



Ionic Concentration Estimation of Urban and Non Urban Water Bodies of Newfoundland and Labrador using Real Time Water Quality Data

Date: July 8th, 2013



**Government of Newfoundland and Labrador
Department of Environment and Conservation
Water Resources Management Division
St. John's, NL, Canada
A1B 4J6**

Abstract

Specific Conductance is one of the continually measured water quality parameters under the Newfoundland and Labrador Real Time Water Quality Monitoring Program. It measures the ability of water to conduct electric current. Its value is dependent on the ionic concentration in water which conveys the current. The greater the ionic concentration in water, the higher will be the value of specific conductance. Common conducting ions found in the waters of Newfoundland and Labrador are sodium, calcium, chloride, and sulphate. These ion concentrations are measured during routine sampling of selected water bodies whereby the traditional grab sample is collected and then shipped to an accredited laboratory for analysis. A relationship can be established between the continuously measured specific conductance and the ionic concentration of sodium, calcium, chloride and sulphate using parametric and non parametric regression models. The established relationship will help predict the ionic concentration in water bodies at any point in time. It will also help investigate the difference in ionic concentration between an impacted and pristine water body. Furthermore ionic concentration due to rainfall, runoff, snowmelt, road side salt applications or accidental spikes can be identified instantaneously and mitigative measures can be implemented more quickly thus minimizing the impact on the aquatic ecosystems. This report studies and compares four water bodies in the island part of Newfoundland and estimates the ionic concentration using continuously measured specific conductance.

Table of Contents

	Page
1. INTRODUCTION	1
1.1. Purpose and Scope	2
1.2. Site and Basin Descriptions	3
1.2.1. Leary's Brook at Clinch Crescent	4
1.2.2. Waterford River at Kilbride	5
1.2.3. Humber River at Humber Village Bridge	6
1.2.4. Rattling Brook Below Bridge	7
2. LITERATURE REVIEW	8
3. DATA MANAGEMENT AND COLLECTION	9
3.1. Data Collection through RTWQ monitoring network	9
3.1.1. RTWQ Instrumentation	9
3.1.2. Real Time Sampling Protocol	10
3.1.3. Grab Sampling Technique	11
3.2. Data Management	11
4. SITE SPECIFIC REGRESSION MODEL	14
4.1. Leary's Brook (LB)	14
4.1.1. Statistical Analysis for Real Time and Grab Sample Data	14
4.1.2. Parametric Regression Model	15
4.1.3. Non Parametric Regression Model	16
4.2. Waterford (WF)	16
4.2.1. Statistical Analysis for Real Time and Grab Sample Data	17
4.2.2. Parametric Regression Model	18
4.2.3. Non Parametric Regression Model	19
4.3. Humber River (HR)	19
4.3.1. Statistical Analysis for Real Time and Grab Sample Data	19
4.3.2. Parametric Regression Model	20
4.3.3. Non Parametric Regression Model	21
4.4. Rattling Brook (RBBB)	21
4.4.1. Statistical Analysis for Real Time and Grab Sample Data	21
4.4.2. Parametric Regression Model	23
4.4.3. Non Parametric Regression Model	23
5. MODEL VALIDATION	24
5.1. Leary's Brook	24
5.1.1. Parametric Regression Model Validation	24
5.1.2. Non Parametric Regression Model Validation	25
5.2. Waterford River	26
5.2.1. Parametric Regression Model Validation	26
5.2.2. Non Parametric Regression Model Validation	27
5.3. Rattling Brook Below Bridge	28
5.3.1. Parametric Regression Model Validation	28
5.3.2. Non Parametric Regression Model Validation	29
6. CONCLUSION AND PATH FORWARD	30
7. ACKNOWLEDGEMENTS	31
8. REFERENCES	32

List of Tables

	Page
Table 2.1: Specific Conductance range, ranks and indicator	8
Table 4.1: Statistical analysis for grab and corresponding real time data in LB	14
Table 4.2: Results for AD Normality Test ($\alpha=0.05$) and Box Plot in LB	15
Table 4.3: OLS model for sodium, calcium, chloride and sulphate in LB	16
Table 4.4: KTRL Model for sodium, calcium, chloride and Sulphate in LB	16
Table 4.5: Statistical analysis for grab and corresponding real time data in WR	17
Table 4.6: Results for AD Normality Test ($\alpha=0.05$) and Box Plot in WR	17
Table 4.7: OLS model for sodium, calcium, chloride and sulphate in WR	18
Table 4.8: KTRL Model for sodium, calcium, chloride and sulphate in WR	19
Table 4.9: Statistical analysis for grab and corresponding real time data in HR	19
Table 4.10: Results for AD Normality Test ($\alpha=0.05$) and Box Plot in HR	20
Table 4.11: OLS model for sodium, calcium, chloride and sulphate in HR	21
Table 4.12: KTRL Model for sodium, calcium, chloride and sulphate in HR	21
Table 4.13: Statistical analysis for grab and real time data in RBBB	21
Table 4.14: Results for AD Normality Test ($\alpha=0.05$) and Box Plot in RBBB	22
Table 4.15: OLS model for sodium, calcium, chloride and sulphate in RBBB	23
Table 4.16: KTRL Model for sodium, calcium, chloride and sulphate in RBBB	23

List of Figures

Figure 1.1: Generic overview of the ionic concentration estimation model	2
Figure 1.2: Geographic location of the four sites chosen for regression model	3
Figure 1.3: Leary's Brook sampling station and drainage area	4
Figure 1.4: Waterford River sampling station and drainage area	5
Figure 1.5: Humber River sampling station and drainage area	6
Figure 1.6: Rattling Brook Below Bridge sampling station and drainage area	7
Figure 3.1: Water Quality Parameter Sensors	9
Figure 3.2: Specific Conductance Sensor	9
Figure 3.3 Calibration and deployment cycle	10
Figure 3.4: Data retrieval and management in ADRS System	13
Figure 4.1: Scatter plot for sodium Vs SC - LB	15
Figure 4.2: Scatter plot for calcium Vs SC – LB	15
Figure 4.3: Scatter plot for chloride Vs SC – LB	15
Figure 4.4: Scatter plot for sulphate Vs SC – LB	15
Figure 4.5: Scatter plot for sodium Vs SC - WR	18
Figure 4.6: Scatter plot for calcium Vs SC – WR	18
Figure 4.7: Scatter plot for chloride Vs SC - WR	18
Figure 4.8: Scatter plot for sulphate Vs SC – WR	18
Figure 4.9: Scatter plot for sodium Vs SC - HR	20
Figure 4.10: Scatter plot for calcium Vs SC – HR	20
Figure 4.11: Scatter plot for chloride Vs SC – HR	20
Figure 4.12: Scatter plot for sodium Vs SC - RBBB	22
Figure 4.13: Scatter plot for calcium Vs SC – RBBB	22

Figure 4.14: Scatter plot for chloride Vs SC – RBBB	22
Figure 5.1: LB - Model Comparison (OLS) with actual grab samples	25
Figure 5.2: LB - Model Comparison (KTRL) with actual grab samples	26
Figure 5.3: WR - Model Comparison (OLS) with actual grab samples	27
Figure 5.4: WR - Model Comparison (KTRL) with actual grab samples	28
Figure 5.5: RBBB - Model Comparison (OLS) with actual grab samples	28
Figure 5.6: RBBB - Model Comparison (KTRL) with actual grab samples	29

1.0 INTRODUCTION

The Water Resources Management Division (WRMD) of the Department of Environment and Conservation (ENVC) of the Province of Newfoundland and Labrador (NL) has established a near real time water quality (RTWQ) monitoring network throughout the province. Continuous RTWQ data is collected using the network. Leary's Brook was the first RTWQ station established in 2001 in the city of St. John's. The network has expanded over the years with representative geographic coverage throughout the Province. The continuous collection of water quality data can be used to monitor the health of aquatic ecosystems, establish trends for water quality measurements and determine when specific water quality events occur.

The information obtained from the network is needed by the WRMD to implement its mandate. It also allows managers and policy makers to make informed decisions on early warning of adverse water quality events. The general public, policy makers, government agencies and private sectors greatly benefit from such timely data and information.

The water quality parameters measured by the NL RTWQ monitoring system are water temperature, pH, dissolved oxygen (DO), specific conductance (SC) and turbidity. Percent saturation and total dissolved solids are two additional parameters calculated from DO and SC. These key indicator parameters provide significant information to better understand the water quality of a particular water body.

Conductivity reveals the presence of dissolved materials in water (Williams, 1966). This dissolved material consists of metallic ions, organic and inorganic materials. Conductivity measured at or corrected to 25°C is called "Specific Conductance". Hence SC is an indirect measure of the amount of dissolved substances (salts) (Williams, 1986).

Sampling performed at a particular point in time is known as a grab sample. The grab sampling is part of the Quality Assurance/Quality Control protocol for NL RTWQ program. Ionic concentration of many dissolved materials in water is measured by means of grab sampling. Some of the parameters measured during a routine grab sample are sodium, calcium, chloride, and sulphate. Laboratory analysis is performed to measure the values of these parameters. Since both the grab sample and the RTWQ measurement are taken at the same point in time it is possible to correlate the grab sample measurement with the RTWQ measurement (see figure 1.1) (Granato and Smith, 1999). Specific conductance, being an indicator for the dissolved materials in water, is more likely to correlate with the ionic concentration of some of the parameters measured during grab sampling (Lind, 1970).

Water quality sampling sites across the island of Newfoundland can be used to analyze the relationship between continuous SC and grab sample measurements of sodium, calcium, chloride and sulphate ions. The sites chosen are Leary's Brook, Waterford River, Humber River and Rattling Brook below bridge. These sites have been sampled extensively for the last four to five years. The sites are selected based on their location

and amount of dissolved solid material received by these water bodies in order that comparative analysis can be drawn from the data obtained.

Figure 1.1 shows a generic overview of the estimation of ionic concentration measurement of grab sample from real time data by applying regression analysis. Real time and grab sample data collected over a period of time are used for statistical analysis to develop a regression model for grab sample estimation. The model provides a measure of strength and variation of the relationship between real time and grab sample data in the four sampling locations.

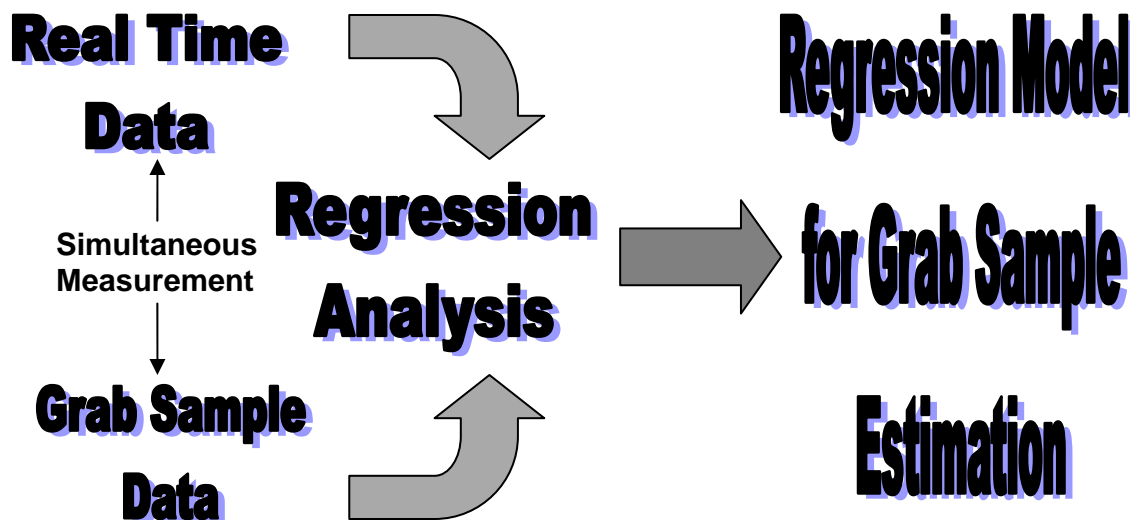


Figure 1.1: Generic overview of ionic concentration estimation model

1.1 Purpose and Scope

The purpose of this report is to estimate ionic concentration measures using regression modeling and making comparative analysis of the results between two urban and two non urban water bodies. Data for both continuous RTWQ and grab sample measurements were collected from January 2006 to September 2010 (January 2007 to September 2010 for Rattling Brook below bridge). This data is used to develop a site specific regression model for sodium, calcium, chloride and sulphate. The results of this report will help to better understand how increased ionic concentration leads to elevated specific conductance at impacted sites. It will also help to estimate ionic concentration in real time using the site specific regression model. Both parametric and non parametric regression analyses are applied to identify which model best fits actual water quality measurements. The results from this report would help the WRMD to save time and effort required for sampling and laboratory analysis. It can also be used as a quality assurance/quality control protocol to estimate individual parameter values. Furthermore, this method would be helpful to estimate water quality variables at discontinued sampling stations.

The report is arranged in the following order:

- Site description and sampling location;
- Literature review;
- RTWQ monitoring, grab sample measurement and data management.
- Statistical analysis on the datasets;
- Site Specific Regression Models;
- Model verification and validation;
- Path forward and conclusion.

1.2 Site and Basin Description

Figure 1.2 shows the location of the four sites from which data is collected. The sites are: Leary's Brook, Waterford River, Humber River and Rattling Brook below bridge. These sites are chosen based on anthropogenic activity and availability of water quality data. Leary's Brook and Waterford River are located in an urban setting with a high level of anthropogenic impact from the surrounding areas. Rattling Brook is non urban but in the middle of a construction site while Humber River is non urban with little impact from surrounding areas. The sites are described in detailed in the following section.

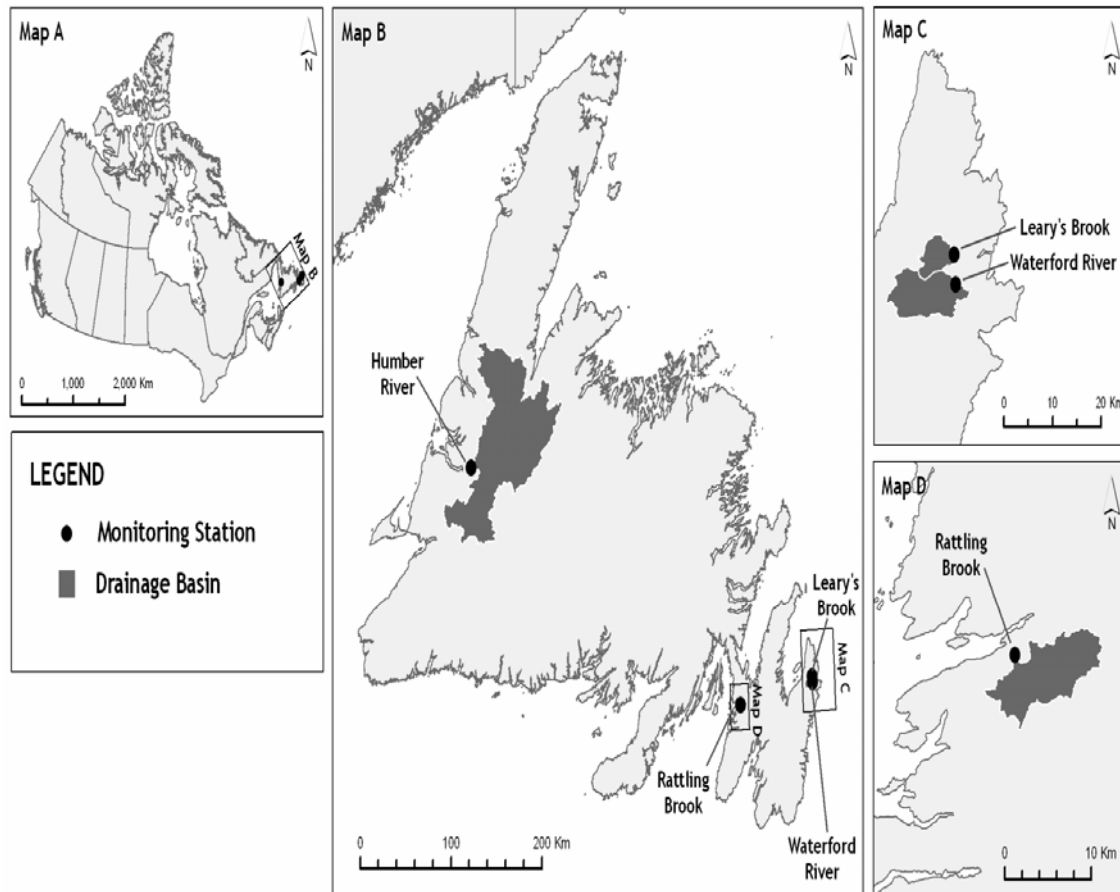


Figure 1.2: Geographic location of the four sites chosen for regression model

1.2.1 Leary's Brook at Clinch Crescent

Leary's Brook was the first RTWQ station in NL established in 2001. The sampling site is located in a developed section of the City of St. John's close to Memorial University. One of the main shopping centers in the city is located immediately upstream of the sampling site where the river is culverted beneath the parking lot. Oxen Pond also drains upstream of the sampling site before it is culverted. Significant urban runoff can be observed in the culvert area as a result of surrounding anthropogenic activities.

Figure 1.3 shows the location of the sampling site along with the drainage basin. As shown, the area is densely surrounded by houses, buildings, and major roads. Road salts are applied during the winter months which affect the water quality within the river. Hence the ionic concentration surges during snowmelt and runoff periods which can be observed by the corresponding increases in specific conductance.

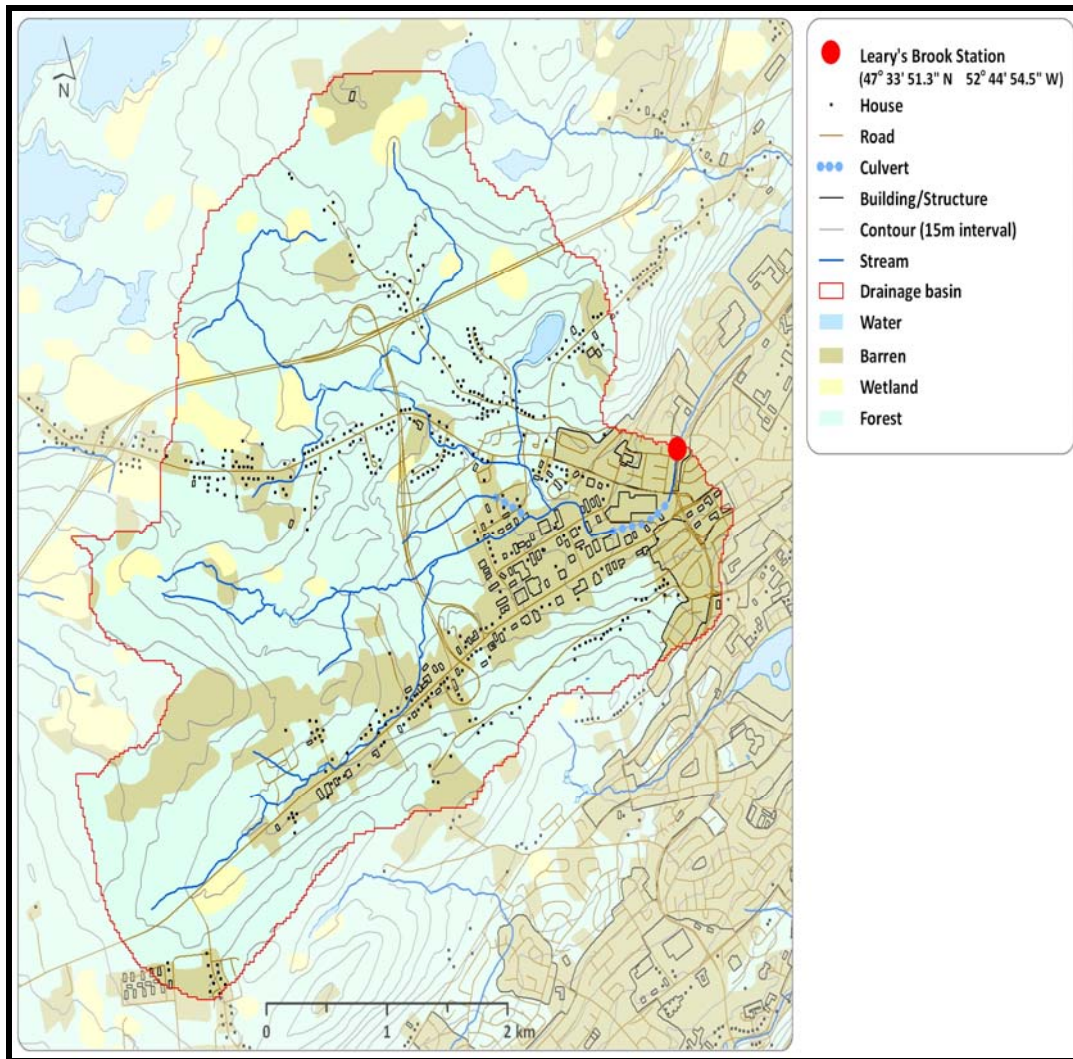


Figure 1.3: Leary's Brook at Clinch Crescent sampling station and drainage area

1.2.2 Waterford River at Kilbride

Waterford River at Kilbride was established in July 2005. The sampling site is situated near the downtown area of the City of St. John's. It is also downstream of Bowring Park, a major recreational park in the City of St. John's. The water passes through numerous rivers and ponds from the Town of Paradise and the City of Mount Pearl before arriving at the sampling site. South Brook drains into the river slightly upstream from the sampling site. The river then flows directly into St. John's Harbor. The river is highly impacted as a result of surrounding anthropogenic influence which affects the quality of water around the sampling site.

Figure 1.4 shows the location of the sampling site along with the drainage basin. As shown, the area around the sampling site is densely surrounded by houses, buildings, roads and highways. Major industrial areas are also located within the drainage basin. Road salts are applied during the winter months which affect the water quality within the river. Hence the ionic concentration surges during snowmelt and runoff periods which can be observed by the corresponding increases in specific conductance.

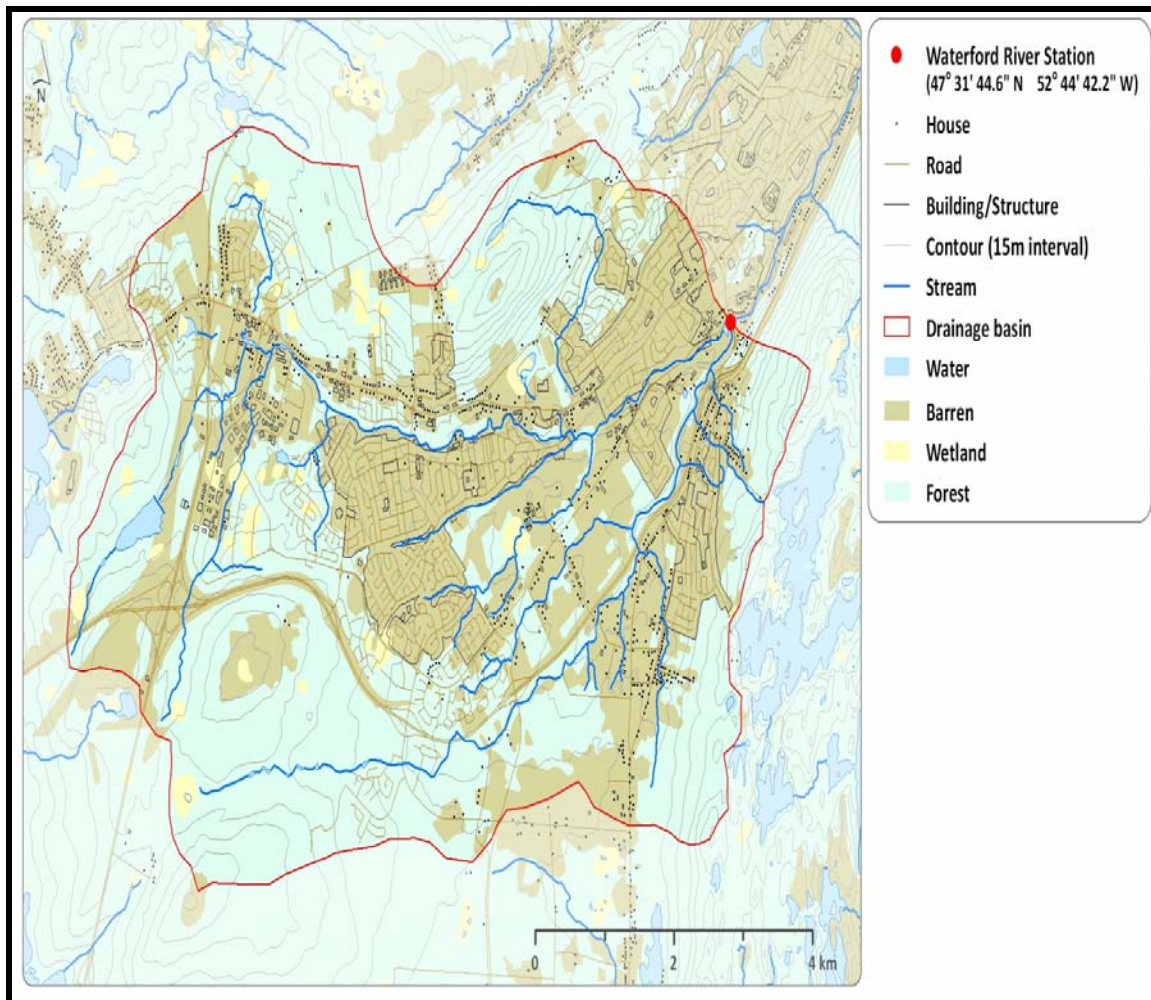


Figure 1.4: Waterford River sampling station and drainage area

1.2.3 Humber River at Humber Village Bridge

The Humber River is the second largest river on the island of Newfoundland. The headwaters of the Humber River flow from the highlands of the Long Range Mountains, through a series of pools and steadies surrounded by extensive boreal forest. It then flows through a deep and heavily forested river valley into a wide flood plain between Adies Lake and Sandy Lake, dominated by extensive marshland. The Humber River tributaries drain the mountainous areas surrounding Grand Lake and Deer Lake, joining the Humber River before it continues along a steep river valley to its outlet at the Humber Arm. The sampling site was established in December 2003 and is located near Humber Village approximately 14 km from the town of Pasadena.

Figure 1.5 shows the sampling location and the drainage basin. The station is classified as non urban. There are a number of small communities located within the watershed but the overall population density is sparse. There are some transportation routes throughout the basin which are salted during the winter months. However, due to the large volume of water within the system, the ionic concentration is diluted.

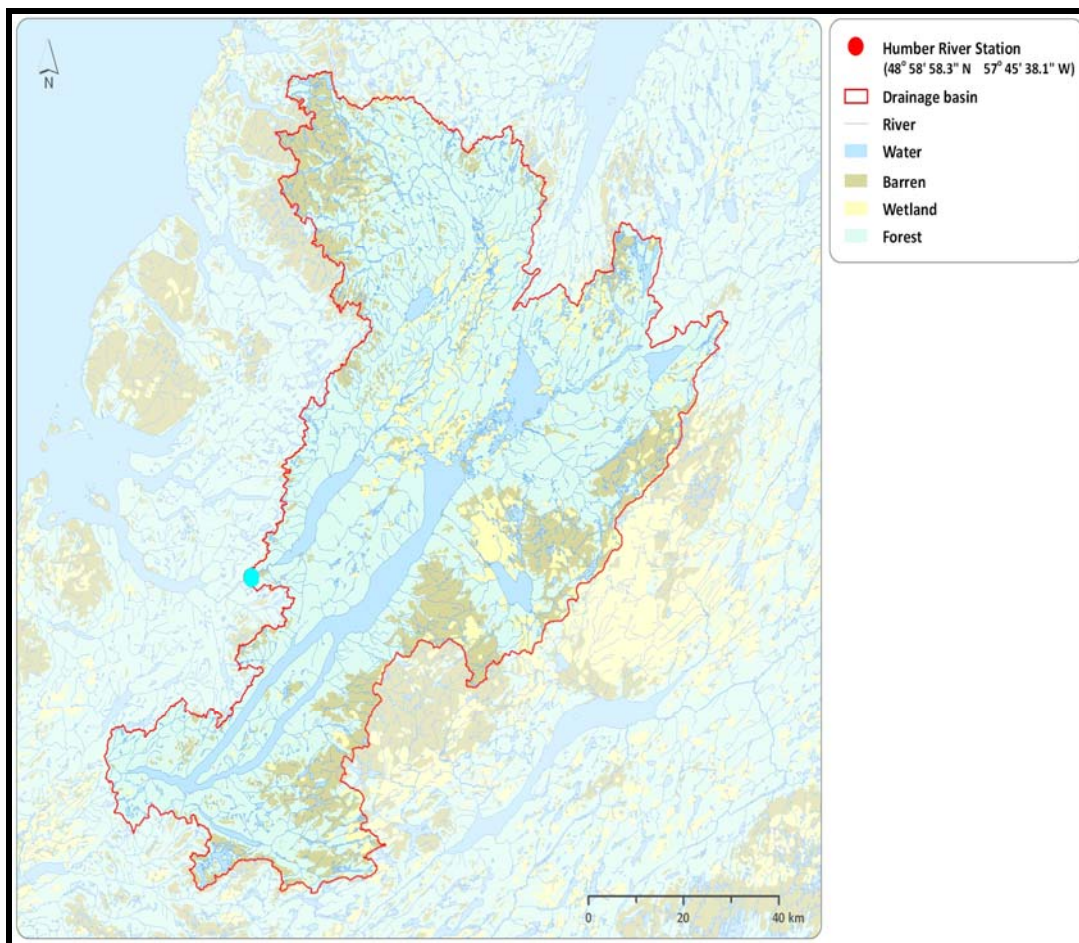


Figure 1.5: Humber River at Humber Village Bridge sampling station and drainage area

1.2.4 Rattling Brook below Bridge

Rattling Brook below bridge was established in December 2006. It is located close to the town of Long Harbour & Mount Arlington Heights, on the southeast of the Avalon Peninsula. It is within the construction zone of Vale NL commercial nickel processing facility. The river is entirely contained within the Vale NL site and is used as a water supply. Major work resulting from the construction of the nickel processing facility is occurring along the river and access to the sampling sites is controlled due to security and safety concerns.

Figure 1.6 shows the sampling station and the drainage basin. There is no permanent population settlement within the watershed. The river channel is generally narrow (<5 meters wide) with moderately deep pools and intermittent braided, shallow areas with some small rapids. The streambed is generally rocky in most locations. There is no major transportation route located within the sampling area. The river is moderately impacted with ionic concentration due to sparse population and the presence of the Vale processing plant and facilities.

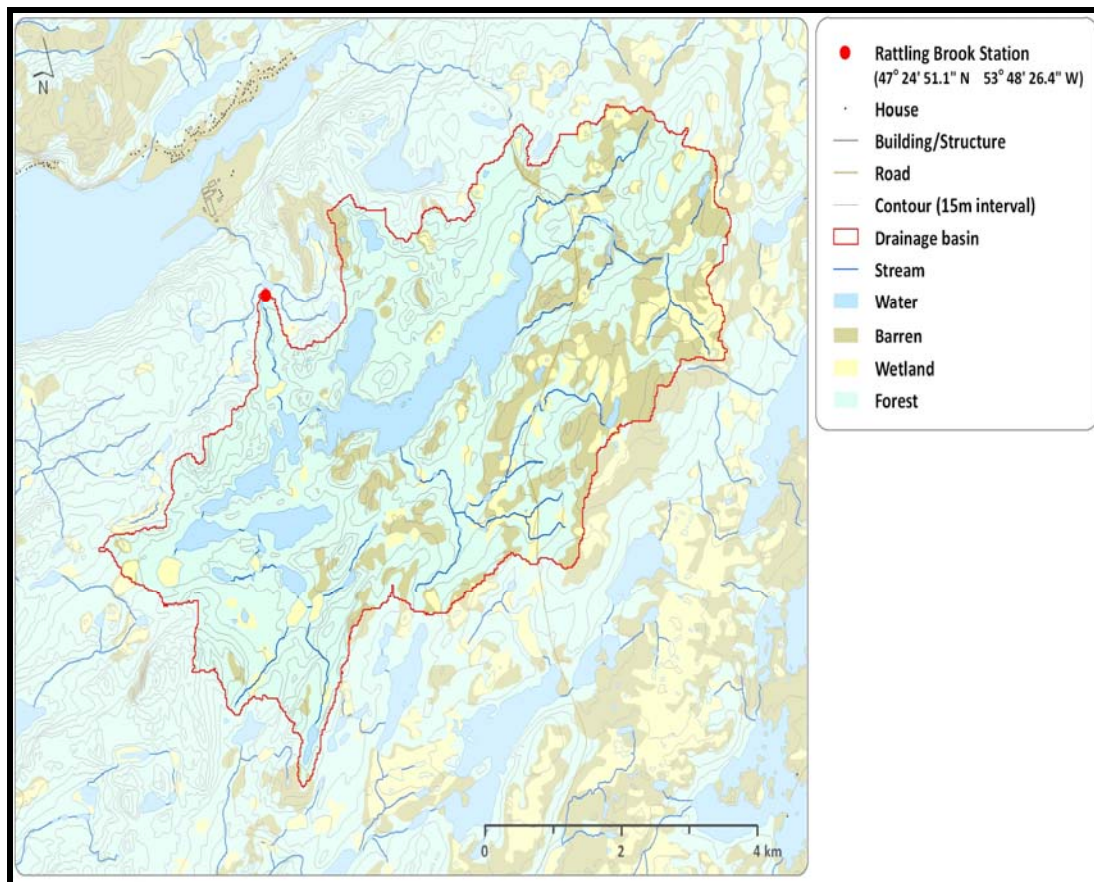


Figure 1.6: Rattling Brook Below Bridge sampling station and drainage area

2.0 Literature Review

Conductivity is a measure of the ability of an aqueous solution to carry an electric current (Eaton et. al. 2005). When measured at or corrected to 25°C, it is called “Specific Conductance”. Specific Conductance (SC) is an indirect measure of the amount of dissolved substances (salts) and may influence the toxicity of many substances. Generally, the ability of pure water to conduct an electric current is low and hence improved water quality lowers the range of SC (Hem, 1982). The SC range along with the ranks and indicator for conducting capability in water is given in the table below.

SC Range ($\mu\text{S}/\text{cm}$)	Ranks	Indicator
0 - 200	Low	Pristine or Background Conditions
200 - 1000	Mid	Normal background Conditions
1000 - 10,000	High	Saline Conditions

Table: 2.1: Specific Conductance range, ranks and indicator

In some of the earlier studies (Williams, 1966, Sing and Kalra, 1975 and Thomas, 1986), linear relationships were identified between the conductivity and total dissolved solids. Studies performed by Lind (1970) had shown that it is possible to estimate concentration of individual ionic constituents from continuous measurement of SC since the ionic composition remains constant even though the concentration may vary as a result of dilution. Miller and others (1988) have shown that estimation of water quality constituents can be applied in locations where continuous water quality measurements have been discontinued and also as a part of the quality assurance/quality control program to verify chemical analyses of discrete water samples. Stevens, O'Bric and Carton (1995) measured the relationship between electrical conductivity and nutrient content of animal slurries using correlation and linear regression analyses. All these studies indicate a strong relationship between specific conductance and selected water quality parameters.

A detailed study by Granato and Smith (1999) in Northborough, Massachusetts showed the application of continuously monitored SC data to estimate road salt concentration. Regression analysis was applied in this study to measure constituent calcium, sodium, and chloride on the basis of continuous records of SC of highway runoff. Christensen, Xiaodong, and Ziegler (2000) and Ryberg (2006, 2007) have also developed regression equations to estimate constituent concentration yields in water bodies in Kansas and North Dakota. Reham El-Korashey (2009) has applied regression analysis to estimate sodium and chloride in Bahr El Baqar Drain in Egypt using electrical conductivity as an explanatory variable. These studies show that regression analysis is the standard method for modeling the relationship between specific conductance and water quality constituents.

The proposed study will not only verify the relationship between specific conductance and water quality constituents but also compare parametric and non parametric regression models. All models will be validated with actual values to identify the performance of the models using parametric and non parametric approaches.

3.0 DATA COLLECTION AND MANAGEMENT

3.1 Data Collection through RTWQ Monitoring Network

The RTWQ Monitoring Network consists of a series of monitoring stations across NL. The continuous collection of water quality data can be used to monitor the health of aquatic ecosystems, establish trends and determine when specific water quality events occur. The network is used to collect key water quality indicator parameters, such as pH, temperature, specific conductance, turbidity and dissolved oxygen and utilize this data to catch emerging water quality events. The RTWQ Monitoring Network has been established through joint partnership with industries and the federal government (Environment Canada).

3.1.1 RTWQ Instrumentation

The RTWQ monitoring instrument is deployed beneath the water's surface in a representative section of the stream in order to measure basic water quality parameters. The instrument which continuously measures RTWQ parameter data at a sampling station is referred to as field sonde. The instrument used to check the accuracy of the field sonde at a sampling site is referred to as Quality Assurance/Quality Control (QA/QC) sonde. The QA/QC sonde is always freshly calibrated before use while the field sonde is calibrated at the time of redeployment at the sampling station and is left at the site for a period of 30 days. Figure 3.1 shows a sonde deployed in a stream and the sensor components of a sonde.



Figure 3.1: Water Quality Parameter Sensors

Figure 3.2 shows a SC sensor removed from a sonde. The SC sensor acquires data by utilizing its open cell design with four graphite electrodes (Campbell Scientific, 2007).



Figure 3.2: Specific Conductance Sensor

A current is passed between two electrodes which are held at a fixed potential. The SC sensor measures the resistance of small electrical currents passing through the pins of a sensor. The level of conductance is directly related to the amount of electrical current passing between the pins. Increased electrical current causes a greater SC measurement (Campbell Scientific, 2007).

3.1.2 Real Time Sampling Protocol

The RTWQ monitoring program follows quality assurance, quality control and quality assessment procedures in order to ensure the effectiveness and reliability of data. Quality Assurance (QA) includes all high-level activities, structures and mechanisms used to ensure and document the accuracy, precision, completeness, effectiveness and representivity of the RTWQ monitoring program. Quality control refers to the technical activities employed to ensure that the data collected are adequate for quality assessment purposes. Quality assessment activities are implemented to quantify the effectiveness of the quality control procedures.

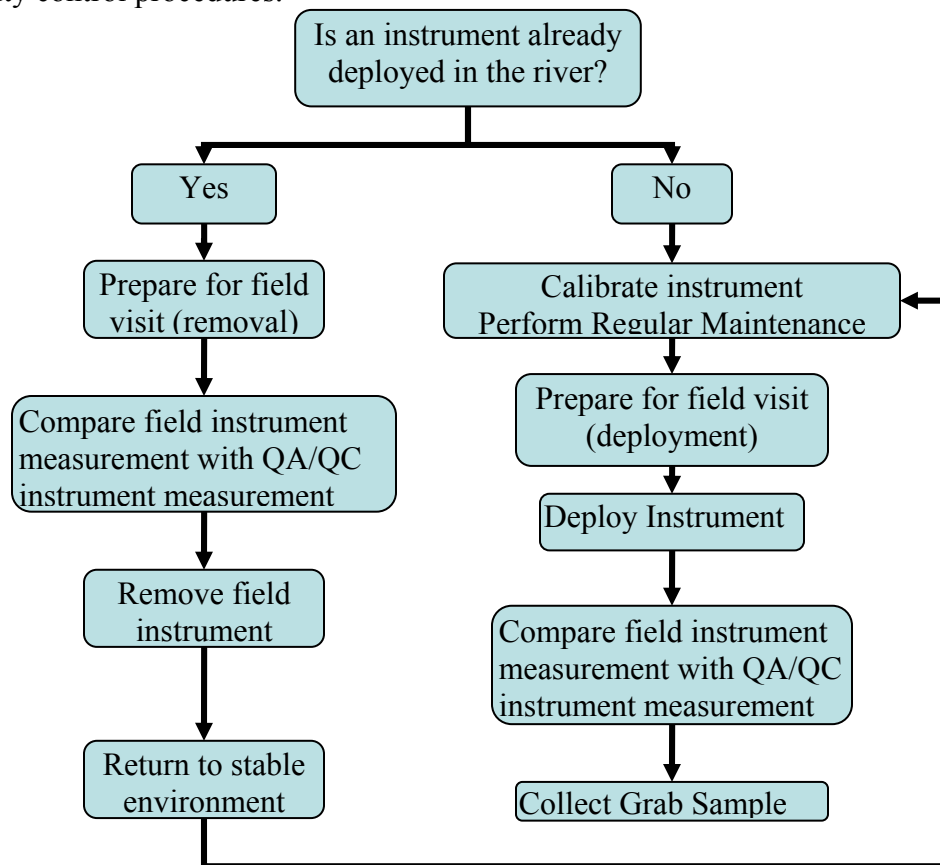


Figure 3.3: Calibration and deployment cycle

In NL, the RTWQ parameter sensors are generally deployed, calibrated and removed on a thirty day cycle as shown in figure 3.3. At the beginning of the deployment period (field sonde deployment) a comparison is made between the field sonde and the QA/QC sonde measurement in order to measure the accuracy of the field sonde measurement. The grab

sample is collected when the sonde is deployed at the beginning of the thirty day period in order that the laboratory results can be compared with a freshly calibrated sonde as it enters deployment. Data collected from the field sonde is compared with QA/QC sonde at the end of deployment period. Any significant shift in field sonde measurement will be identified at this time. To further measure the accuracy of the field sonde measurement, all the collected data is entered into an Excel spreadsheet which is transformed into automated graphs. These graphs are used to identify whether field sonde readings are within acceptable ranges for the given sensors.

3.1.3 Grab Sampling Technique

In order to perform a grab sample a plastic bottle container is obtained from an accredited lab (for major ions a 1000 ml bottle is used). The container is labeled properly in order to identify the sample number, station, date, time and any relevant sample information (CCME, 2011). The cap for the container is then removed and a sample is collected from the stream at a depth of approximately 0.3 m under and against the downward current of water. The opening of the container must face upstream. The container is rinsed twice with the stream water to ensure that a representative sample is collected. Less accessible water body may require the use of poles and buckets to collect grab samples. The bottles must remain capped until samples are collected and stored under clean conditions ensuring that the inside of the bottle is not touched at any times to avoid contamination.

Once the grab sample is collected and analyzed in an accredited lab, the results for the parameter values are returned to the WRMD on a monthly basis. Sodium, calcium, chloride and sulphate are few of the many parameters measured in a typical grab sample.

3.2 Data Management

Figure 3.3 shows the WRMD procedure for real time data transmission and management. All RTWQ parameter data is retrieved through the Automatic Data Retrieval System (ADRS), a tool developed by the WRMD. It is a series of microcomputer based programs which automatically collects, processes and distributes the near real time water quality, stream flow and climate data. In this system, data is collected from remote sites via the Geostationary Operational Environmental Satellite (GOES) system. When the field sonde takes a reading, it is immediately transmitted to the National Environmental Satellite Data Information System that is operated by NOAA in Maryland, USA. Other data collection techniques are also utilized as necessary. Data for real time stations at Leary's Brook and Waterford River are collected using a dial-up modem.

Real time stations are continuously recording large amounts of data with intervals ranging from every fifteen minutes to an hour. The interval length is dependent on the method of data collection (GOES/dial up). For dial up data collection, the interval is generally shorter than the GOES data collection due to the flexibility of frequent transmission without loss of data. Water quality changes in streams can happen much quicker than they would in a remote or pristine site and hence dial-up data collection would be suitable for this purpose.

All parameter data are logged by a datalogger at the RTWQ monitoring stations and transmitted through GOES/dial-up into the Oracle database server maintained by the WRMD. The ADRS stores and collects raw data downloaded from the GOES/dial-up system. It then populates RTWQ graphs of each parameter at each real time station online allowing a visual representation of the parameters. These graphs aid in identifying trends over specific time periods and provide a method for tracking any disturbances or changes in water characteristics. The graphs are available for public viewing and are updated approximately every two hours. The obtained data can be exported and downloaded along with previously obtained data, as specified by the user.

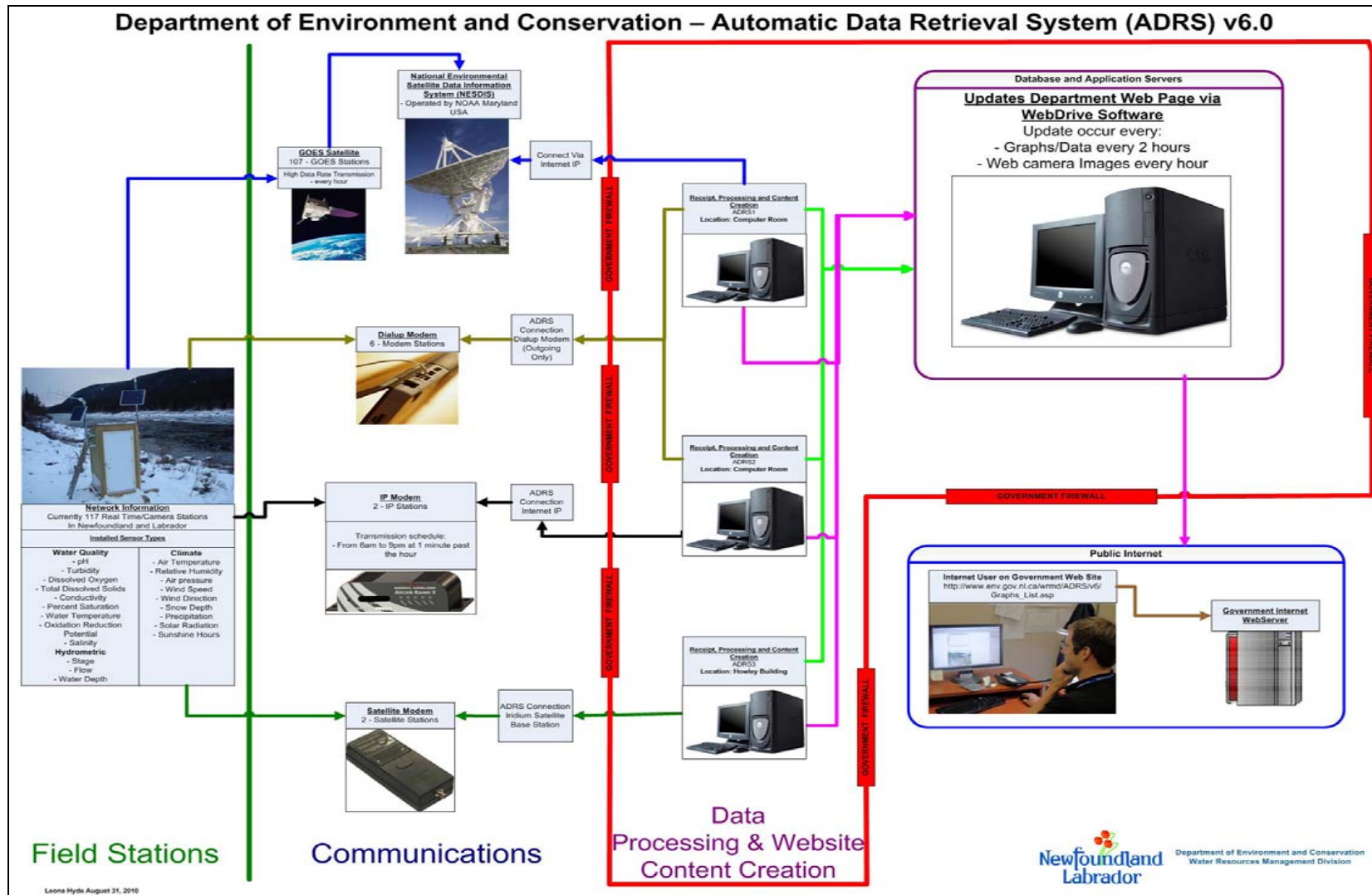


Figure 3.4: Data retrieval and management in ADRS System

4.0 SITE SPECIFIC REGRESSION MODEL

Regression analysis is a widely used technique in defining the mathematical relationship between two or more variables. Regression analysis has been applied to estimate individual water quality variables (Hem, 1992). In order to apply regression analysis method, one must ensure that the data obtained fulfills the pre-conditions required for regression analysis (Helsel and Hirsch, 2002). Since the model for same parameter can change from one stream to another, the data used to develop the model must be site specific for a water body in order to reflect the local changes in the variables. This would also ensure the accuracy of the model locally.

The parametric Ordinary Least Square (OLS) has been applied in many studies (Granato, 1999, Christensen, 2000, 2002) to estimate water quality constituents. The non parametric alternative Kendall Theil Robust Line (KTRL) can also be applied (Helsel and Hirsch, 2002, Granato, 2006) to model between two variables. KTRL is resistant to the effects of outliers and non normality in residuals that commonly characterize hydrologic data sets. In this study, both OLS and KTRL have been applied at the four sampling stations to derive site specific regression model.

4.1 Leary's Brook (LB)

Five years of data from January 2006 to September 2010 were used to develop site specific models for sodium, calcium, chloride and sulphate using specific conductance as a predictor. Statistical analysis, parametric and non parametric modeling was applied to these data.

4.1.1 Statistical Analysis for Real Time and Grab Sample Data

The table below shows statistical measurements of grab sample parameters and the corresponding specific conductance data obtained during this period in real time.

Table 4.1: Statistical analysis for grab data and corresponding real time data in LB

Statistical Measurement	Real Time Parameter	Grab Sample Parameters			
	Cond ($\mu\text{S}/\text{cm}$)	Na (mg/l)	Ca (mg/l)	Cl (mg/l)	SO ₄
Minimum	148.10	26.00	4.20	35.00	6.00
Maximum	1346.00	270.00	16.00	420.00	18.00
Mean	450.59	77.77	8.25	122.90	9.81
Standard Deviation	268.72	53.25	3.04	86.64	2.69
Median	360.20	63.00	8.00	94.00	9.00
5th Percentile	220.00	35.50	4.70	53.50	7.50
95th Percentile	1045.00	190.00	14.00	305.00	15.50

The statistical measurements show high variations in sodium and chloride while a low variation in calcium and sulphate values of the collected parameters. The high variations can be due to increased snowmelt or storm runoff that takes place during seasonal weather changes.

The Anderson Darling Normality test was performed for specific conductance, sodium, calcium, chloride and sulphate using the collected data. Box plot analysis was also used to detect the presence of outliers. The results of the analysis are shown in the table below.

Table 4.2: Results for AD Normality Test ($\alpha = 0.05$) and Box Plot Outlier in LB

Parameter	AD Normality Test		Box Plot Outlier
	Normal/Non Normal	P-value	Present/Absent
Specific Conductance	Non Normal	<0.005	Present
Sodium	Non Normal	<0.005	Present
Calcium	Normal	0.059	Absent
Chloride	Non Normal	<0.005	Present
Sulphate	Non Normal	<0.005	Present

The above table shows that all parameters are non normal with the exception of calcium. Outliers were also detected in all parameter data except calcium. Scatter plot graphs were obtained using Minitab™ to check for linear patterns for sodium, calcium, chloride and sulphate with respect to specific conductance (SC). The scatter plots display linear patterns for each of the parameters against SC.

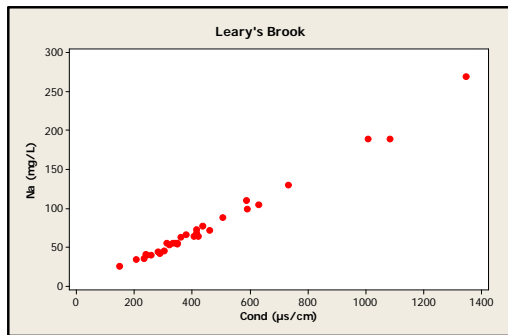


Figure 4.1: Scatter plot for sodium vs SC - LB

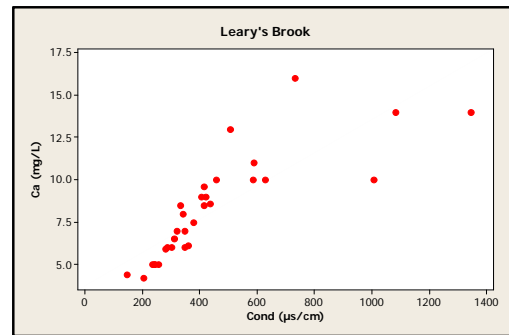


Figure 4.2: Scatter plot for calcium vs SC - LB

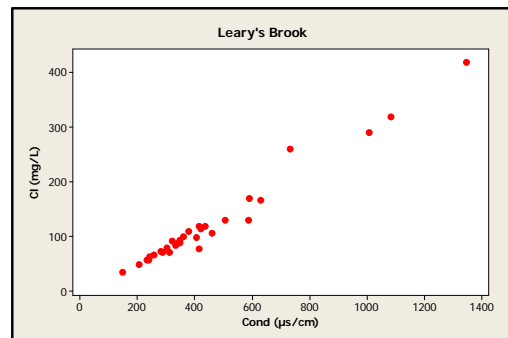


Figure 4.3: Scatter plot for chloride vs SC - LB

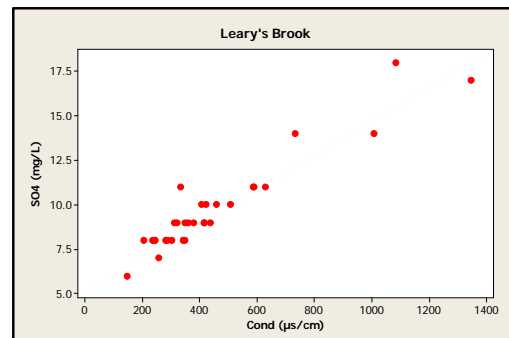


Figure 4.4: Scatter plot for sulphate vs SC - LB

4.1.2 Parametric Regression Model

Due to non normality and the presence of outliers in most of the above parameter data values, log transformation was performed on the original data. Ordinary least square was applied on the log transformed data using Minitab™. Thirty grab samples were used for developing the model. Bias correction (Duan, 1983) was performed on the log transformed model. The results are shown in the table below:

Table 4.3: OLS model for sodium, calcium, chloride and sulphate in LB

[Cond: Specific Conductance; R-square: the proportion of variation in the response data that is explained by the predictor; P-value: tells whether or not the association between the response and predictor(s) is statistically significant]

Computed Variable	Computed Variable Range	Regression Model	R-square	P-Value	Bias Correction (Duan, 1983)
Sodium	Na: 26 - 270	$\log(\text{Na}) = -0.979 + 1.08 \times \log(\text{Cond})$	98.70%	0	0.992
Calcium	Ca: 4.2 - 16	$\log(\text{Ca}) = -0.811 + 0.654 \times \log(\text{Cond})$	80.60%	0	1.015
Chloride	Cl: 35 - 420	$\log(\text{Cl}) = -0.878 + 1.11 \times \log(\text{Cond})$	96.70%	0	1.028
Sulphate	SO ₄ : 6 - 18	$\log(\text{SO}_4) = -0.220 + 0.461 \times \log(\text{Cond})$	88.40%	0	1.005

The above table shows that specific conductance can explain the variation of most of the above parameters (indicated by high R-square values). The association between the specific conductance and the above parameters is statistically significant as indicated by P-values < 0.05. The association is stronger in sodium and chloride in comparison with calcium and sulphate.

4.1.3 Non Parametric Regression Model

Non parametric KTRL was applied to the original data. The obtained models along with the non parametric statistics are shown in the table below:

Table 4.4: KTRL Model for sodium, calcium, chloride and sulphate in LB

[Cond: Specific Conductance; Median Deviation – a non parametric measure of variability; Median Absolute Deviation – median of estimator of spread in the population of residual errors;]

Computed Variable	Regression Model	Median Deviation	Median Absolute Deviation	Confidence Interval (5 th /9 th Percentile)
Sodium	$\text{Na} = -2.78 + 0.182 \times \text{Cond}$	- 2.841	3.952	0.1727 - 0.195
Calcium	$\text{Ca} = 2.13 + 0.0162 \times \text{Cond}$	- 0.807	0.943	0.011 - 0.0202
Chloride	$\text{Cl} = -14.57 + 0.301 \times \text{Cond}$	2.97	4.93	0.278 - 0.32
Sulphate	$\text{SO}_4 = 5.78 + 0.0089 \times \text{Cond}$	- 0.0517	0.439	0.007 - 1.13

The above table shows the KTRL model for sodium, calcium, chloride and sulphate in Leary's Brook. A low value of median deviation and MAD would indicate a good fit for KTRL model. The values are higher for sodium and chloride as a result of high variation in the original data.

4.2 Waterford River (WR)

Five years of data from January 2006 to September 2010 were used to develop site specific models for sodium, calcium, chloride and sulphate using specific conductance as a predictor. Statistical analysis, parametric and non parametric modeling was applied to these data.

4.2.1 Statistical Analysis for Real Time and Grab Sample Data

The table below shows statistical measurements of grab sample parameters and the corresponding specific conductance data obtained during this period in real time.

Table 4.5: Statistical analysis for grab and corresponding real time data in WR

Statistical Measurement	Real Time Parameter	Grab Sample Parameters			
	Cond ($\mu\text{S}/\text{cm}$)	Na (mg/l)	Ca (mg/l)	Cl (mg/l)	SO ₄ (mg/l)
Minimum	235.00	33.00	5.00	51.00	7.00
Maximum	1417.00	280.00	21.00	550.00	22.00
Mean	529.45	92.57	11.09	146.73	12.00
Standard Deviation	303.04	61.78	3.38	110.70	3.53
Median	438.50	68.00	11.00	110.00	11.00
5th Percentile	255.77	41.90	6.95	59.45	8.45
95th Percentile	1152.95	221.00	16.55	342.00	20.20

The statistical measurements show high variations in sodium and chloride while a low variation in calcium and sulphate values of the collected parameters. The high variations can be due to increased snowmelt or storm runoff that takes place during seasonal weather changes.

The Anderson Darling Normality test was performed for specific conductance, sodium, calcium, chloride and sulphate using the collected data. Box plot analysis was also used to detect the presence of outliers. The results of the analysis are shown in the table below:

Table 4.6: Results for AD Normality Test ($\alpha = 0.05$) and Box Plot Outlier in WR

Parameter	AD Normality Test		Box Plot Outlier
	Normal/Non Normal	P-value	Present/Absent
Specific Conductance	Non Normal	<0.005	Present
Sodium	Non Normal	<0.005	Present
Calcium	Normal	0.319	Present
Chloride	Non Normal	<0.005	Present
Sulphate	Non Normal	<0.005	Present

The above table shows that all parameters are non normal with the exception of calcium. Outliers were detected in all parameter data. Scatter plot graphs were obtained using Minitab™ to check for linear patterns for sodium, calcium, chloride and sulphate with respect to specific conductance. The scatter plots display linear patterns for each of the parameters against SC.

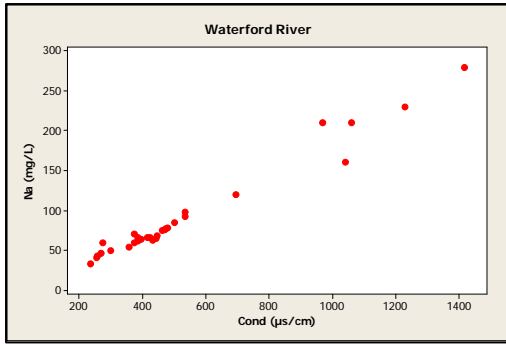


Figure 4.5: Scatter plot for sodium Vs SC - WR

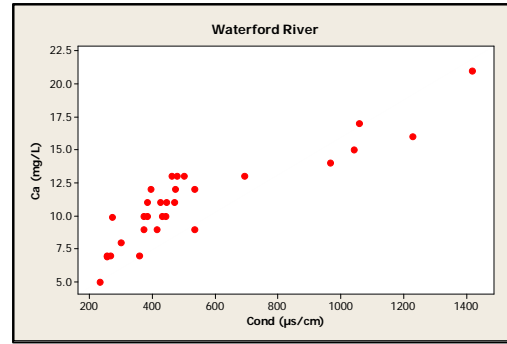


Figure 4.6: Scatter plot for calcium Vs SC - WR

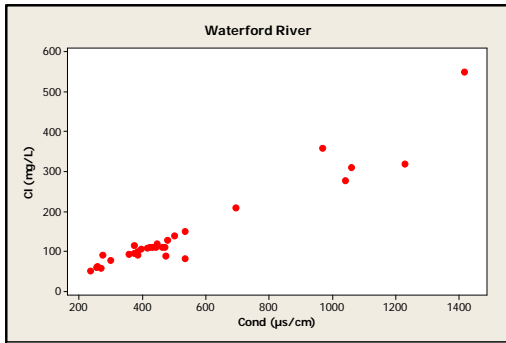


Figure 4.7: Scatter plot for chloride Vs SC - WR

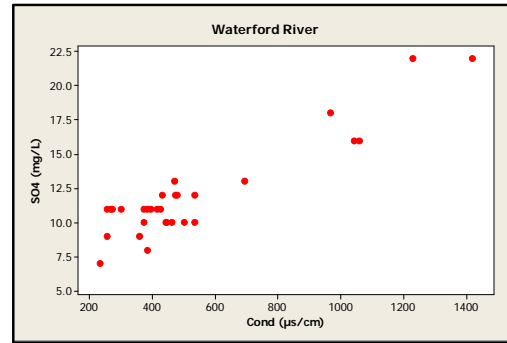


Figure 4.8: Scatter plot for sulphate Vs SC - WR

4.2.2 Parametric Regression Model

Due to non normality and the presence of outliers in the above parameter data values, log transformation was performed on the original data. Ordinary least square was applied on the log transformed data using Minitab™. Twenty nine grab samples were used for developing the model. Bias correction (Duan, 1983) was performed on the log transformed model. The results are shown in the table below:

Table 4.7 OLS model for sodium, calcium, chloride and sulphate in WR

Computed Variable	Computed Variable Range	Regression Model	R-square	P-Value	Bias Correction (Duan, 1983)
Sodium	Na: 33 - 280	$\log(\text{Na}) = -1.02 + 1.09 \times \log(\text{Cond})$	96.50%	0	1.022
Calcium	Ca: 5 - 21	$\log(\text{Ca}) = -0.494 + 0.569 \times \log(\text{Cond})$	77.40%	0	1.009
Chloride	Cl: 51 - 550	$\log(\text{Cl}) = -0.990 + 1.15 \times \log(\text{Cond})$	91.70%	0	1.023
Sulphate	SO4: 7 - 22	$\log(\text{SO4}) = -0.182 + 0.466 \times \log(\text{Cond})$	72.40%	0	1.011

The above table shows that specific conductance can explain the variation of most of the above parameters (indicated by high R-square values). The association between the specific conductance and the above parameters is statistically significant as indicated by P-values < 0.05. The association is stronger in sodium and chloride in comparison with calcium and sulphate.

4.2.3 Non Parametric Regression Model

Non parametric KTRL was applied to the original data. The obtained models along with the non parametric statistics are shown in the table below:

Table 4.8: KTRL model for sodium, calcium, chloride and sulphate in WR

Computed Variable	Regression Model	Median Deviation	Median Absolute Deviation	Confidence Interval (5 th /9 th Percentile)
Sodium	$Na = -15.33 + 0.19 \times Cond$	4.307	4.834	0.165 - 0.206
Calcium	$Ca = 5.69 + 0.012 \times Cond$	-1.085	1.422	0.00936 - 0.0178
Chloride	$Cl = -20.57 + 0.297 \times Cond$	5.717	8	0.265 - 0.328
Sulphate	$SO_4 = 7.013 + 0.009 \times Cond$	0.1713	1.06	0.00528 - 0.0117

The above table shows the KTRL model for sodium, calcium, chloride and sulphate in Waterford River. A low value of median deviation and median absolute deviation would indicate the goodness of fit for KTRL model. The values are higher for sodium and chloride as a result of high variation in the original data.

4.3 Humber River (HR)

Five years of data from January 2006 to September 2010 were used to develop site specific model for sodium, calcium, chloride and sulphate using specific conductance as a predictor. Statistical analysis, parametric and non parametric modeling were applied on these data.

4.3.1 Statistical Analysis for Real Time and Grab Sample Data

The table below shows statistical measurements of grab sample parameters and the corresponding specific conductance data obtained during this period in real time. Due to lack of enough variability sulphate data is not included.

Table 4.9: Statistical analysis for grab and corresponding real time data in HR

Statistical Measurement	Real Time Parameter	Grab Sample Parameters		
	Cond ($\mu S/cm$)	Na (mg/l)	Ca (mg/l)	Cl (mg/l)
Minimum	25.50	2.00	3.70	3.00
Maximum	43.40	3.60	5.90	5.00
Mean	34.94	2.53	4.34	4.03
Standard Deviation	4.51	0.54	0.50	0.50
Median	35.60	2.60	4.10	4.00
5th Percentile	28.50	2.00	4.00	3.80
95th Percentile	39.50	3.20	5.00	5.00

The statistical measurement shows little variation in values of the collected parameters. The Anderson Darling Normality test was performed for specific conductance, sodium, calcium, chloride and sulphate using the collected data. Box plot analysis was also used to detect the presence of outliers. The results of the analysis are shown in the table below:

Table 4.10: Results for AD Normality Test ($\alpha = 0.05$) and Box Plot Outlier in HR

Parameter	AD Normality Test		Box Plot Outlier
	Normal/Non Normal	P-value	
Specific Conductance	Normal	0.146	Absent
Sodium	Non Normal	<0.005	Absent
Calcium	Non Normal	<0.005	Present
Chloride	Non Normal	<0.005	Present

The above table shows that all parameters are non normal with the exception of specific conductance. Outliers were present in calcium and chloride data.

Scatter plot graphs were obtained using Minitab™ to check for linear patterns for sodium, calcium, chloride and sulphate with respect to specific conductance. The scatter plots fail to display linear patterns for each of the parameters against SC.

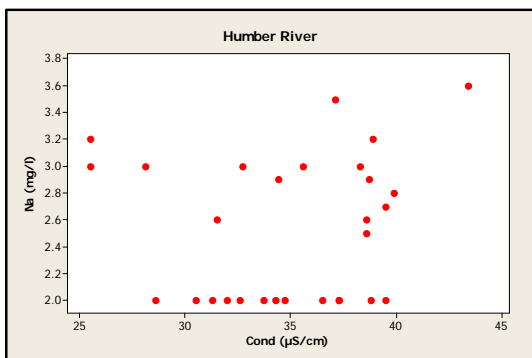


Figure 4.9: Scatter plot for sodium Vs SC - HR

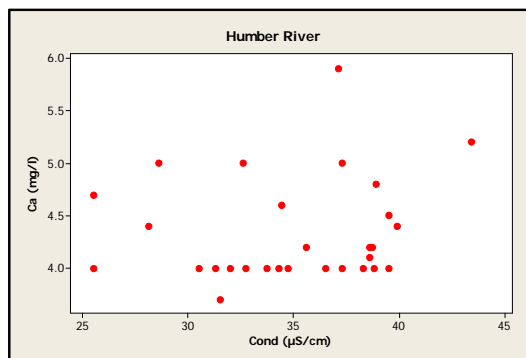


Figure 4.10: Scatter plot for calcium Vs SC - HR

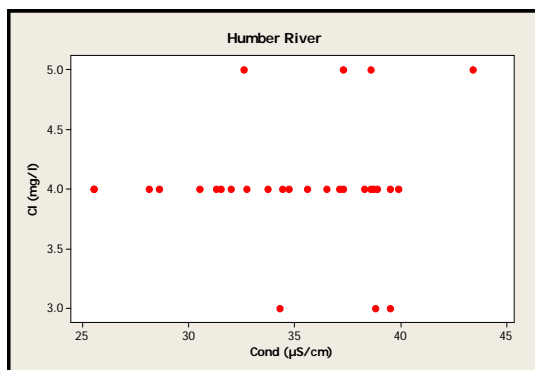


Figure 4.11: Scatter plot for chloride Vs SC - HR

4.3.2 Parametric Regression Model

Due to non normality and the presence of outliers in some of the above parameter data, log transformation was performed on the original data. Ordinary least square regression was applied on the log transformed data using Minitab™. Twenty eight grab samples were used in the modeling. The results for regression analysis are shown in the table below:

Table 4.11: OLS Model for sodium, calcium and chloride in HR

Computed Variable	Computed Variable Range	Regression Model	R-square	P-Value	Total Samples
Sodium	2-3.6	$\log(\text{Na}) = 0.54 + 0.103 \times \log(\text{Cond})$	0.40%	0.734	28
Calcium	3.7-5.9	$\log(\text{Ca}) = 1.14 + 0.092 \times \log(\text{Cond})$	1.3%	0.549	28
Chloride	3-5.0	$\log(\text{Cl}) = 1.32 + 0.019 \times \log(\text{Cond})$	0%	0.916	28

The above table shows that specific conductance cannot explain the variation of the above parameters (low R-square values). The association between the specific conductance and the above parameters is not statistically significant, as indicated by P-value > 0.05.

4.3.3 Non Parametric Regression Model

Non parametric KTRL was applied to the original data. The obtained models along with the non parametric statistics are shown in the table below:

Table 4.12: KTRL model for sodium, calcium, chloride and sulphate in HR

Computed Variable	Regression Model	Median Deviation	Median Absolute Deviation	Confidence Interval (5 th /9 th Percentile)
Sodium	$\text{Na} = 2.6 + 0 \times \text{Cond}$	0	0.599	0 – 0.075
Calcium	$\text{Ca} = 4.1 + 0 \times \text{Cond}$	0	0.099	0 – 0.046
Chloride	$\text{Cl} = 4 + 0 \times \text{Cond}$	0	0	0 - 0

Since the original data is nonlinear for sodium, calcium and chloride, the KTRL model is unlikely to be a good estimator for these parameters. The scatter plot in figures 4.9-11 shows that a good portion of the data points are laying in a straight line with no variation in slope. As a result the conductivity value in the resulting regression model will not contribute in estimation of sodium, calcium and chloride.

4.4 Rattling Brook below bridge (RBBB)

Four years of data from January 2007 to September 2010 were used to develop site specific regression model for sodium, calcium and chloride using specific conductance as a predictor. Due to lack of adequate variation sulphate data is not included.

4.4.1 Statistical Analysis for Real Time and Grab Sample Data

The table below shows statistical measurements of grab sample parameters and the corresponding specific conductance data obtained during this period.

Table 4.13: Statistical analysis for grab and corresponding real time data in RBBB

Statistical Measurements	Real Time Parameter	Grab Sample Parameters		
	Cond (µS/cm)	Na (mg/l)	Ca (mg/l)	Cl (mg/l)
Minimum	27.20	3.00	1.20	5.00
Maximum	41.50	5.20	3.00	9.00

Mean	34.13	4.38	1.76	6.59
Standard Deviation	3.50	0.46	0.34	0.95
Median	35.10	4.40	1.70	6.00
5th Percentile	28.6	3.9	1.34	5.4
95th Percentile	38.78	5.02	2.1	8

The Anderson Darling Normality test was performed for specific conductance, sodium, calcium and chloride using the collected data. Box plot measurements were used to detect the presence of outliers. The results of the analysis are shown in the table below.

Table 4.14: Results for AD Normality Test ($\alpha = 0.05$) and Box Plot Outlier in RBBB

Parameter	AD Normality Test		Box Plot Outlier
	Normal/Non Normal	P-value	Present/Absent
Specific Conductance	Non Normal	<0.005	Present
Sodium	Normal	0.125	Absent
Calcium	Non Normal	0.022	Present
Chloride	Non Normal	<0.005	Present

The above table shows that all parameters except sodium are non normal. Outliers were present in all parameter data except sodium.

Scatter plot graphs were obtained using Minitab™ to check for linear patterns for sodium, calcium, chloride and sulphate with respect to specific conductance. The scatter plots fail to display linear patterns for each of the parameters against SC.

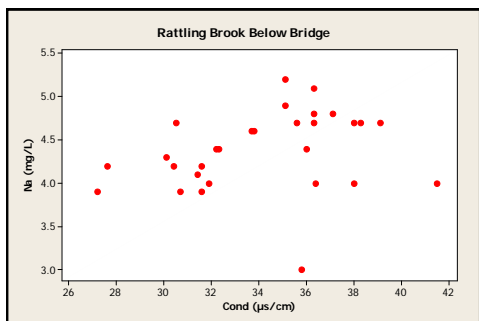


Figure 4.12: Scatter plot for sodium Vs SC - RBBB

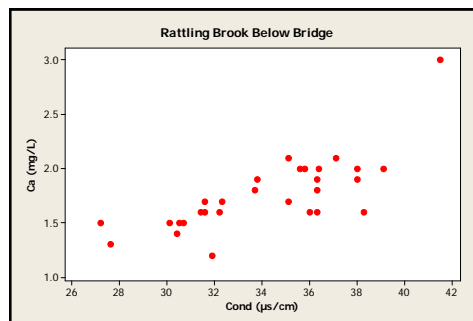


Figure 4.13: Scatter plot for calcium Vs SC - RBBB

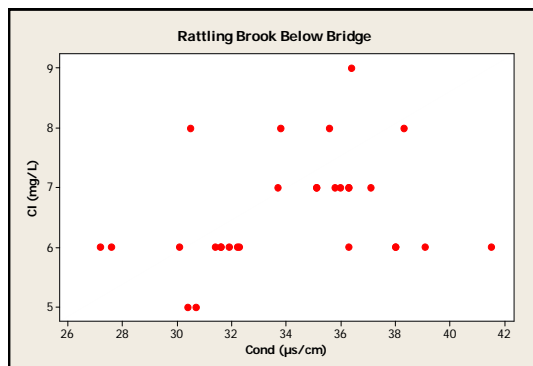


Figure 4.14: Scatter plot for chloride Vs SC - RBBB

4.4.2 Parametric Regression Model

Due to non normality and the presence of outliers in some of the above parameter data values, log transformation was performed. Ordinary least square regression was applied on the log transformed data using Minitab™. Twenty nine grab samples were used in the modeling. The results for regression analysis are shown in the table below:

Table 4.15: OLS Model for sodium, calcium and chloride in RBBB

Computed Variable	Computed Variable Range	Regression Model	R-square	P-Value	Total Samples
Sodium	Na: 3.0 - 5.2	$\log(\text{Na}) = 0.234 + 0.265 \times \log(\text{Cond})$	6.20%	0.192	28
Calcium	Ca: 1.2 - 3.0	$\log(\text{Ca}) = -1.76 + 1.31 \times \log(\text{Cond})$	58.10%	0	28
Chloride	Cl: 5.0 - 9.0	$\log(\text{Cl}) = 0.092 + 0.472 \times \log(\text{Cond})$	12.30%	0.103	28

The above table shows that specific conductance cannot explain the variation of the above parameters (indicated by low R-square values). The association between the specific conductance and the above parameters, except calcium, is not statistically significant as indicated by P-value > 0.05.

4.3.3 Non Parametric Regression Model

Non parametric KTRL was applied to the original data. The obtained models along with the non parametric statistics are shown in the table below:

Table 4.16: KTRL Model for sodium, calcium, chloride and sulphate in RBBB

Computed Variable	Regression Model	Median Deviation	Median Absolute Deviation	Confidence Interval (5 th /9 th Percentile)
Sodium	$\text{Na} = 2.996 + 0.04 \times \text{Cond}$	0.106	0.264	0 - 0.083
Calcium	$\text{Ca} = -0.54 + 0.0638 \times \text{Cond}$	0.099	0.0158	0.033 - 0.083
Chloride	$\text{Cl} = 6 + 0 \times \text{Cond}$	0	1	0 - 0.185

Since the original data is nonlinear for sodium and chloride, the KTRL model is unlikely to be a good estimator for these parameters. The scatter plot in figures 4.12 and 4.14 shows that a good portion of the data points are laying in a straight line with no variation in slope. As a result the conductivity value in the resulting regression model will not contribute in estimation of sodium and chloride.

5.0 MODEL VALIDATION

The obtained results show that it is possible to predict ion concentration of sodium, calcium, chloride, and sulphate from real time data as long as there is enough variation within the parameter values of grab sample data. An urban water body is more likely to show this variation due to high impact from surrounding population, industries, road salts etc and seasonal changes such as snowmelt and runoff. This has been observed in the cases of Leary's Brook and Waterford River. A non urban water body is less likely to be impacted due to fewer anthropogenic influences from the surroundings. This can be observed in the cases of Humber River and Rattling Brook Below Bridge. One exception to this was the variation of calcium values for Rattling Brook which showed statistical significance in the resulting model with the specific conductance values (indicated by R-square and low P value). This can be partially explained by the construction work performed by Vale Inco which may affect the quality of water periodically.

The models in the previous section that have shown strong linear relationships with ionic constituents must be validated in order to identify how similar the estimated values are in comparison with real values. The grab samples collected after the model development were used for this purpose. The grab samples that have been collected for model development is defined as calibration grab samples. The grab samples that have been collected after the model development is defined as validation grab samples.

The models obtained for Leary's Brook, Waterford River and Rattling Brook below Bridge are used for validation. Both parametric (OLS) and non parametric model (KTRL) were compared with calibration and validation grab sample values in order to see which of the two models gives a better fit. The Humber River is not used since there was no variation in any one of its parameter values.

5.1 Leary's Brook

5.1.1 Parametric Regression Model Validation

The graphs in figure 5.1 shows the parametric version of the ion concentration estimation and validation of parameters used for Leary's Brook modeling. The Ordinary Least Square models from table 4.3 are used to estimate ion concentration (sodium, calcium, chloride and sulphate) in real time. The model is represented by a line in the graph. The corresponding grab sample values are placed as points within the graph to see how closely they fit to the model. As shown in the graph, the grab sample lies very close to the regression model line.

In order to validate the model ten additional grab samples are used after the model development. The validation grab samples are represented in different color in order to distinct calibration and validation samples. As shown in the graph the validation grab sample values fit closely to the predicted values in the model. The fitness of the values is closer for sodium and chloride as expected from the model results (table 4.3).

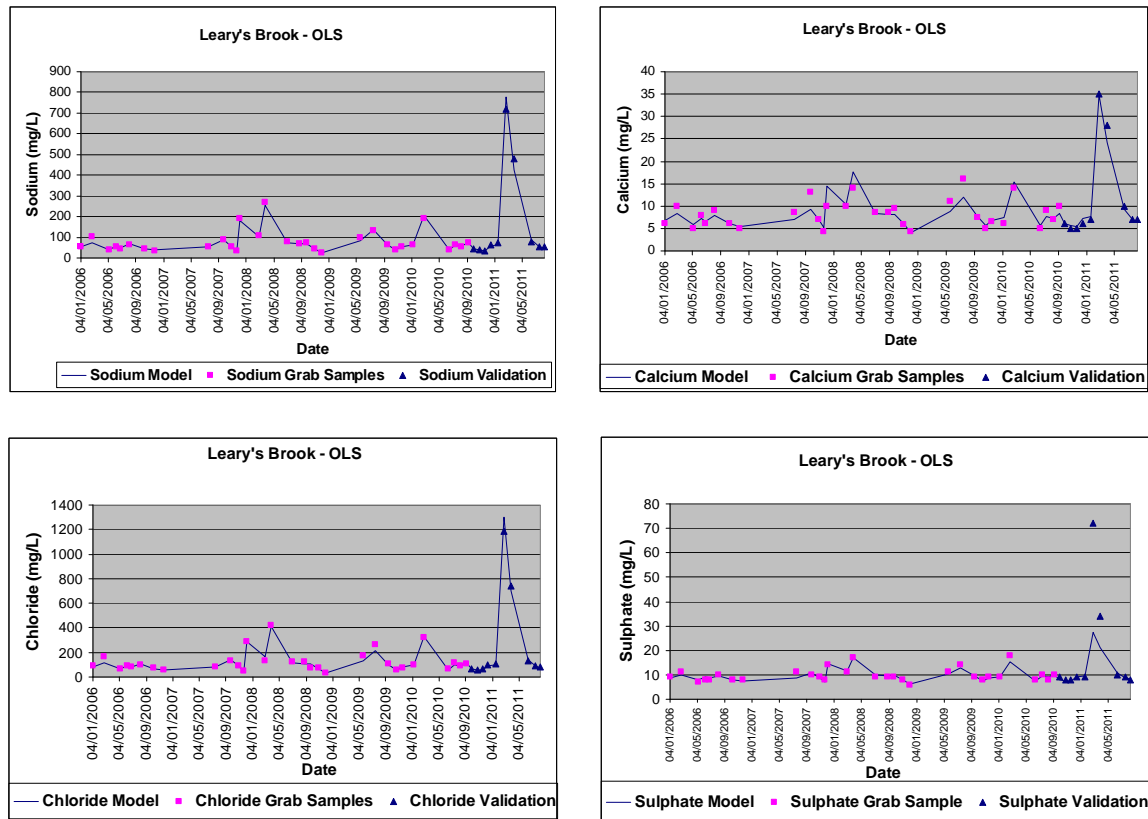


Figure 5.1: LB - Model Comparison (OLS) with actual grab samples

5.1.2 Non Parametric Regression Model Validation

The graph in figure 5.2 shows the non parametric version of the ion concentration estimation and validation of parameters used in Leary's Brook modeling. Kendall Theil Robust Line models from table 4.4 are used to estimate ion concentration (sodium, calcium, chloride and sulphate) in real time. The model is represented by a line in the graph. The corresponding grab sample values are placed as points within the graph to see how closely it fits to the model. As shown in the graph, the grab sample lies very close to the regression model line.

In order to validate the model ten additional grab samples are used after the model development. The validation grab samples are represented in different color in order to distinct calibration and validation samples. As shown in the graph the validation grab sample values fit closely to the predicted values in the model. The fitness of the values is closer for sodium and chloride as expected from the model results (table 4.4).

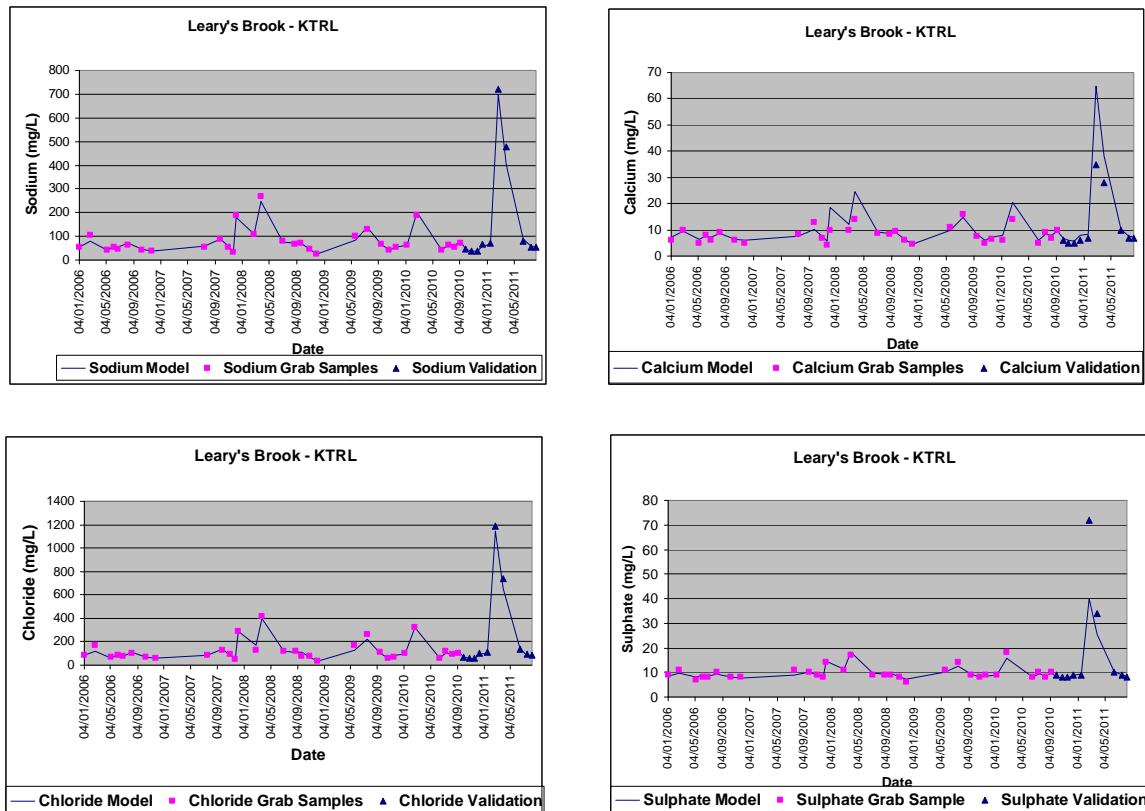


Figure 5.2: Leary's Brook: Model Comparison (KTRL) with actual grab samples

5.2 Waterford River

5.2.1 Parametric Regression Model Validation

The graphs in figure 5.3 shows the parametric version of the ion concentration estimation and validation of parameters used in Waterford River modeling. The Ordinary Least Square models from table 4.5 are used to estimate ion concentration (sodium, calcium, chloride and sulphate) in real time. The model is represented by a line in the graph. The corresponding grab sample values are placed as points within the graph to see how closely it fits to the model. As shown in the graphs, the grab samples lie very close to the regression model line.

In order to validate the model ten additional grab samples are used after the model development. The validation grab samples are represented in different color in order to distinct calibration and validation samples. As shown in the graphs the validation grab samples fit closely to the predicted values in the model. The fitness of the values is closer for sodium and chloride as expected from the model results (table 4.5).

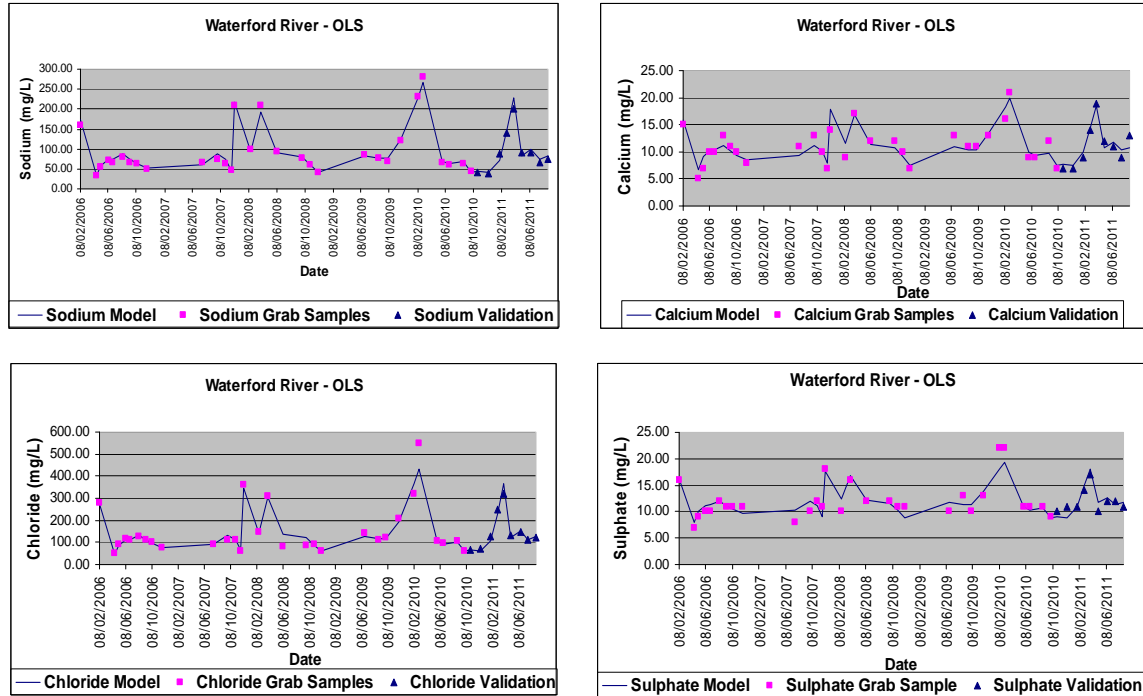


Figure 5.3: WR - Model Comparison (OLS) with actual grab samples

5.2.2 Non Parametric Regression Model Validation

The graph in figure 5.4 shows the non parametric version of the ion concentration estimation and validation of parameters used in Waterford River modeling. Kendall Theil Robust Line models from table 4.6 are used to estimate ion concentration (sodium, calcium, chloride and sulphate) in real time. The model is represented by a line in the graph. The corresponding grab samples are placed as points within the graph to see how closely it fits to the model. As shown in the graphs, the grab samples lie very close to the regression model line.

In order to validate the model ten additional grab samples are used after the model development. The validation grab samples are represented in different color in order to distinct calibration and validation samples. As shown in the graph the validation grab samples fit closely to the predicted values in the model. The fitness of the values is closer for sodium and chloride as expected from the model results (table 4.6).

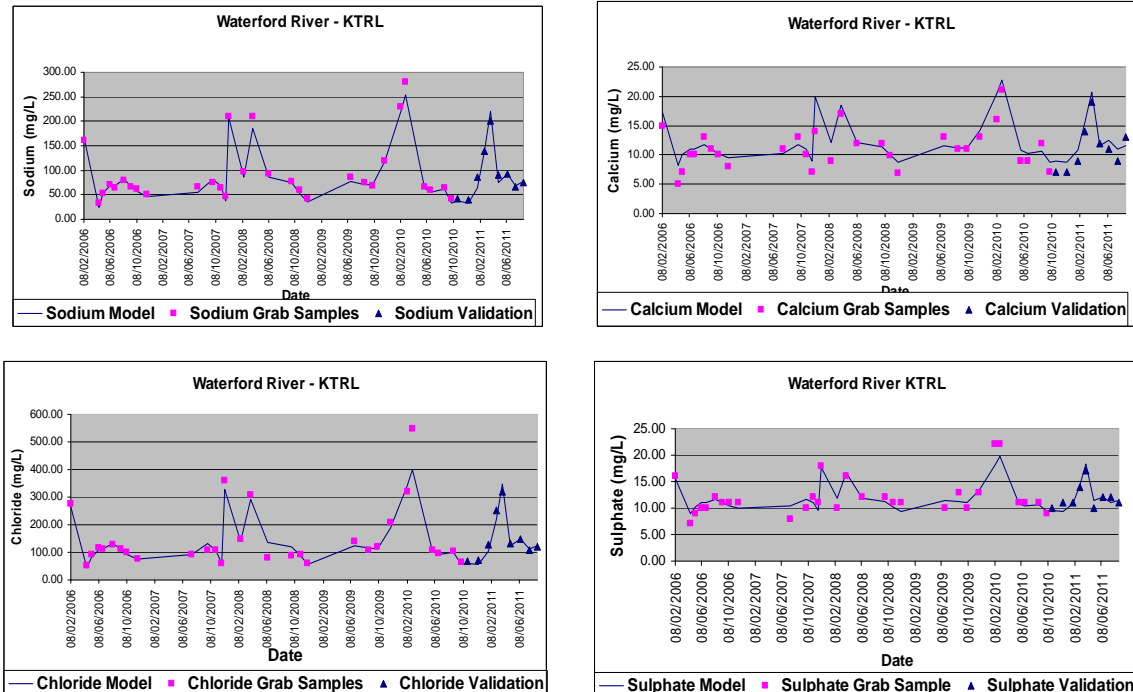


Figure 5.4: WR - Model Comparison (KTRL) with actual grab samples

5.3 Rattling Brook Below Bridge

5.3.1 Parametric Regression Model Validation

The only parameter that regressed close for Rattling Brook Below Bridge is calcium. The graph below shows the estimation of calcium ion concentration for Rattling Brook below bridge using specific conductance as a predictor. The model from Table 4.15 is used to estimate calcium ion concentration in real time. The corresponding grab sample values were placed within the graph to see how closely they fit within the model. As shown the grab sample(s) lies reasonably closely to the regression model line. In order to validate the model three more grab sample values for sulphate are used after the model development. As shown in the graph these grab sample values lies close to the predicted values of the model.

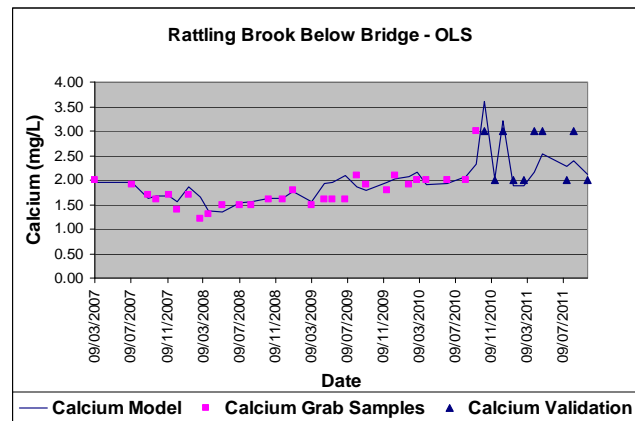


Figure 5.5: RBBB - Model Comparison (OLS) with actual grab samples

5.3.2 Non Parametric Regression Model Validation

The KTRL model from Table 4.16 is used to estimate calcium ion concentration in real time. The corresponding grab sample values were placed within the graph to see how closely they fit within the model. As shown, the grab samples lie close to the regression model line.

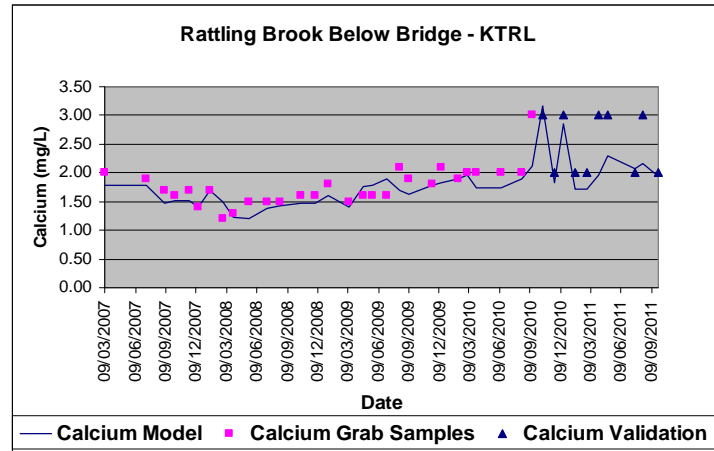


Figure 5.6: RBBB - Model Comparison (KTRL) with actual grab samples

6.0 Conclusions and Path Forward

It can thus be observed that increased variation within grab sample measurement leads to a better regression model. This has been observed in the case of Leary's Brook and Waterford River as well as for calcium in Rattling Brook below bridge. The variation in the level of ionic concentration is largely due to the presence of anthropogenic influence along with low flow conditions. In the case of Humber River with little anthropogenic influence, the ionic concentration of most parameter measurements were below the detection limit, and hence it was difficult to apply any statistical tests to identify if a relationship exists between real time parameters and grab samples. The high flow of water in that river dilutes most of the parameter concentrations which is resonated in the low measurements of parameter values.

For the models where a good estimation of ionic concentration was obtained, there was little difference between the parametric and non parametric approaches. Both approaches were able to predict ionic concentration within reasonable bounds using specific conductance as a predictor.

This study will aid in estimating ionic concentration in real time for the sites where a good fit for regression was obtained. It will also reduce the time delay required to measure water quality constituents at the laboratory by estimating ionic concentration instantaneously. The models obtained from this study can be applied to obtain real-time graphs for sodium, calcium, chloride and sulphate for Leary's Brook and Waterford River by the WRMD.

The parameters where specific conductance was able to provide a good estimation can be viewed graphically in real time. In order to maintain the accuracy of the model, it needs to be calibrated every year with newer available grab samples. This will adjust the model accuracy based on the updated grab sample values.

Potential parameters of interest can be estimated in emerging real time sites using real time parameters as predictors by applying the methodological analysis applied in this study. One such parameter can be total suspended solids (TSS) which can be estimated using real time turbidity. This would be beneficial to industries monitoring real time water quality parameters and would like to ensure that the TSS values are in compliance with the current regulations. Other potential usefulness of this study can be the impact of water quality due to road salts. The models used in this study can be applied in identifying instantaneous increase of sodium and chloride in water and its effect on the quality upon increased road salt application.

7.0 Acknowledgments

The statistical regression analysis and modelling work contained within this report was carried out by Mr. Shibly Rahman (Environmental Scientist - WRMD, Department of Environment and Conservation). Technical advice, comments and project direction were provided by Renee Paterson (Real time Water Quality Program Coordinator). The report was thoroughly reviewed by colleagues Kyla Brake and Rob Wight with valuable insights being provided. Technical advice relating to the "Data Collection and Management" portion of the report was provided by Leona Hyde. Finally, assistance was provided by Keith Abbott in the generation of the station location maps.

References

- Campbell Scientific Canada Corp, 2007. Hydrolab Advanced Maintenance Workshop Manual, Advanced Maintenance and Calibration, Water Quality Instrumentation Training, p. 20.
- Canadian Council of Ministers of the Environment, 2011, Protocols Manual for Water Quality Sampling in Canada, PN 1461, ISBN 978-1-896997-7-0 PDF, 180 pgs.
- Christensen, V.G., Jian, Xiaodong, and A.C. Ziegler, 2000. Regression Analysis and Real Time Water Quality Monitoring to Estimate Constituent Loads and Yields in the Little Arkansas River, South-Central Kansas, 1995-99. U.S. Geological Survey Water-Resources Investigations Report 00-4126, 36 p.
- Christensen, V.G., Rasmussen, P.P., Ziegler, A.C., 2002. Real time water quality monitoring and regression analysis to estimate nutrient and bacteria concentrations in Kansas streams, Water Science and Technology, Vol. 45, No. 9, pp. 205-219.
- Duan, N., 1983, Smearing estimate - a nonparametric retransformation method. Journal of the American Statistical Association, Vol. 78, No. 383, p. 605-610
- Eaton, A. D., Clescere, L. S., Rice E. W., Greenberg, A. E., 2005. Standard Methods for the Examination of Water & Wastewater, American Public Health Association 21st Edition.
- Granato, G. E. & Smith, K. P. 1999. Estimating concentrations of road-salt constituents in highway-runoff from measurements of specific conductance. Northborough, Mass.: U.S. Dept. of the Interior, U.S. Geological Survey .
- Granato, G.E., 2006. Kendall-Theil Robust Line (KTRLLine - version 1.0) - A visual basic program for calculating and graphing robust nonparametric estimates of linear-regression coefficients between two continuous variables: Techniques and Methods of the U.S. Geological Survey, book 4, chap. A7, 31 p.
- Hem, J.D., 1982. Conductance: a collective measure of dissolved ions: in R.A. Minear and L.A. Keith (eds.), Water Analysis, v. 1, part 1. Inorganic Species, New York, NY, Academic Press, p. 137-161.
- Hem, J.D., 1992. Study and interpretation of the chemical characteristics of natural water (3rd ed): U.S.Geological Survey Water-Supply Paper 2254, 263 p.
- Helsel D.R., and Hirsch, R.M., 2002, Statistical methods in water resources—Hydrologic analysis and interpretation: Techniques of Water-Resources Investigations of the U.S. Geological Survey, chap. A3, book 4, 510 p.

Lind, C.J., 1970. Specific conductance as a means of estimating ionic strength: U.S. Geological Survey Professional Paper 700-D, p. D272-280.

Miller, R.L., Bradford, W.L., and Peters, N.E., 1988, Specific conductance: theoretical considerations and application to analytical quality control: U.S. Geological Survey Water-Supply Paper 2311, 16 p.

Reham El-Korashey, 2009. Using Regression Analysis to Estimate Water Quality Constituents in Bahr El Baqar Drain, *Journal of Applied Sciences Research*, 5(8): 1067-1076.

Ryberg, K.R., 2006. Continuous Water-Quality Monitoring and Regression Analysis to Estimate Constituent Concentrations and Loads in The Red River of the North, Fargo, North Dakota, 2003-05, U.S. Geological Survey Water Resources Investigations Report 2006-5241, pp: 35.

Ryberg, K.R., 2007. Continuous water-quality monitoring and regression analysis to estimate constituent concentrations and loads in the Sheyenne River, North Dakota, 1980–2006: U.S. Geological Survey Scientific Investigations Report 2007–5153, pp: 22.

Singh, T. and Kalra, Y.P., 1975. Specific conductance method for in situ estimation of total dissolved solids. *J. Am. Water Works Assoc.*, v. 67, no. 2. p. 99.

Stevens, R. J., O'Bric C. J. and Carton O. T., 1995. Estimating nutrient content of animal slurries using electrical conductivity. *The Journal of Agricultural Science*, v. 125, no. 2, pp 233-238.

Thomas, A. G., 1986, Specific conductance as an indicator of total dissolved solids in cold dilute waters. *Hydrological Sciences Journal* v. 31, no. 1, p. 81-92.

Williams, W. D. 1966, Conductivity and the concentration of total dissolved solids in Australian lakes. *Australian Journal of Marine and Freshwater Research* 17, 169-76.