Alaska Department of Environmental Conservation Virtual Beach Waterbody Modeling Report Ketchikan and Kenai Beaches, Alaska

Prepared by: J. Petitt[1]

Abstract

During summer 2022, the Alaska Department of Environmental Conservation pilot tested the U.S. Environmental Protection Agency's indicator pathogen predictive model building software, Virtual Beach V. 3.0.7 at eight Alaskan beaches. Pathogen levels were predicted for the date and time of routine sample collection at six Ketchikan beaches and two Kenai beaches over the recreational season. Samples were collected at each Kenai Beach four times over the recreation season before, during, and after the July personal use fishery. Samples were collected six times at each Ketchikan beach from June through September. Predictions were made for each routine sample collected using a gradient boosted model and multiple linear regression. These predictions were compared to their concurrent routine sample results to ground truth the predictive models. The beaches with the lowest performing gradient boosted models were Rotary Park Pool in Ketchikan and North Kenai Beach. Multiple linear regression performed slightly better at Rotary Park and North Kenai beaches, with comparably acceptable model performance at North Kenai Beach (75% correct). The beaches with the best performing gradient boosted models were for Ketchikan at Mountain Point Cultural Foods and South Point Higgins. Using either gradient boosted models or multiple linear regression in Virtual Beach to predict pathogen levels may be a substitute for monitoring at all beaches studied except Rotary Park Pool in Ketchikan.

[1] Nonpoint Source Pollution, Water Quality, Division of Water, Department of Environmental Conservation

Basic Waterbody Information

Table 1. Beaches Information, Kenai Alaska

Region	Kenai Peninsula	
Assessment Unit ID	AK_B_2030218_002	AK_B_2030218_003
Assessment Unit Name	North Kenai Beach	South Kenai Beach
Location description	North and South Kenai Beaches located at the outlet of the Kenai River into Cook Inlet	
Hydrologic unit code	190203021907	
Water Type	Marine	
Area sampled	Sites within the boundary of the personal use dipnet fishery	
Time of year sampled	June through September	

Table 2. Beaches Information, Ketchikan Alaska

Region	Ketchikan					
Assessment Unit ID	AK_B_1010205_003	AK_B_1010208_001	AK_B_1010204_007	AK_B_1010204_006	AK_B_1010204_008	AK_B_1010204_003
Assessment Unit Name	KB-Herring Cove	KB-Mtn Pt Cultural Food	KB-Sunset 403	KB-SP Higgins	KB-Thomas Basin	KB-Rotary Pool
Location description	Beaches located on the western edge of Revillagigedo Island along the Clarence Strait. Accessible by the Tongass Hwy					
Hydrologic unit code	1901010-20510	1901010-20801	190101020403			
Water Type	Marine					
Area sampled	Beaches where primary contact is common and/or pathogen exceedances frequently occur					
Time of year sampled	July through August					

Virtual Beach Predictions Evaluation

Background

The Beaches Environmental Assessment and Coastal Health (BEACH) Act was passed by the U.S. Congress in 2002 in response to increased occurrences of water-borne illnesses. The U.S. Environmental Protection Agency (EPA) administers grant funds to states, tribes and territories under the Act to establish monitoring and public notification programs. In Alaska the BEACH funds come through the Alaska Department of Environmental Conservation (ADEC) and are typically administered as grants to local communities. The BEACH program has established national marine water quality monitoring and reporting standards for fecal waste contamination and notifies the public when levels exceed state standards. The pathogens of interest are fecal coliform bacteria and enterococci bacteria.

Both Ketchikan and Kenai beaches have an elevated probability of human contact with marine water. The six selected Ketchikan beaches are frequented by dog walkers and recreators, notably Rotary Park Pool which is a popular swimming hole for children as it is warm and shallow. The North and South Kenai Beaches are positioned at the outlet of the Kenai River and are busiest during the dipnet personal use fishery that occurs July 10th through July 31st. Both Ketchikan and Kenai beaches have been monitored for marine pathogens under the Alaska BEACH Program since 2017 and 2010 respectively (ADEC 2022.c, SAWC 2022). Water quality at both beaches has periodically exceeded the criteria for marine pathogens (18 AAC 70.020(14))[2].

During the summer of 2021, historic pathogen and environmental data was compiled and a method for calibrating models was suggested in the 2021 Virtual Beach Report (ADEC 2021). Pathogen monitoring continued in 2022 to collect samples to ground truth the predictive models.

---

[2] Water Quality Standards 18 AAC 70, Alaska Department of Environmental Conservation, amended March 5, 2020

Figure 1. Ketchikan Beaches

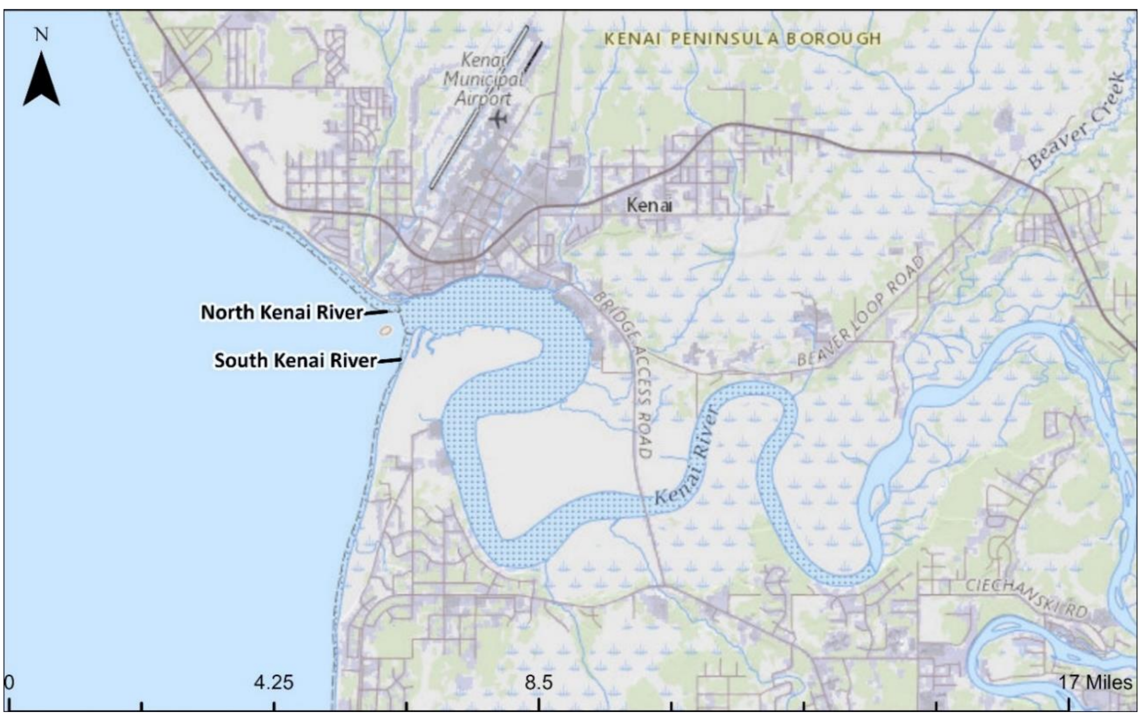

Figure 2. Kenai Beaches

Objective

- Ground truth statistical models for two Kenai beaches and six Ketchikan beaches by predicting the enterococci levels of samples collected during summer 2022.
- Develop standard operating procedures and training materials that can be used by future grantees or stakeholders interested in predicting pathogen levels on Kenai or Ketchikan beaches.

Quality Assurance Review

See the Kenai River Beaches and Ketchikan Beaches 2022 waterbody field reports (DEC) for quality assurance review of pathogen monitoring.

The retrieval of weather and tide level data was automated using R programming version 4.2.1 and remained consistent throughout the season. Both training data[3] and independent variable data was downloaded, processed, and analyzed with R code[4], ensuring that the predictions made over the season were reflective of actual weather changes. Record of code, predictions, and diagnostic plots were kept throughout the season.

Methods

For pathogen sampling protocol see respective Kenai[5] and Ketchikan[6] Beaches Quality Assurance Project Plans (QAPPs). See the 2021 Virtual Beach Report for an outline of methods used in the 2022 Virtual Beach pilot test project.

The following adaptations were made to the 2021 Virtual Beach methods to increase accuracy of predictions or to deal with glitches that arose while running models on certain beaches.

The gradient boosted model (GBM) module would often generate errors when attempting to run models with training data from Sunset Beach in Ketchikan, possibly due to the lower frequency of exceedances in the historic data. The multiple linear regression (MLR) module alone was used on Sunset Beach.

Water temperature was added to the Ketchikan training data after the 2021 Virtual Beach methods were established, but before the 2022 pilot test began. Water temperature data was retrieved using the NOAA CO-OPS Application Programming Interface for Data Retrieval with R programming.

[3] Training data refers to the historical beaches' dataset, enterococci values and corresponding weather data, used to train the models.
[4] Weather and tide data processing code developed in house. Precipitation processing code adapted from Precip_Accumulation_calc.R by Rebecca Bellmore, Southeast Alaska Watershed Coalition
[5] Quality Assurance Project Plan, Kenai River BEACH Program, Alaska Department of Environmental Conservation, February 2022.
[6] Quality Assurance Project Plan, Ketchikan BEACH Program, Alaska Department of Environmental Conservation, April 2022.

Predictions made with GBM closely follow the peaks and troughs of exceedances and non-exceedances on an observed vs. predicted time series, but the predictions are often muted, or not as high as the highest observed values and not as low as the lowest observed values. See Figure 3. To unmute the predictions, a regression line was made in Excel of observations (independent variable) and predictions (dependent variable). The resulting regression equation (y=mx+b) was used to stretch the predictions, to force the high values higher and the low values lower. The intercept (b) was subtracted from the predictions, and that new value was then divided by the regression coefficient (m). See Figure 4 for an example of an unmuted time series plot.

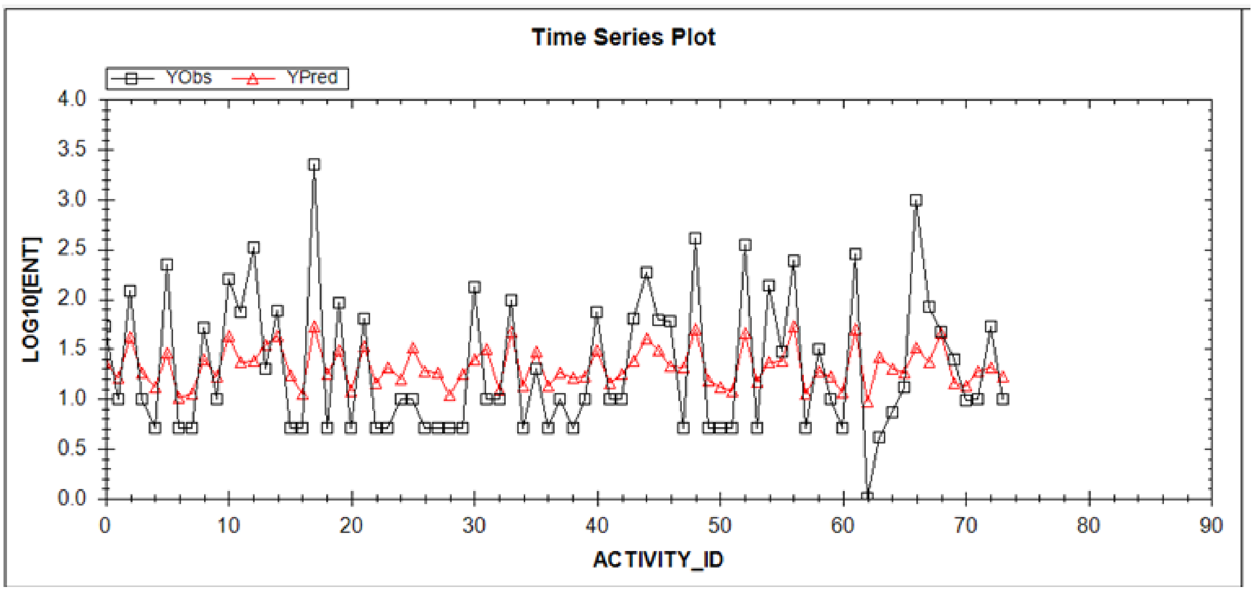

Figure 3. Example of muted time series plot of observations and GBM Virtual Beach predictions. Data displayed is from South Point Higgins Beach, Ketchikan, AK.

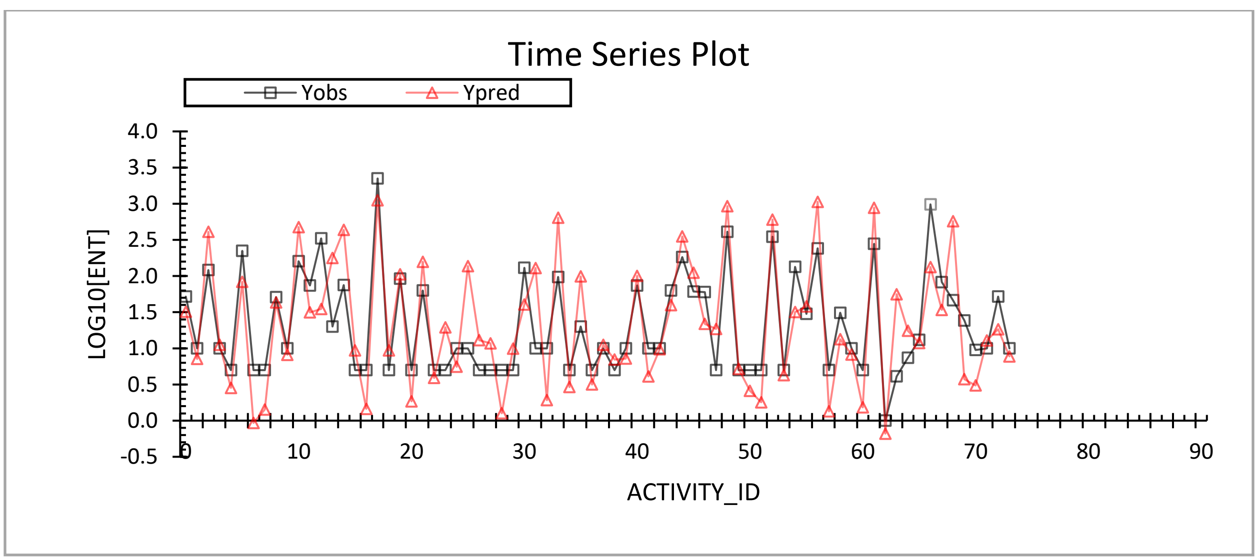

Figure 4. Example of unmuted time series plot of observations and corrected predictions. Data displayed is from South Point Higgins Beach, Ketchikan, AK.

Ketchikan Beaches Results

Over the course of the 2022 monitoring season for all beaches, only one exceedance in a monitoring sample occurred at Rotary Park Pool. GBM provided correct predictions in relation to Water Quality Criteria[7] 100% of the time at Mountain Point Cultural Foods and South Point Higgins Beach, 67% for Herring Cove, 17% for Rotary Park Pool, and 67% for Thomas Basin. Multiple Linear regression provided correct predictions 67% of the time for Mountain Point Cultural Foods, 83% for both Rotary Park Pool and Thomas Basin, and 100% for Herring Cove, South Point Higgins, and Sunset Beach. See Appendix figures 9-19 for graphs of predictions and sample results.

Table 3. Ketchikan beaches percent correct predictions.

Beach	GBM, percent correct	MLR, percent correct	Average, percent correct
Mountain Point Cultural Foods	100%	67%	67%
Herring Cove	67%	100%	100%
Rotary Park Pool	16%	67%	50%
South Point Higgins	100%	100%	100%
Sunset	*	100%	*
Thomas Basin	67%	100%	83%

* GBM not run due to malfunction

Table 4. Ketchikan beaches data summary. Bold values indicate quality criteria exceedances.

Beach	Date	GBM Prediction (MPN)	MLR Prediction (MPN)	Average of GBM and MLR	Actual Sample Result (MPN)
Mountain Point Cultural Foods	6/2	14.2	**343.7**	**179.0**	20
	6/20	35	**721.8**	**378.4**	41
	6/30	24	66	45.0	10
	7/28	13	29.08	21.0	20
	8/18	55.3	67.08	61.2	10
	9/1	15	43.8	29.4	10
Herring Cove	6/2	13.4	38.8	26.1	10
	6/20	13	109.5	61.3	10
	6/30	**203**	29	116.0	20
	7/28	37	45.09	41.0	10
	8/18	35.6	14.7	25.2	128
	9/1	**138**	37.29	87.6	121
Rotary Park Pool	6/2	53.7	18.1	35.9	**197**
	6/20	**699.9**	40.2	**370.1**	30
	6/30	**594**	**428.5**	**511.3**	40

[7] Water Quality Standards (WQS) (18 AAC 70.020(14)) primary contact criteria for marine waters (enterococci > 130 MPN/100ml)

Beach	Date	GBM Prediction (MPN)	MLR Prediction (MPN)	Average of GBM and MLR	Actual Sample Result (MPN)
	7/28	98	35	66.5	10
	8/18	**175.7**	39.3	107.5	20
	9/1	**136**	44.8	90.4	10
South Point Higgins	6/2	11.9	25.7	18.8	10
	6/20	4.4	20	12.2	52
	6/30	8	20.8	14.4	10
	7/28	10	26.8	18.4	10
	8/18	4.7	11.3	8.0	20
	9/1	25.6	17.8	21.7	20
Sunset	6/2	*	13.3	*	10
	6/20	*	13	*	10
	6/30	*	14	*	10
	7/28	*	13	*	10
	8/18	*	9.9	*	10
	9/1	*	13	*	10
Thomas Basin	6/2	10.1	12	11.1	10
	6/20	**210.9**	42	126.5	118
	6/30	132	73	102.5	10
	7/28	6	22	14.0	109
	8/18	33	75.18	54.1	31
	9/1	**885**	25	**455.0**	10

* GBM not run due to malfunction

Kenai Beaches Results

The collected samples showed water quality criteria exceedances once at North Kenai Beach and once at South Kenai Beach during the 2022 monitoring season. GBM correctly predicted pathogen levels in relation to the water quality criteria 75% of the time for South Kenai Beach, and 50% of the time for North Kenai Beach. MLR correctly predicted pathogen levels 75% of the time for both Kenai Beaches. See appendix for graphs of results. When taking the average of the results from both model types, predictions were correct 75% of the time at both Kenai Beaches. See Appendix figures 5 and 6 for graphs of predictions and sample results.

Table 5. Kenai beaches percent correct predictions

Beach	GBM, percent correct	MLR, percent correct	Average, percent correct
South Kenai Beach	75%	75%	75%
North Kenai Beach	50%	75%	75%

Table 6. Kenai beaches data summary. Bold values indicate water quality exceedances.

Beach	Date	GBM (MPN)	MLR (MPN)	Average of GBM and MLR	Sample result (MPN)
South Kenai Beach	6/30	39	26.4	32.7	62
	7/12	**183.5**	**140.6**	**162.05**	**435**
	7/27	34	116	75	39
	8/9	**499**	**157.6**	**328.3**	117
North Kenai Beach	6/30	52	26.4	39.2	25
	7/12	**191**	24.6	107.8	66
	7/27	2	16.2	9.1	22
	8/9	8	19.9	14	**248**

Conclusion

Pathogen values predicted with Virtual Beach should not be used to determine impairment of a waterbody and should only be used to notify the public of likely exceedance of water quality criteria.

Virtual Beach correctly predicted one (at South Kenai Beach) of the three observed exceedances of water quality criteria that occurred in the routine samples collected at all beaches over the 2022 monitoring period. In total, the GBM predicted 11 false exceedances and MLR predicted four false exceedances in 2022. The frequency of predicted exceedances by Virtual Beach was higher than what was observed. Therefore, Virtual Beach may not be appropriate for determining waterbody impairments. However, Virtual Beach may be a valuable tool to protect human health due to the model's tendency to skew towards predicting exceedances.

Historic beach data influenced the behavior and fit of the models produced by Virtual Beach. Predictions at Ketchikan beaches tended to be of higher magnitude than what was observed, with the maximum value being 885 MPN at Thomas Basin. The beaches where more false positives occurred were beaches that historically have had a higher number of exceedances. These beaches were more likely to predict an exceedance than the beaches with less exceedances in their training data under the same environmental conditions. Beaches that historically have had a lower frequency of exceedances, such as Sunset Beach, tended to predict low bacteria values with each model run. The training data contained so few exceedances, that this may have been the cause for malfunction when attempting to run the GBM models for Sunset Beach. MLR models alone were run for Sunset Beach and did well at predicting the non-exceedances, but the lack of exceedances in samples collected at Sunset beach may be cause for excluding it from future prediction projects.

Differences in geographical characteristics or proximity to nonpoint pollution sources could influence the frequency of exceedances at some beaches. One example is Rotary Park Pool, where sandbars block the flow of water creating warmer and more stagnant conditions. Since the ocean water temperature data used to train the models comes from a NOAA station a few miles away, it may not capture the warmer, stagnant conditions at Rotary Park Pool. The lack of accuracy (16%) of Rotary Park Pool bacteria predictions may be justification to exclude the beach from further analysis with Virtual Beach unless point specific water temperature data can be sourced for this beach.

For the 2022 Virtual Beach Pilot project, only GBM predictions were posted online along with the results from sample collection. The predictions were published as either “Exceedance”, meaning values above the 130 MPN water quality criteria threshold, or “Okay”, meaning values below the 130 MPN threshold. If the models are used to predict marine pathogen exceedances in the future, both GBM and MLR predictions could be published as two separate values or averaged. The decision threshold could also be adjusted, and three descriptions or colors could be assigned to low, medium, and high predictions. For example, a green color could be published when the models predict values from 0-130. A yellow color could be published when values from 131-260 are predicted and red when values of 261+ are predicted.

Virtual Beach, using either gradient boosted models, multiple linear regression, or an averaging of both, may be a viable substitute for pathogen monitoring at all beaches studied except Rotary Park Pool in Ketchikan. Ultimately, maintaining a public outreach campaign to remind beach users to wash after recreating in marine water, regardless of the most recent sample results or predictions, will help prevent illness from accidental ingestion of marine pathogens.

References

ADEC (Alaska Department of Environmental Conservation). 2020. Water Quality Standards: 18 AAC 70.

ADEC (Alaska Department of Environmental Conservation). 2021. 2021 Virtual Beach Report, Kenai, Alaska. Prepared by J. Petitt. Division of Water, Soldotna, AK.

ADEC (Alaska Department of Environmental Conservation). 2022.a. Kenai BEACH Water Quality Monitoring and Bacteria pathogen Detection: Quality Assurance Project Plan, V. 6.

ADEC (Alaska Department of Environmental Conservation). 2022.b. Quality Assurance Project Plan, Ketchikan BEACH Program.

ADEC (Alaska Department of Environmental Conservation). 2022.c. Waterbody Field Report, Kenai River North and South Beaches, Kenai, Alaska. Prepared by S. Apsens and J. Petitt. Division of Water, Soldotna, AK.

Bellmore, R. Southeast Alaska Watershed Coalition. 2021. Precip_Accumulation_calc.R. R Programming Script.

SAWC (Southeast Alaska Watershed Coalition). 2022. 2017-2021 Ketchikan Beach Monitoring Comprehensive Report.

Appendix

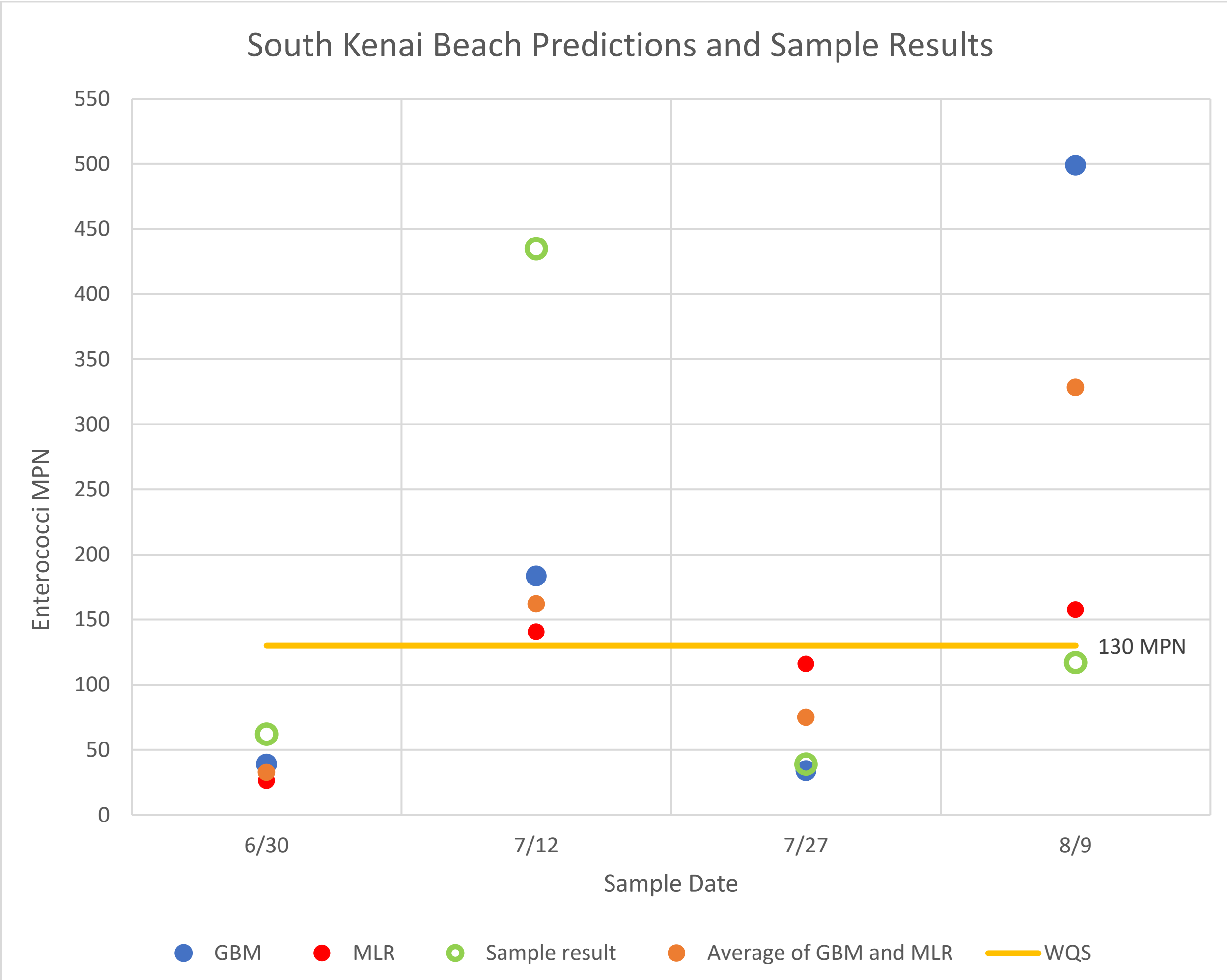

Figure 5. South Kenai Beach GBM and MLR Predictions and Sample results. Water Quality Standards (WQS) (18 AAC 70.020(14)) primary contact criteria for marine waters (enterococci > 130 MPN/100ml)

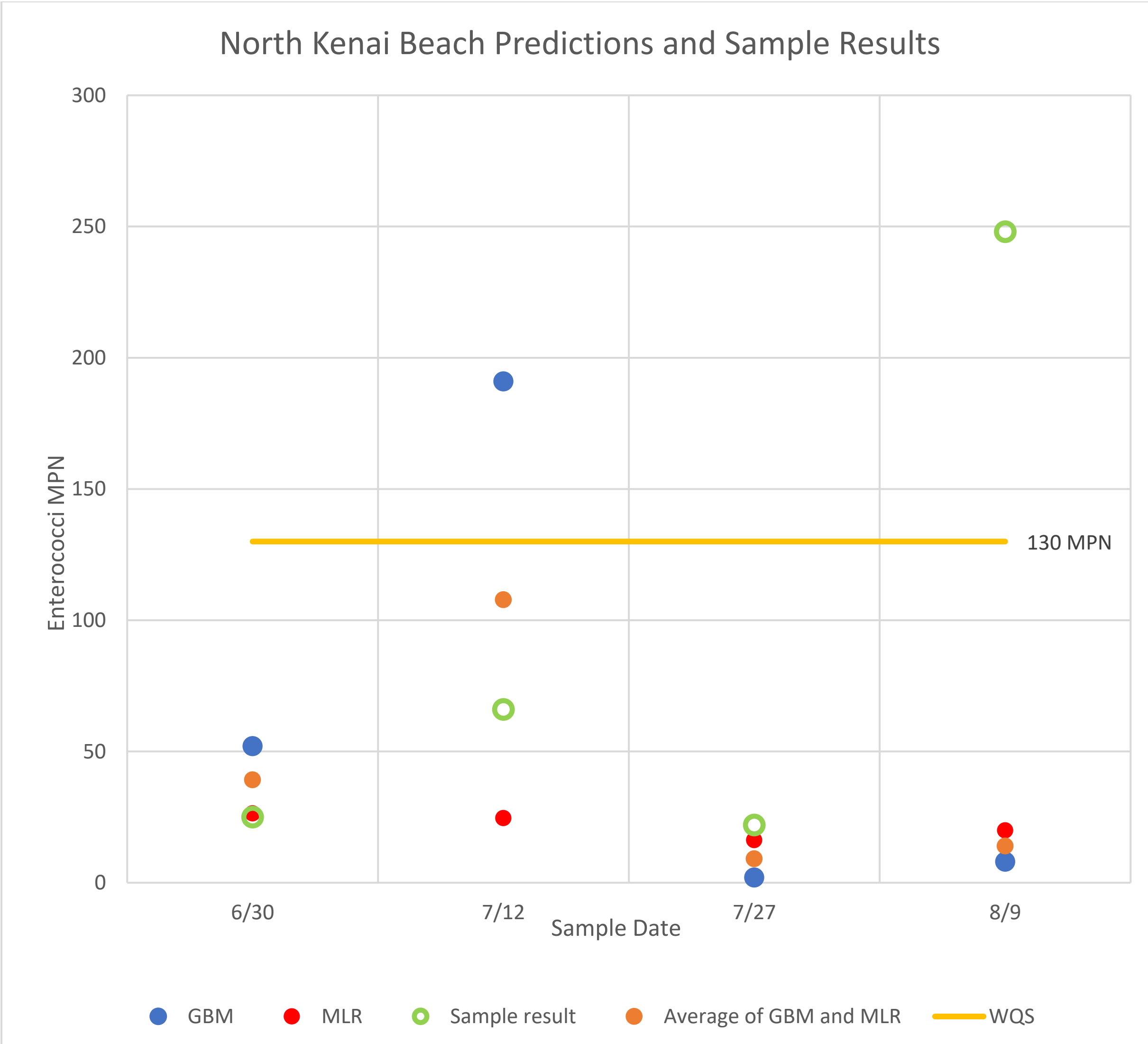

Figure 6. North Kenai Beach GBM and MLR Predictions and Sample Results. Water Quality Standards (WQS) (18 AAC 70.020(14)) primary contact criteria for marine waters (enterococci > 130 MPN/100ml)

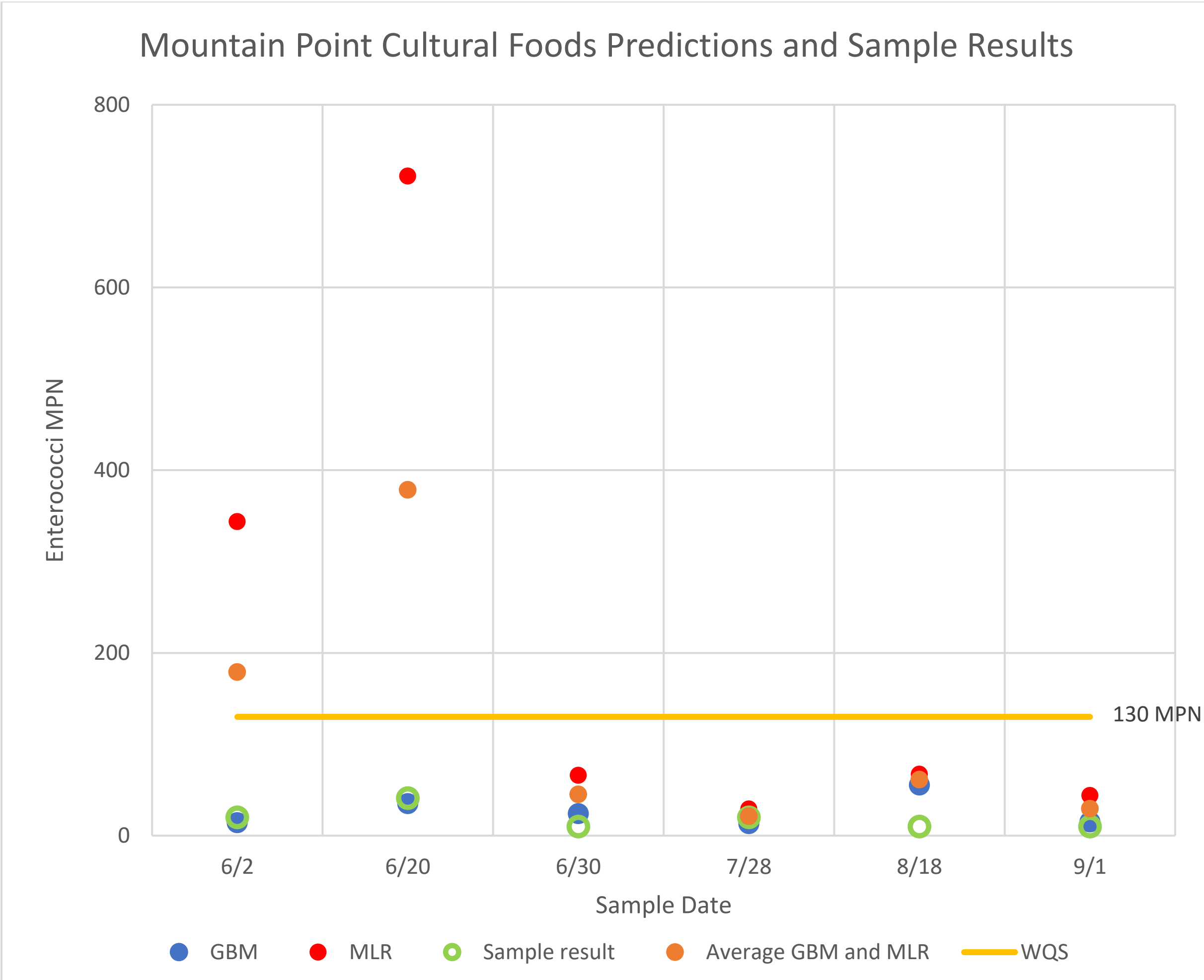

Figure 8. Mountain Point Cultural Foods GBM and MLR Predictions and Sample results. Water Quality Standards (WQS) (18 AAC 70.020(14)) primary contact criteria for marine waters (enterococci > 130 MPN/100ml)

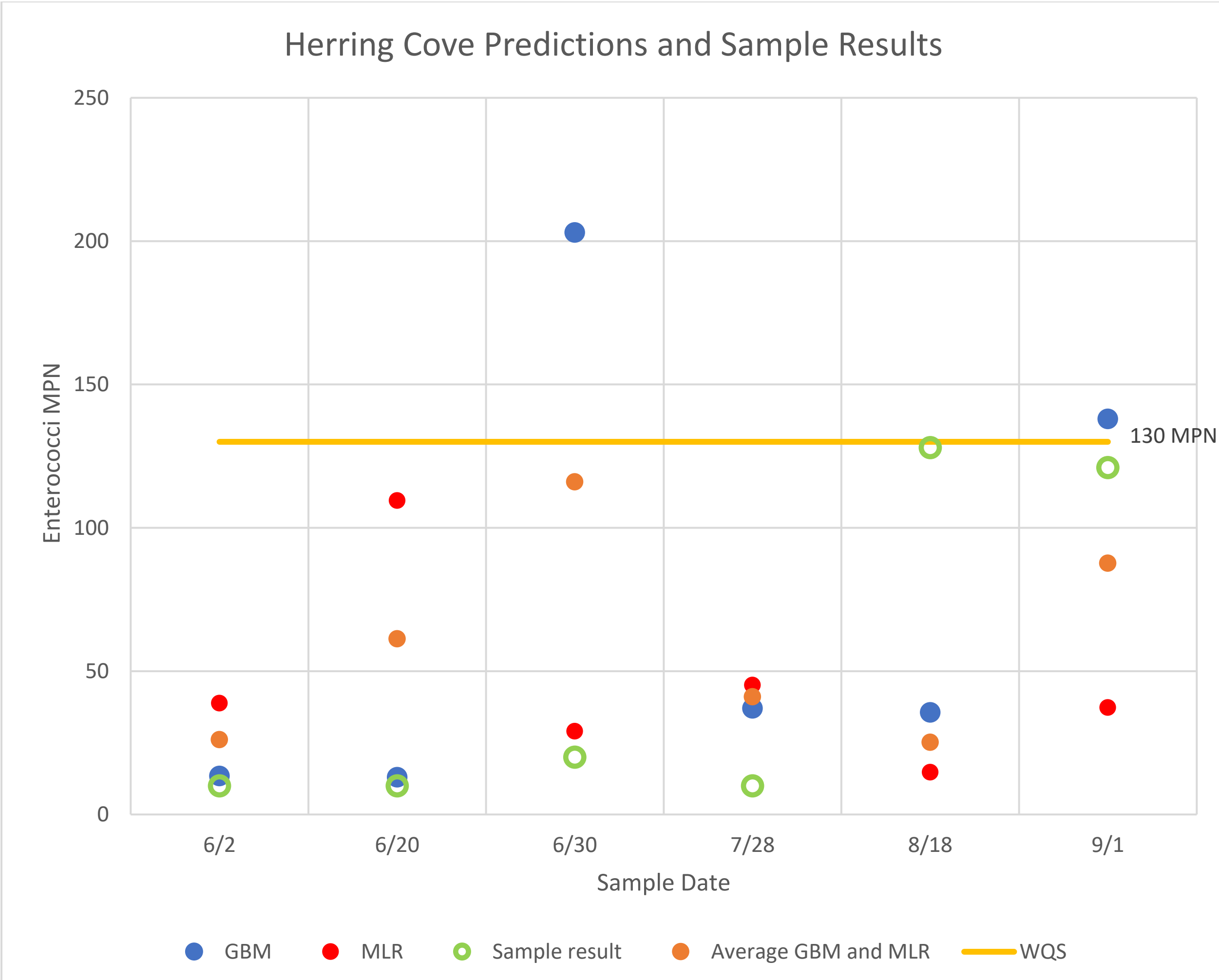

Figure 9. Herring Cove GBM and MLR Predictions and Sample results. Water Quality Standards (WQS) (18 AAC 70.020(14)) primary contact criteria for marine waters (enterococci > 130 MPN/100ml)

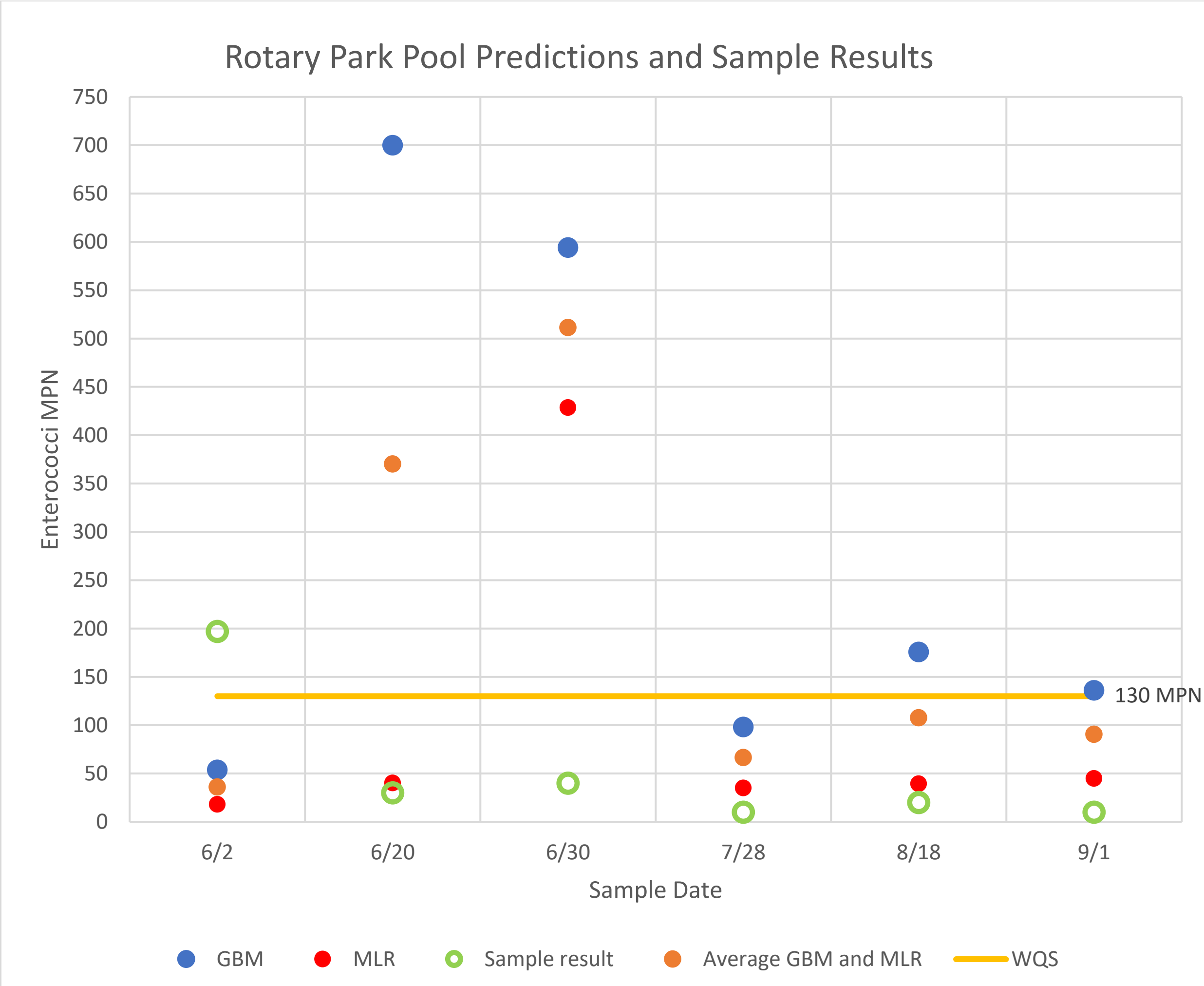

Figure 10. Rotary Park Pool GBM and MLR Predictions and Sample results. Water Quality Standards (WQS) (18 AAC 70.020(14)) primary contact criteria for marine waters (enterococci > 130 MPN/100ml)

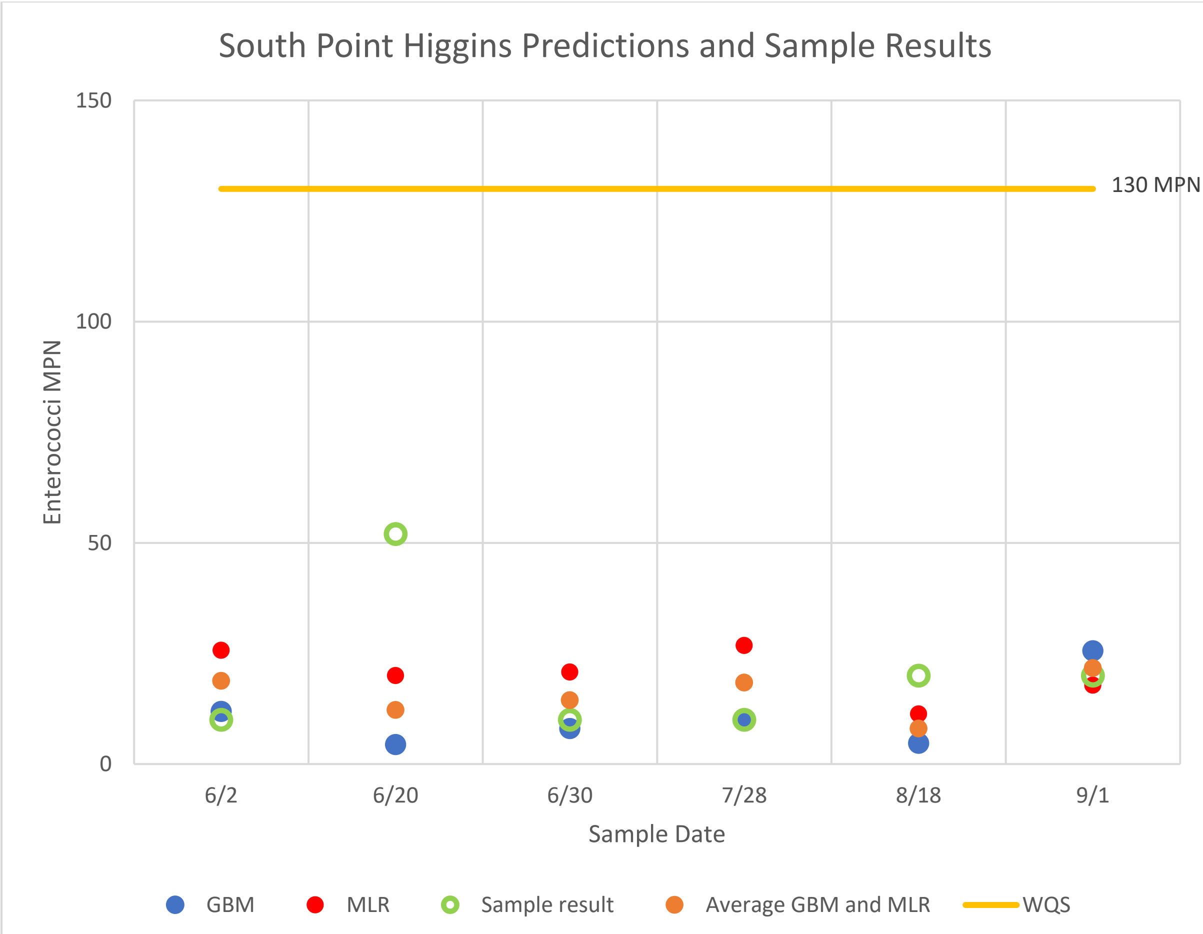

Figure 11. South Point Higgins GBM and MLR Predictions and Sample results. Water Quality Standards (WQS) (18 AAC 70.020(14)) primary contact criteria for marine waters (enterococci > 130 MPN/100ml)

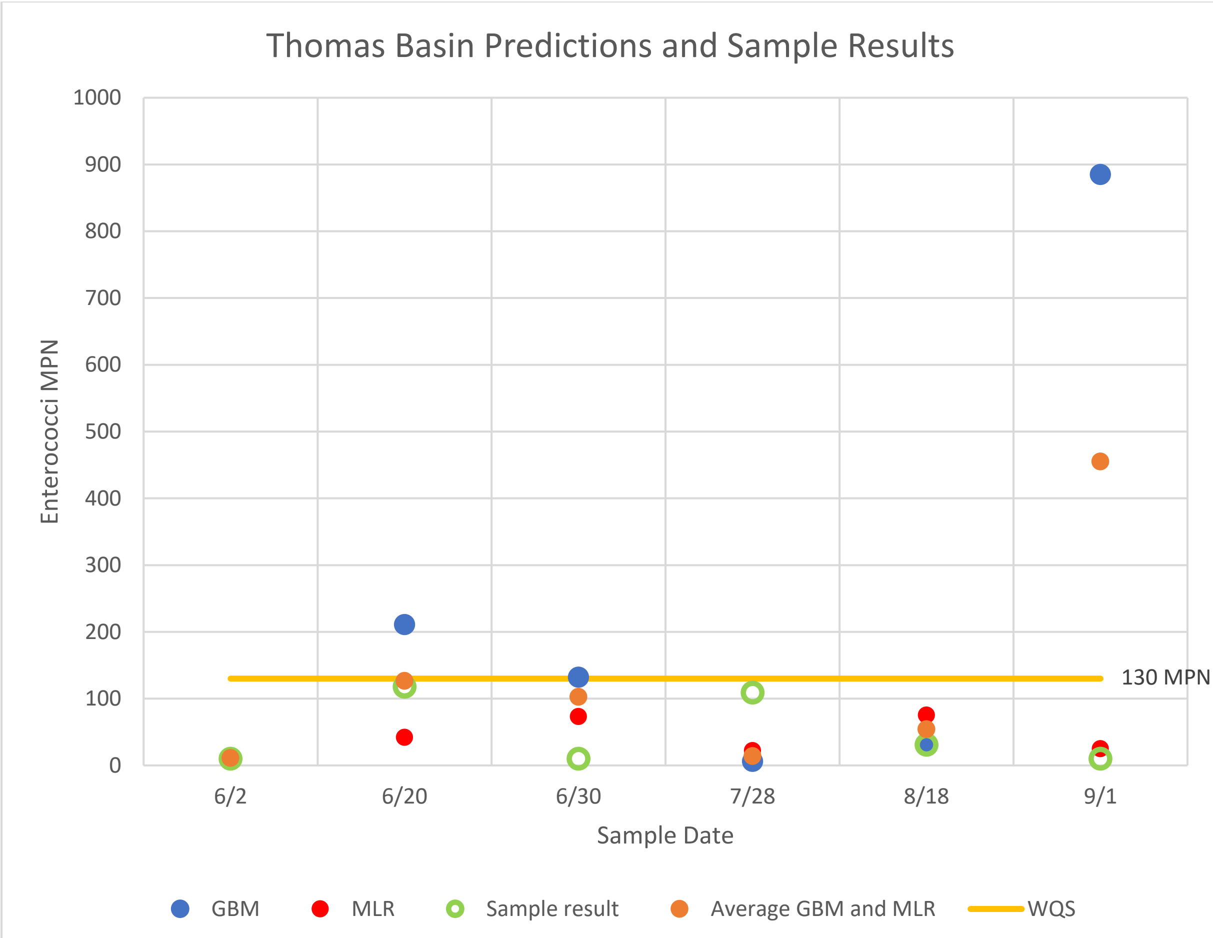

Figure 12. Thomas Basin GBM and MLR Predictions and Sample results. Water Quality Standards (WQS) (18 AAC 70.020(14)) primary contact criteria for marine waters (enterococci > 130 MPN/100ml)

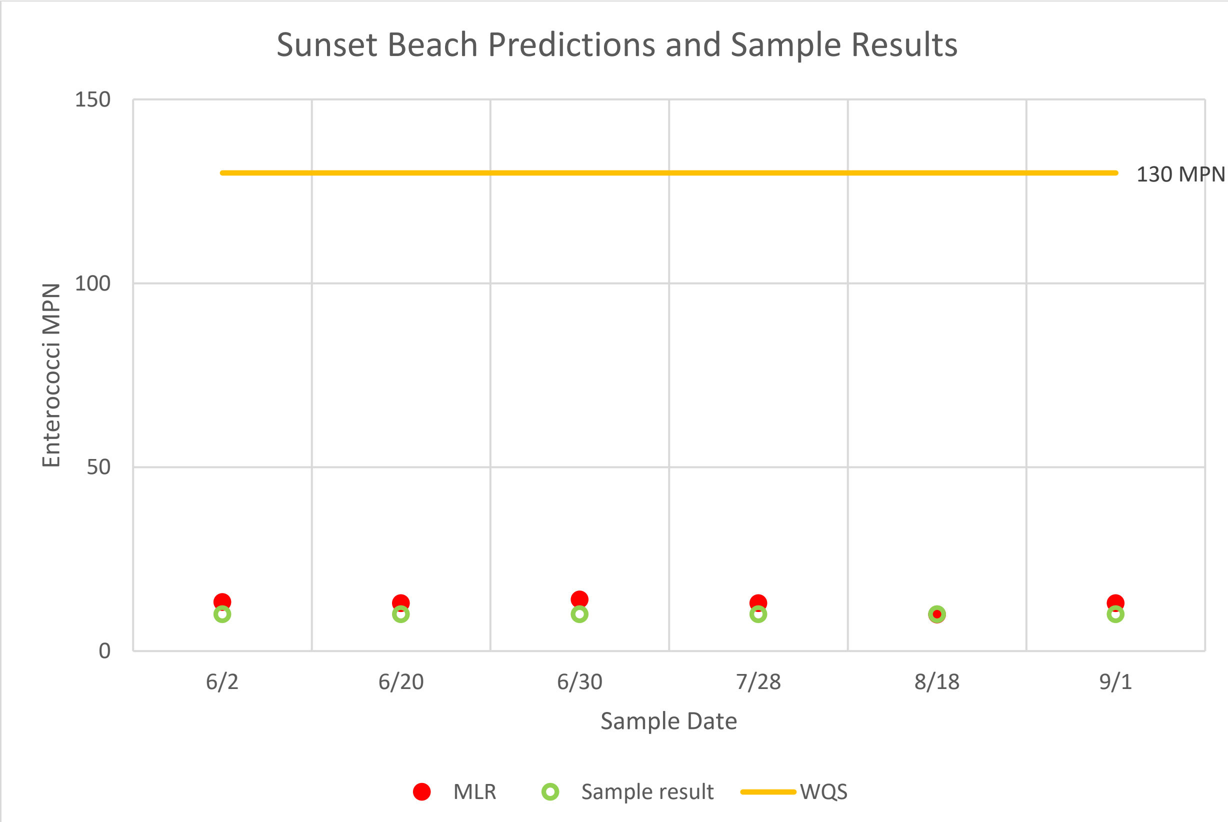

Figure 13. Sunset Beach GBM and MLR Predictions and Sample results. Water Quality Standards (WQS) (18 AAC 70.020(14)) primary contact criteria for marine waters (enterococci > 130 MPN/100ml)