End.
- Sampling from March 2022 – May 2023 with 114 samples above 104 MPN/100mL included in the dataset.
- Stations GAL083, GAL084, and GAL085 were new stations added in the middle of the year that are still included in the dataset.



T5D1

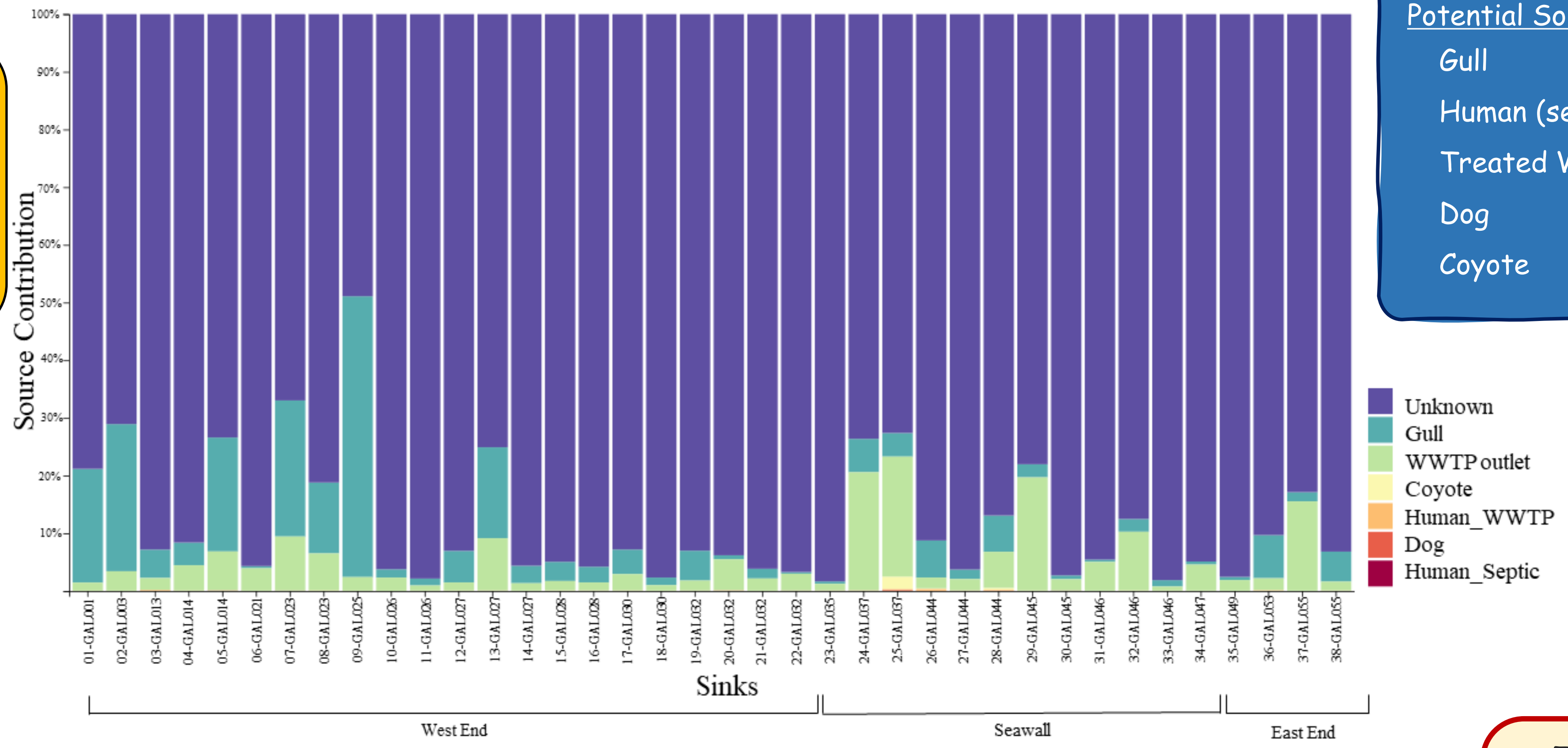
Environmental, Enterococci, and Molecular Marker Dataset

DNA-based Microbial Source Tracking

Gull was the most frequently detected and largest source overall

Widespread, but low (<1%) levels from human sources (WWTP, septic)

Low levels from dog and coyotes



Samples from July and August were used for DNA-based source tracking. This process compares the bacterial community in potential sources (e.g., sewage) to that in environmental sinks (e.g., water).

Potential Sources Examined:

- Gull
- Human (septic and WWTP sewage)
- Treated WWTP effluent (outlet)
- Dog
- Coyote

"Unknown" sources largely represent naturally occurring marine organisms such as cyanobacteria

Treated WWTP effluent (outlet) impacts detected at several stations especially the Seawall and East End

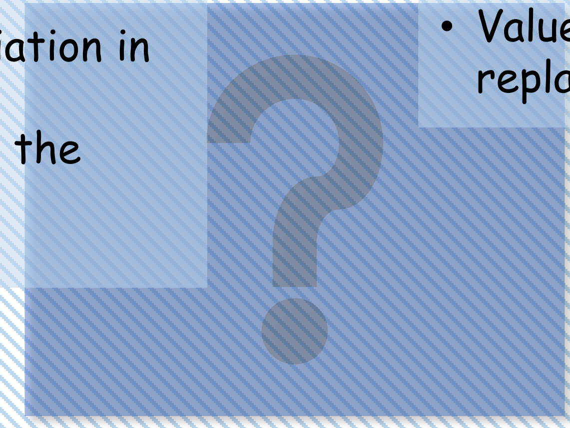
T5D2

Outputs from Statistical Relationship Tests

Enterococci result were not found to correlate with marker abundance

Source	Kendall Tau Correlation
Human	-0.142
Dog	0.296
Gull	0.07

- Samples positive for a qPCR marker were compared to enterococci levels using the Kendall Tau method
- No correlation was found
- Lack of correlation may be due to variation in the persistence of the markers in the environment and differences between the target organisms' ecology and that of enterococci.



Environmental variables were not found to correlate with marker abundance

- Rainfall data, as in previous tasks, was correlated to human marker abundance using rolling seven-day sums
- Water level data, unlike in previous tasks, was correlated using the average for the day
- Values below 430 (limit of quantification) were replaced with 215 (half of the limit)