Designing an Automated Water Quality Monitoring System for West and Rhode Rivers

Alex Anvari, Jenny Delos Reyes, Ehsan Esmaeilzadeh, Ali Jarvandi, Nicholas Langley, Keyssi Rivera Navia, *Member, IEEE*

***Abstract*— The Chesapeake Bay is the largest estuary in the United States, where its watershed is home to more than 3,600 species of plants and animals and more than 16.6 million people. However one of its major issues is water pollution. Good water quality is vital for the health of all these plants, animals and people. In order to act upon this problem and restore the water, there is the need to monitor the water quality. There are currently several organizations and agencies monitoring different parts of the bay and working to restore the bay. [1]This project analyzes innovative ways to improve water quality monitoring in the West and Rhode rivers. The water in these rivers is currently monitored by a River Keeper. There is the need for an improved system with higher frequency of data input, higher accuracy of higher quality sensors, and a wider range of parameters being monitored.**

The motivation behind this project is to develop a transfer function between water quality and source of pollution. An improved model will allow the river keeper to have a better understanding of the conditions of the water and track the sources of pollution. With this new system, he will be able to act upon this acquired data and help to restore these rivers and subsequently the Chesapeake Bay.

This design evaluated various sensor alternatives, transmission technologies, and used GIS mapping software in order to implement an automated water monitoring system for the West and Rhode rivers. A notional utility curve between available sensors and transmission techniques was developed where preliminary results indicate a system will fit the river keepers' needs and desired goals

I. Introduction

The current environmental condition of the Bay is now only slightly better than the worst it has ever been. Each year more and more pollutants find their way into the bay. Growing development along the water front replaces vegetation vital to the ecosystem and increases of the amount of runoff and pollutants are collected in the water. Over the decades one of the most prevalent businesses on the water was the farming/collection of oysters. This business, along with the populace of oysters has declined and is now in danger. With the overharvesting and decline of oysters from disease the ability of nature to filter and clean the waters of the bay has drastically reduced. Combined with the ever growing population and development in the area, the waters of the bay are unable to deal with the increased pollution and effects of storm water runoff. Without filter feeders to limit excess nutrients entering the water, algae has grown almost unhindered. With excess nutrients such as nitrogen and phosphorous present in the water due to the run off of fertilizer from farms and sewage treatment plants, algae have been allowed to grow in excess. The blooms of algae obscure water clarity and block sunlight from submerged aquatic vegetation such as under water grasses. The blooms will also deplete dissolved oxygen causing massive fish kills in the rivers of the Chesapeake. [14]

A. *Problem Statement and Motivation*

Environmental regulations require the monitoring of the environmental state of the West and Rhode River in order to preserve or improve its water quality.[4] The current system in place for the West and Rhode River water basin requires travel to 32 locations on the West and Rhode Rivers. At each location a manual sensor sample for dissolved oxygen, conductivity, and bacteria are recorded by hand. And a rough estimate of turbidity is recorded using a Secchi disk. Testing occurs once a week during the months from May through October, where a set of volunteers would survey sampling sites on the West River and another set of volunteers would sample on the Rhode River. The process of travelling and collecting data about takes two hours to complete. [5] Upon completion, the manually recorded data is then given to a webmaster for input into a web server. The water collected for bacteria sampling is sent to a lab at a nearby community college. There is an apparent time delay between each of these weekly cycles, and there is an apparent chance that data may be incorrectly recorded or lost. Because of these possibilities the River-Keeper has very little time to investigate and act on poor sources of water quality, nor has he have the ability to accurately gauge the West and Rhode Rivers' overall current state. There are also a limited number of parameters tested in the current system. Important parameters such as pH, salinity, phosphorous, and nitrogen, which contribute to water quality, are not tested.[5]

B. *Scope*

This study was restricted to the Chesapeake Bay and specifically the West and Rhode Rivers water quality improvement. Within this context the scope of the project began with understanding the water quality monitoring systems currently used. Plans were then made to improve the monitoring of water quality within the mentioned context, where the scope of the system was to investigate potential use of a new sensor or sensors; fixed sensors and moving sensors. The sensors overall benefit to the quality of monitoring water was tested against the current system as well as other sensor alternatives. The scope of the system included investigation of the implementation of technology that will allow the collected data to be transmitted wirelessly to a central location. A utility

This work is sponsored by the West/Rhode River-keeper, Chris Trumbauer and Sivix Corporation

The faculty advisor is Dr. George Donohue, from the Department of Systems Engineering and Operations Research, George Mason University, Fairfax, VA 22030 USA (e-mail: gdonohue@gmu.edu)

function has also been developed for each sampling site, to give recommendations as optimum places to sample or to locate various fixed sensors. Historical data collected at each sampling point and spatial temporal data from a GIS map will be used as scored inputs for the utility function.

II. Valuing System Alternatives

After determining the scope, the project team developed a plan that would quantify the goals of the system and its objectives in a measured utility function. Measurements of utility and lists of requirements were then made so that system alternatives were tested against each other. Recommendation based on potential utility feedback from water quality monitoring experts, as well as careful analysis has been derived.

A. Value Hierarchy

Below are the utility values derived from the input of the West and Rhode River-Keeper on what a water quality monitoring system should perform. Criteria for each portion of the utility was derived by several interviews with the West and Rhode River-Keeper. Total utility is broken down into operational quality, maintainability, and reliability.

$$\textbf{U(Total): } (0.4)*W_O + (0.2)*W_M + (0.4)*W_R$$

$W_O = (0.3)F + (0.7)P$

F = Frequency of the times to test the water measured in times/day

P = Number of parameters to test

$W_M = (0.8)C + (0.2)B$

C = Time between calibrations measured in weeks

B = Battery lifetime measured in months, per hour s ampling time

$W_R = (0.4)T + (0.4)D + (0.2)S$

T = Mean Time Between Failure measured in years

D = Profundity measured in meters

S = Weather survivability measured in number of seasons usable

B. Derived Requirements and Corresponding Goals

Goal: The system shall provide information about the defined parameters of water quality in pre-determined cycles.

1. Parameters: water temperature, salinity, dissolved oxygen, turbidity, pH, nitrogen and phosphorus levels
2. Extend coverage and gathering frequency of data from once a week to every 15 minutes
3. Provide automated data analysis
4. Provide figures of merit suitable for use in a new GIS water quality analysis system under concurrent development

Goal: The system shall limit the number of personnel required at regular testing sessions.

1. Ability to define testing frequency and conditional testing times
2. Present data on a web portal provided by the River-Keeper
3. Sensor to have an interface to receive and transfer data to user
4. Provide compatible data to the GIS mapping system under concurrent development for the West/Rhode Rivers

III. Design Alternatives

A. Sensor Alternatives

In order to improve the water quality monitoring system of the West and Rhode rivers, the team first investigated different sensor alternatives. The team gathered five different sensor alternatives which were representative of the most popular sensors in the water quality monitoring market. Then analysis of the various sensors was conducted to understand the utility that an individual sensor could provide to the system as a whole.

YSI 6600 V2

The YSI 6600 V2 is a multi-parameter sonde and measures up to 16 parameters in severe fouling environments for extended periods. The sensor can measure the desired nine parameters mentioned in the requirements list: conductivity, temperature, specific conductance, salinity, pressure, depth, pH, dissolved oxygen Concentration (mg/L) and saturation, turbidity (NTU), chlorophyll fluorescence (% full scale), chlorophyll concentration (µg/L), blue-green algae fluorescence (% full scale), and blue-green algae Concentration (cells/ml). The sonde has two optical ports, conductivity/temperature port, or rapid pulse DO port, and a pH port. The calibration for most of the parameters is due once a month. YSI 6600 also features anti-fouling kit, which saves time and reduces operating costs by extending the deployment period of sondes. These kits include sonde guards, probe wipers, hardware fabricated from copper alloys as well as copper tape for the probe bodies, and longer and thicker wiper pads. [7] [8]

Troll 9500 Professional XP

The Troll 9500 multi-parameter sonde is suitable for use on surface, ground water interaction, watershed and source water protection, and aquaculture. This sonde has IP67 waterproof battery compartment, USB or RS232 connections with capability to accept SDI-12, and it can be easily calibrated in less than five minutes. The Professional XP is the newest version of this sonde with the ability of measuring conductivity, pH, dissolved oxygen, salinity, temperature, depth, turbidity, ammonium, and the nitrate. Calibration is required once every three months for most of the sondes, and the battery type is 2D-size lithium with estimated life time of 11 months assuming 15-minute sampling interval. A data logger needs to be purchased in order to read and record the data. [7] [8] [13]

YSI 6820

6829 V2 is a sonde for profiling and spot-sampling in rivers, lakes, wetlands, wells, estuaries, and coastal waters. This version of 6820 has 2 optical ports: conductivity/temperature port, pH port, and one choice of nutrients (Nitrate, Phosphorus.) Available optical sensors include ROX optical dissolved oxygen, blue-green algae, chlorophyll, turbidity, or rhodamine. 6820 features RS-232 and SDI-12 interface communications and easily interfaces with any data collection platforms for long-term deployment. YSI 6820 V2 can be combined with anti-fouling kit to offer longer deployment time, and has a built-in data logger. [8]

600 XL

YSI 600 XL is a handheld sampling device that will allow a maximum of five desired parameters to be measured simultaneously in real time. It is compact, submersible to 61 meters (200 ft) and employs YSI's sensor reliability and parameter measurement systems. The system can quickly sample fresh, salty, and polluted waters. 600 XL does not feature a built-in data logger and therefore has to be combined with YSI 650 MDS display-data logger to record and transfer the data. The battery life is approximately 30 hours for spot sampling depending on the number of stations. It measures conductivity, temperature, pH, and dissolved oxygen. The unit can go reach depths of 61 meters, but it is limited by cable length. The maximum sampling frequency is once every 4 seconds, and calibration is the same as that of all other YSI alternatives. [8]

YSI 85

This instrument enables simultaneous measurement of dissolved oxygen, conductivity, salinity, and temperature. It has a battery life of 30 hours, and is similar to the 600 XL in that can retrieve one sample of the mentioned parameters every four seconds. This is the current sensor being used by the River-Keeper. [8]

B. *Transmission Alternatives*

After first documenting the capabilities of the various sensor alternatives, the team also obtained data on various transmission technologies and their capabilities of gathering data in a timely manner. The team evaluated the transmission alternatives also paying particular attention to its capability of receiving and sending data over various environments.

Short-Range Transmission vs. Long-Range Transmission

Radio frequency transmission is best used for short range transmission. The capabilities of radio frequency transmission can satisfy the system's overall objective, and includes features that would allow both fixed sensors and moving sensors to gather data from the water and transmit data to a nearby device that could display or collect the information. Long-range transmission is associated with cellular telemetry and also provides the capability of the short range radio frequency transmission, but it has added features of GSM (Global System for Mobile Communications) or GPRS (General Package Radio Service), which allows for faster retrieval of data over various mediums and various distances.[6]

IV. Analyses of System Alternatives

A. *Sensors*

The value hierarchy and its corresponding utility function equation were used to test the different sensors. The corresponding criteria under the three major parts of the value hierarchy were then given a particular utility curve that reflected the variability of the utility as a particular area of the criteria was changed. The team determined that the Mean Time Between Failure, Sensor Battery Lifetime, and Time Between Calibration of Sensors has a linear relationship. As time between failure increased, as battery lifetime increased, and as time between calibrations increased, the sensor would experience a gain in utility in a linear fashion. Profundity, Number of Parameters Measured, and Sampling Frequency behaves in an exponential fashion. After having determined the sub-utility behavior on a utility curve, the team then placed the actual values of each of the components into the function. The team retrieved various results for each of the different sub-utility criteria. The results are shown below:

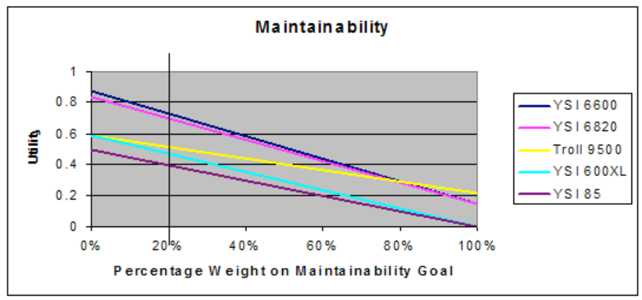

Figure 1: Maintainability

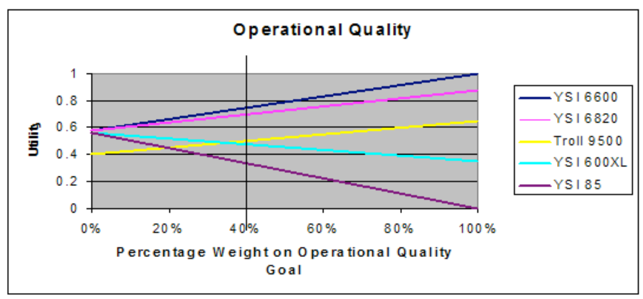

Figure 2: Operational Quality

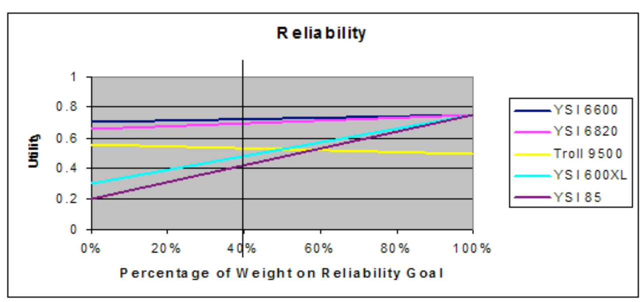

Figure 3: Reliability

On the x axis the percent of weight for that particular portion of the value function is shown, on the *y* axis is the utility score. The vertical line shown within the graph shows the percent of weight actually given to that portion of the utility function. With each of the different components of overall utility, the YSI-6600 scores better than the other sensors. The YSI-6600 is similar to the Troll 9500 as well as the YSI 6820, and therefore a detailed sensitivity analysis was conducted to see how varying weights affect the outcome of the utility of the sensor.

A cost analysis was done on the total system costs relating to sensor choice after utilities were found. This allowed a representation of value by providing a cost per utility. Costs were found using Net Present Value(NPV) for periods of one to five years. And based on these total system life-cycle costs, a cost per utility per year was found. The YSI 6820 came out at the best value for its associated utility. And with a subjective look, the 6820 also falls into key grant-funding brackets that are available to the River-Keeper, thus the YSI emerged as a final selection.

B. *Transmission Devices*

As both the short-range and long-range transmission alternatives meet our overall objective, the alternatives were analyzed based on cost with the added utility of cellular technology. The cost difference between radio frequency and GSM/GPRS technology was negligible in respect to the added functionality of long-range transmission, and thus the GSM/GPRS technology is the preferred system alternative. The transmission device is being provided by Sivix Corporation.

V. INTERFACES

Due to the multiple devices used in this system, interface is a vital part of overall capability. Poorly matched interface systems will limit functionality and hamper performance.

In order to allow proper data transfer, the sensor will be using an RS-232 serial cable that will be adapted to USB in order to properly connect to the separate transmission device. The transmission device will send out the data periodically; access to this data will be provided on a web interface. The web interface will also provide analytical outputs, including graphical representations.

A. RS-232. A standard for serial binary data signals, RS-232 is a common standard for data transmission from and to external devices, including the water quality industry. The cable can be adapted to USB through a serial-to-USB adapter used for computer external devices, allowing for a USB device to serve as a communication medium between the Sensor and the User.

B.*Sensor Output Transmission*. The Sivix transmission device uses GSM/GPRS cellular based data transfer abilities. The specially modified device will have a computing processor that uses telnet based coding in order to control the telnet firmware on the YSI 6820 Sensor. The telnet output of the YSI 6820 will be recorded into a Comma Delimited Value(CSV) file which will be transmitted over the GSM Network to a central server. GSM provides the generic cellular service, including additional features of SMS Text functions. GPRS provides the infrastructure for packet based internet connections, thus allowing the uploading of CSV files. [6]

C. *Fixed/"Moving" Sensor*. The model of a "moving" sensor system provides the ability for one single sensor-transmission system, being transported by boat, to provide the improved sampling to all sites. Added utility occurs with the placement of a fixed sensor at a specific sampling location. This allows for water sampling to take place without the user being on the water to accompany the sensor. The transmission device allows for data collection and external control. However with the sensor being fixed, without the additional cost of a secondary system, the user would need to remove the fixed sampling setup in order to use the system for other sampling locations. However, this allows for the user to begin with a "moving" sensor operational scenario until the testing session is complete, then at the end of the process place the sensor into a fixed-location mode, allowing for data collection during off-testing hours.

VI. SAMPLING SITE OPTIMIZATION

After having found the best alternatives for both sensors, and transmission devices the team investigated actual ways in which to determine the value of each sampling site. Statistical analysis was completed to find correlation between different sampling sites, where recommendations were made to see if certain sampling sites could be eliminated. GIS software was used to input historical spatial temporal data, wherein the team was able to interpolate and visualize areas of poor water quality, areas near run-off, as well as areas that were highly volatile over time. These values were then used as parts of the utility for each of the revised sampling sites, where the most volatile, the most poor of water quality, and the proximity to areas near unbounded run off were given the highest value. The team decided that this was where the fixed sensors would be located.

A. *Statistical Analysis of Correlation*

Based on the overall testing goals, the most valuable data came from the points that either provided a high correlation with the mean value of all stations, or points that underwent sudden changes in one or more parameters. As part of our effort to reduce the number of testing sites required, we developed efficient combinations of testing stations that provided either a high correlation with the mean or serve as a tracking station with high volatility. A parameter was defined as the average deviation from the mean for each station and the lowest resulting numbers identified the most representative points. Volatility is defined by the standard deviation of data samples from the different sampling locations, the following graph shows the standard deviation for the sampling sites in question, certain sites being clear indicators of high volatility.

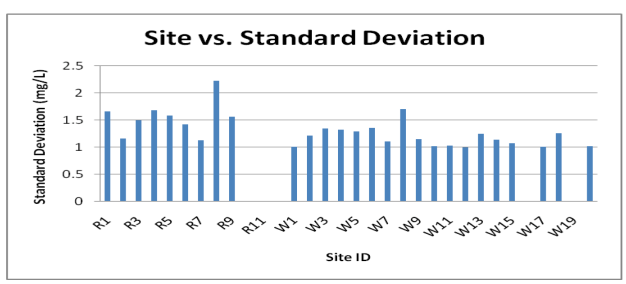

Figure 4: Standard Deviation

Also, the average volatility of each site was defined and the ones with the highest value were chosen as "critical" points, or locations that could provide timestamps to sources of poor water quality. Since most parameters followed a linear path during the year and due to the lack of data on other parameters, dissolved oxygen remains the main parameter to define the utility of testing sites.

A two sample t-test was conducted for sampling sites within 50 yards of each other. The data input used to measure the correlation of the different sites were the different parameters and its historical data measurements in the year 2008. The list below shows the sampling sites that are in close proximity to each other, and the p-value associated with the possible correlation. P-values of 0.1 or higher are instances where we reject the null hypothesis, meaning the two sampling sites are not correlated. Thus, sampling both points are necessary for sampling. P-values that are less than 0.1 are sampling sites that show correlation. Thus, one of the points could suffice for the other.

Ho: $\mu_1=\mu_2$

Ha: $\mu_1\neq\mu_2$

$\alpha = 0.1$ was used to reject the Ho

Site ID	Site ID	P-Value
R1	R2	0.07
R5	R6	0.7
W8	W9	0.0285
W8	W15	0.3914
W12	W13	0.7071
W12	W4	0.2049
W1	W8	0.00246
W1	W15	0.000001
W2	W3	0.8
R2	R7	0.0000001
W5	W20	0.07
W6	W17	0.735

Figure 5: Correlation Table

B. Utility Function for Sampling Site

In the current system, the method that determined the sites for sampling was its vicinity to swimming areas. The team developed a scaling function to redefine the best places for sampling the water. This new system being developed is to find the overall water quality of the rivers, and to also help restore its water quality. The team decided that two different sampling site categories could help satisfy these objectives. The team formed an overall water quality utility function, where the sampling sites would better monitor the general water quality of the rivers. The team also formed a Safety utility, to ensure that swimming areas and densely populated areas would be monitored closely for bacteria.

Overall Water Quality Utility. The input into determining the weights of this utility function was determined by the input of the West and Rhode River-Keeper. This utility function was based mainly on various geographic points, where the locations near the river's mouth, and locations near tributaries and creeks were the factors considered and weighed equally important. These places were scored high as places to sample.

The water quality sampling sites underwent minor revisions. Unlike sample points for bacteria which are based upon common swimming areas, these points were based upon creeks, tributaries, and areas where the flow of the river can be assessed. Because these sampling points were based on separate criteria, we chose to edit them in order that these points could better fit the functions they serve. In the West River it was found that many of the water quality sampling sites were well placed. Sampling locations that were placed in the middle of creeks that drained into larger parts of the river were kept such as those at John's Creek (W13), Smith Creek (W12), and other small creeks. In instances where two points that are located in close proximity (~50 yds or less) to one another where both sampled water quality, one point would be eliminated. The point with greater depth would be chosen over the shallower point. This was done because data at these distances was seen as redundant and time could be better spent increasing the scope of water quality in the river. Other revisions that were suggested were the adding of a few points. These points would be one at the mouth of Muddy Creek (approx. one mile west of R1) in the north western part of the Rhode River, a sampling location at the mouth of the Rhode river where it drains into the West river, and a sampling location in the middle of the West river, approx. 1.5 mile south west of the river's mouth (W6). Over all, the existing sampling locations were well placed and only minor revisions or suggestions were made. It is felt that several new locations, along with a faster sensor to test these locations would give the River Keeper a greater capability in monitoring the water of the West and Rhode rivers. See Figure 8.

Safety Utility. The input into determining these weights was also from the input of the River Keeper and other water quality monitoring experts. They had scored factors of common swimming area, surrounding population density, and potential run- off area as the highest in overall importance. Volatility of the parameters were scored lower. The seasonal and weather changes determined the volatility of the parameters and is not necessarily an indicator of a highly hazardous area. Below is the utility function equation that would determine the best locations for sampling. The weights were derived from surveying the River-Keeper. A one to ten scale measure would determine the overall importance of factors to consider when determining where to sample. A value of 1 was given to factors of least importance, a value of 10 for those factors with greatest importance. The sum of all the different scores for each factor was taken, where the quotient of the individual score divided by the overall sum determined the weights.

	Importance Factor Value	Weight
Population/Swimming Area	9	0.27

Run Off	8	0.24
Distance	5	0.15
Bacteria	5	0.15
Dissolved Oxygen	3	0.09
Secchi Depth	3	0.09

Figure 6: Safety Utility Scores and Weights

The overall value of a potential sampling point would be the sum product of the weights and its corresponding score. The scores of the individual factors were derived from the output of the ArcGIS interpolation algorithm. The team took the average of the historical data of each sampling point and interpolated that value to understand the general behavior of that parameter over a specified area. GIS data smoothing algorithm would allow parameters of water quality to be accurately measured in between two or more sampling sites. Given the correct data and equations, GIS could project what a value of water quality may be by using existing data. This would allow the river keeper to essentially see what the overall quality of the water is with accuracy rather than knowing the quality of water at select points.

The score was then given to these different locations where the higher end of the scale would be given a score of 4, and the lower end would be given a score of 1. The values in between 1 and 4 were determined respectively by the scale values in the ArcGIS software. Below is the graph of the utility values for each of the sampling sites.

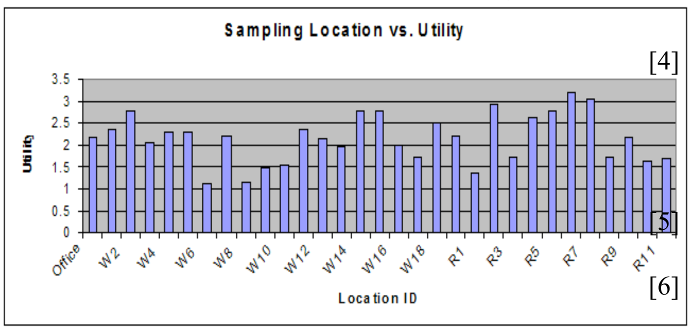

Figure 7: Sampling Location vs. Utility

Below is a mapped recommendation for revision of sampling points, where the top 20 in utility are featured as well as the mentioned revisions for overall water quality sampling.

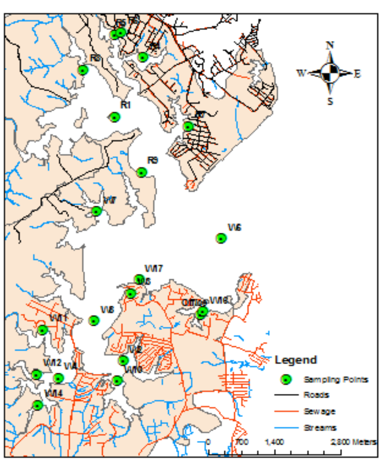

Figure 8: Recommended Sampling Points

VII. Conclusion

Based on analysis with both objective and subjective understanding of the system being improved, certain design alternatives were finalized as the recommended system design choices. The YSI 6820 came out as the greatest value for its functional utility, this will be capable of both a fixed and moving sensor based on the River-Keeper's needs at the time of use. The sensor will accompany the Sivix GSM/GPRS transmitter with telnet programming in order to provide input and output services. Sampling site recommendations listed previously should be taken into account for water quality sampling. The web-portal shall enable external content such as the GIS mapping and data interpretation. This design as a whole will provide increased utility and functionality to the West/Rhode River-Keeper.

References

[1] Maryland Department of Natural Resources. (2008, November 13). *Eyes on the Bay*. [Online] Available: http://mddnr.chesapeakebay.net/

[2] YSI Environmental Monitoring. (2008, October 26 02). YSI Products. [Online] Available: https://www.ysi.com/

[3] West/Rhode River-Keeper. (2008, October 12). Yearly West/Rhode River Report. [Online] Available: http://www.westrhoderiverkeeper.org/

[4] U.S. Environmental Protection Agency. (2008, October 15). Water Quality Standards for Coastal and Great Lakes Recreation Waters. [Online] Available: http://www.epa.gov/fedrgstr/EPA-WATER/2004/November/Day-16/w25303.htm

[5] Trumbauer, Chris. West/Rhode River-Keeper. Contact: riverkeeper@westrhoderiverkeeper.org

[6] Sivix Corporation. Sivix Liaison: Rees, Will. Contact: rees.will@gmail.com

[7] Global Water Instruments Inc. (2008, November 02). WQMS Water Quality Monitoring System. [Online] Available: http://www.globalw.com/products/wqms.html

[8] RSHydro. (2008, November 25). Water Quality Measurement Tools. [Online] Available: http://www.rshydro.co.uk/

[9] In-Situ Inc. (2008, November 23). Troll Water Quality Tools. [Online] Available: http://www.in-situ.com/

[10] National Instruments. (2008, November 28). Developer Zone. [Online] Available: http://zone.ni.com/dzhp/app/main

[11] CICEET. (2008, November 25). Project Brief. [Online] Available: http://ciceet.unh.edu/briefs/caffrey_brief

[12] William and Mary Department of Fisheries Science. (2008, November 27). YSI 650 Equipment. [Online] Available: http://www.fisheries.vims.edu/

[13] Elcee Instrumentation and Services Sdn Bhd. (2008, November 23). Troll Monitoring Equipment. [Online] Available: http://www.elcee-inst.com.my/