Data 201: Data Interpretation

Field Data Analysis

Luigi Limsiaco
000928887

Table of Contents

List of Tables

List of Figures

1. Field Data Organization and Descriptive Statistics

1.1. Population

The population for the self-directed study are the wadeable streams and small rivers in Alberta, since Stream Visual Assessment Protocol (SVAP) is only applicable for streams and rivers that are shallow enough to be sampled without the use of a boat [1].

1.2. Parameters

This study aims to determine if there are differences between water quality classification of a stream or river when assessed using SVAP and in-situ water testing methods. Specifically, the parameters included in this study are the following:

- **SVAP Score**: a numerical score representing the visual elements of the population. These elements would include the channel condition, hydrological alteration, riparian zone, bank stability, water appearance, nutrient enrichment, canopy cover, manure presence, instream fish cover, pools, riffles embeddedness, invertebrate habitat, salinity, and barriers to fish movement.
- **Water Quality Score**: The quantitative attribute of the population as a representation of the health of the river. The specific water quality parameters are pH, electrical conductivity, dissolved oxygen, and turbidity. These parameters will be compared and interpolated from water quality guidelines to be quantified as a single numerical value ranging from 0 to 10.

1.3. Samples

The samples included in this study are sections of the Galatea and Kananaskis Bridge Site. A total of eight sections or sampling locations were taken for this study. Each area are in different locations along the length of the rivers and will exhibit differences in visual elements. Although found in the same river, these sampling locations can also exhibit differences in water quality parameters thus a comparison between SVAP and WQ scores can be made.

1.4. Statistics

Since it is physically and financially near impossible for the group to gather SVAP and water quality data from all sections of wadeable streams and small rivers in Alberta, these data will only be gathered from the eight sampling locations to represent the whole population. The following statistics are included in this field study:

» **SVAP Score**: a cumulative score of the visual elements of the decided sections of the Galatea and Ribbon Creek. These scores will be gathered by the assessors using a checklist adapted from the Natural Resources Conservation Science (NRCS) [1].

» **Water Quality Score:** These are water chemistry parameters that will be measured using a ProQuatro Multiparameter Meter and a HACH Turbidimeter. The specific parameters measured are pH, electrical conductivity, dissolved oxygen, and turbidity. These water chemistry parameters will be scored from 0 - 10, using different water quality guidelines.

1.5. Sampling

This study will adopt a **convenience sampling** method due to the limitations of being only able to sample from two locations and certain sampling locations were selected due to ease of sampling and safety considerations. Since the statistical analysis required by the study design call for random sampling, these change in sampling method must be noted and considered when formulating conclusions.

1.6. Descriptive Statistics

The table below summarizes the descriptive statistics calculated for this study.

Table 1. Summary of Descriptive Statistics for SVAP and Water Quality Scores

Statistics	Mean	Median	Mode	Standard Deviation
SVAP Score	7.583	7.769	7.769	0.631
Water Quality Score	8.792	9.000	9.000	0.765

From table 1, it can be observed that Water Quality Score have higher values for central tendencies (mean, median, and mode) compared to SVAP Score. It can also be observed that the standard deviation of Water Quality Score is higher compared to the standard deviation of SVAP Score. This indicates that the data for Water Quality Score has a greater spread or degree of variation compared to SVAP Score.

1.7. Histogram

The following figures below presents the histograms generated for all the statistics gathered in this study.

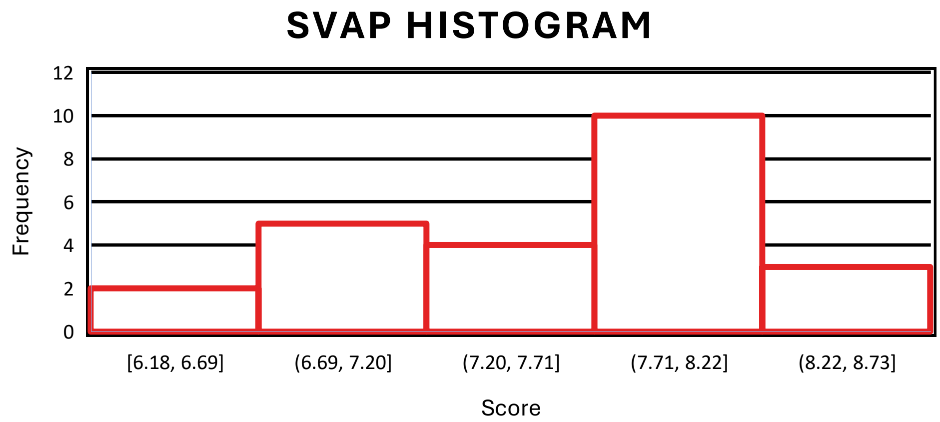

Figure 1. Histogram of SVAP Scores

Figure 1 shows a skewed distribution for SVAP results. The peak is found in the range (7.71, 8.22] and the distribution is unimodal and left-skewed.

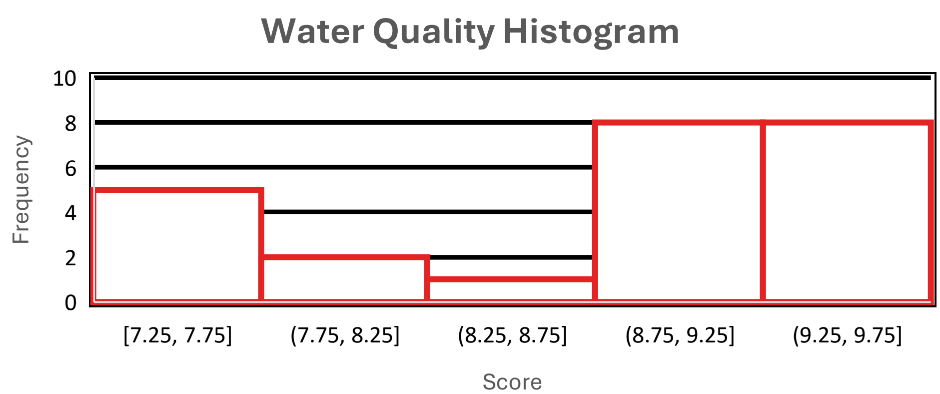

Figure 2. Histogram of Water Quality Scores

Figure 2 presents a histogram for water quality scores. The data shows a non-symmetric bimodal distribution with maximum values at ranges (8.75, 9.25] and (9.25, 9.75] and minimum value at range (8.25, 8.75].

1.8. Descriptive Chart

The bar chart below shows the mean values for both the SVAP Score and the Water Quality Score with corresponding standard deviation values.

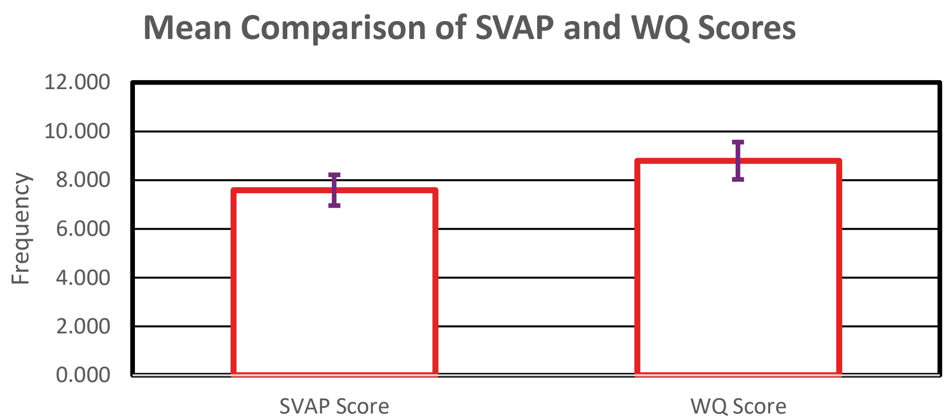

Figure 3. Mean Scores for SVAP and Water Quality with corresponding standard deviation as standard error bar

The mean values for SVAP Score and Water Quality are 7.58 and 8.59 respectively. The standard deviation values are 0.631 from SVAP Score and 0.765 for Water Quality Score.

2. Confidence Interval

2.1. Distribution

For both the SVAP Score and the Water Quality Score, **student's t distribution** was used for the sample mean distribution since figure 1 and figure 2 show that the data was non-normally distributed, and population standard deviation is unknown for both parameters.

2.2. Mean and Standard Deviation

The overall mean SVAP Score is **7.583** and the overall mean Water Quality Score is **8.792** for all eight sample locations. These values were computed using `AVERAGE()` function of excel as seen in the accompanying excel file.

The standard deviations of SVAP Score and Water Quality Score are **0.398** and **0.724** respectively. These values were computed using `STDEV.S()` function in excel.

2.3. Sample Size, n

The sample size used in this study is eight (8). This value was arbitrarily decided on by the group considering resource and time limitations during the conduct of the study.

2.4. Degrees of Freedom, DF

The degrees of freedom are computed using the equation below:

$$DF = n - 1$$

Where n is the sample size. For this study the degree of freedom is $DF = 8 - 1 = \mathbf{7}$

2.5. Standard Error

Standard error is computed the formula below:

$$Standard\ error = \frac{s}{\sqrt{n}}$$

Where $\sigma_{\bar{x}}$ is the standard deviation of the sample means, and n is the sample size.

For this study the standard error computation is computed using excel with the formula `=STDEV.S(range)/SQRT(COUNT(range))`. The computed standard errors are **0.141** and **0.256** for SVAP Score and Water Quality Score respectively.

2.6. t-Score

The t-score is computed using the excel formula `=T.INV.2T(probability, deg_freedom)`. The computed values are **2.365** for both SVAP and Water Quality with $\alpha = 0.05$ and $DF = 7$.

2.7. Confidence Interval

Confidence interval is computed using the formula below:

$$CI = \mu_{\bar{x}} \pm \left(t_{\left(\frac{\alpha}{2},\ n-1\right)} \times \frac{s}{\sqrt{n}} \right)$$

For SVAP Score:

$$CI = 7.583 \pm (2.365 \times 0.141)$$
$$CI = 7.583 \pm 0.333$$
$$\boldsymbol{CI} = (\mathbf{7.250}, \mathbf{7.916})$$

For WQ Score:

$$CI = 8.792 \pm (2.365 \times 0.256)$$
$$CI = 8.792 \pm 0.605$$
$$\boldsymbol{CI} = (\mathbf{8.187}, \mathbf{9.397})$$

Based on the computed confidence intervals, there is a 95% confidence that the mean population SVAP Score falls within the range (7.250, 7.916) and that the mean population WQ Score is within the range (7.674, 8.884).

3. Paired t-test

The self-directed study that was conducted does not compare between two different sites (Galatea and Kananaskis Bridge Site) but rather was a comparison between two different methods (SVAP and WQ) of quantifying the quality of a stream, thus t-test was conducted to determine if there are significant differences between the means obtained from using the different methods.

3.1. Paired Differences and Overall Mean Difference

The table below shows the mean SVAP scores and WQ Scores for each sample site and the corresponding difference (WQ Score - SVAP Score) in scores for each site.

Table 2. Mean SVAP and WQ Scores for each sampled site and corresponding paired differences

Sample	SVAP Score	WQP Score	Difference
Galatea SP-1	7.32	8.00	0.68
Galatea SP-2	7.71	9.00	1.29
Galatea SP-3	7.84	8.92	1.08
Galatea SP-4	8.05	9.00	0.95
Ribbon SP-1	6.96	9.00	2.04
Ribbon SP-2	7.94	9.50	1.56
Ribbon SP-3	7.13	9.50	2.37
Ribbon SP-4	7.72	7.42	-0.30

The overall mean difference ($\bar{x}_d$) is the mean of the Difference column of Table 2 and is computed using excels `AVERAGE()` function. The computed value for $\bar{x}_d$ is **1.208**.

3.2. Standard Deviation of Differences

The standard deviation of differences was calculated using `STDEV.S()` function of excel, where the inputs are the values from the Difference column of Table 2 and the resulting value is **0.829**.

3.3. Probability Distribution of Differences

A probability distribution was generated for the difference between the SVAP and WQ Scores for each sampling locations.

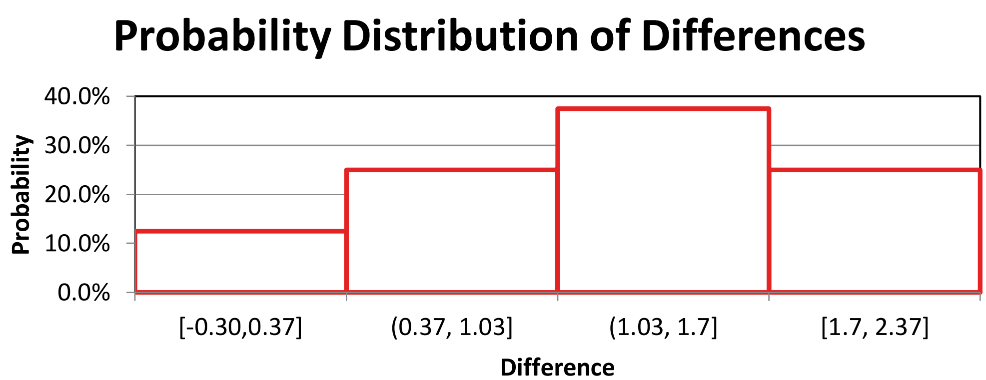

Figure 4. Probability Distribution of the Differences between WQ Score and SVAP Score

The probability distribution shown by Figure 4 shows a left skewed distribution for the differences between WQ and SVAP Score. However, it must be noted that the sample size is not sufficient to make conclusive decisions regarding the actual distribution of the differences.

3.4. Sample Size and Degree of Freedom

The sample size used for conducting the t-test is **8** and the degree of freedom used is **7**. Justifications and computations for the sample size and degree of freedom was discussed in sections 2.3 and 2.4.

3.5. Hypothesis

The null and alternative hypotheses used for the t-test of this self-directed study are:

H_0: There are no significant differences in the scores obtained using SVAP and WQ.

$(H_0: \mu_{SVAP} = \mu_{WQ})$

H_a: There are significant differences in the scores obtained using SVAP and WQ.

$(H_a: \mu_{SVAP} \neq \mu_{WQ})$

3.6. Hypothesis Test

Data Analysis tool of excel was used to conduct a paired t-test with two sample means with a hypothesized mean difference of 0 and an alpha of 0.05. The table below shows the output of the t-test provided by excel.

Table 3. Paired t-Test Output from Excel Data Analysis Tool for WQ Score and SVAP Score

Paired t-Test	***WQ Score***	***SVAP Score***
Mean	8.7917	7.5834
Variance	0.5238	0.1588
Observations	8	8
Pearson Correlation	*-0.0076*	
Hypothesized Mean Difference	0.0000	
df	7	
t Stat	*4.1230*	
P(T<=t) one-tail	*0.0022*	
t Critical one-tail	*1.8946*	
P(T<=t) two-tail	*0.0044*	

t Critical two-tail	*2.3646*	

3.7. Summary

Table 4 below shows the summary of necessary values obtained from the hypothesis test using paired t-test to determine if there are significant differences between the values obtained using two different methods, namely SVAP and WQ, to quantify the quality of a stream.

Table 4. Descriptives and paired t-test summary for the comparison of SVAP and WQ mean scores of 8 sampling locations at Galatea and Kananaskis Bridge site streams on given conditions of hypothesized mean difference = 0 and alpha = 0.05.

Parameter	Mean	Std Dev	Std Err	df	t	p
SVAP	7.583	0.398	0.141	7	*4.1230*	*0.0044*
WQ	8.792	0.724	0.256			

From the results of the t-test where $p < 0.05$, there is sufficient evidence, with 95% confidence, to reject the null hypothesis. There are significant differences between scores obtained using SVAP (M = 7.583, SD = 0.398) and WQ (M = 8.792, SD=0.724) for conditions: t(7) = *4.1230* and p =*0.0044*. It must be taken into consideration that the result of this analysis is based on an arbitrarily chosen sample size due to limitations stipulated in section 2.3. Power analysis must be conducted to determine the actual sample size required for the appropriate significance level, statistical power, and effect size necessary to have statistically valid conclusions.

Figure 5 below presents the mean scores for SVAP and WQ and their corresponding standard error bars.

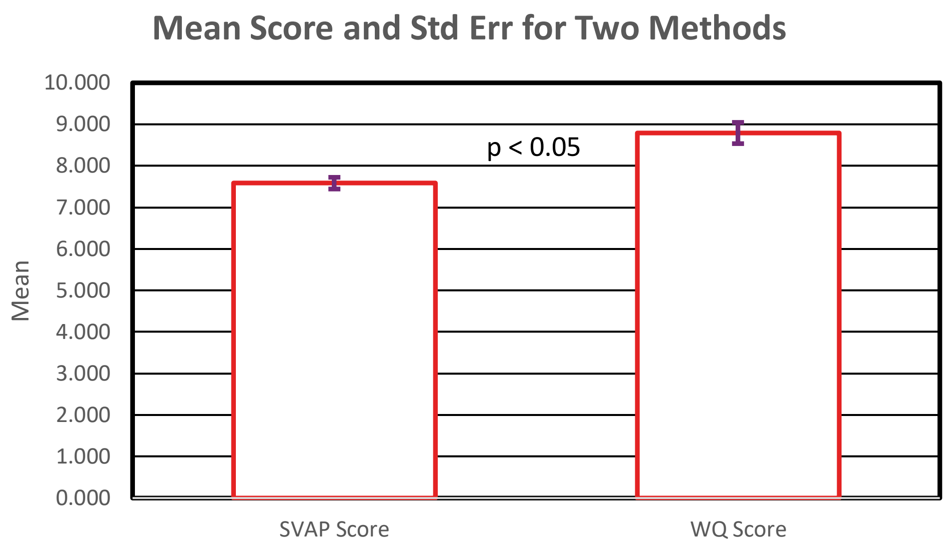

Figure 5. Mean scores for SVAP and WQ with standard error bars and t-test result

The figure above visually shows that WQ has a higher mean compared to SVAP. Since there is sufficient evidence to reject the null hypothesis, it can be accepted and that there are significant differences in the means of SVAP and WQ scores. Together with the observations from Figure 5, it can be inferred that WQ Score is significantly higher than SVAP Score. From the results obtained from the statistical analysis, it can be concluded that quantifying stream health using SVAP provides a significantly lower value compared to using WQ.

4. ANOVA

4.1. Hypothesis

One-Way Analysis of Variance (ANOVA) was used to test the following hypotheses:

H_0: There are no significant differences in the scores obtained using SVAP and WQ.

$(H_0: \mu_{SVAP} = \mu_{WQ})$

H_a: There are significant differences in the scores obtained using SVAP and WQ.

$(H_a: \mu_{SVAP} \neq \mu_{WQ})$

4.2. Hypothesis Test

Data Analysis Tool of Excel was used to conduct an ANOVA: Single Factor with an alpha of 0.05 to determine if there are significant differences between the mean score obtained using SVAP and the mean score obtained using WQ. The table below presents the result generated by excel.

Table 5. Mean scores analysis of variance (ANOVA) with alpha of 0.05 for SVAP and WQ of 8 sampling locations in Galatea and Kananaskis River Bridge Site

Summary

Groups	Count	Sum	Average	Variance
SVAP Score	8	60.667	7.583	0.159
WQ Score	8	70.333	8.792	0.524

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit

Between Groups	5.839326	1	5.839326	*17.108884*	*0.0010082*	*4.6001099*
Within Groups	4.778252	14	0.341304			
Total	**10.617578**	**15**				

4.3. Summary

Table 5 below shows the summary of necessary values obtained from hypothesis test using ANOVA: Single Factor to determine if there are significant differences between the values obtained using two different methods, namely SVAP and WQ, to quantify the quality of a stream.

Parameter	Mean	Std Dev	Std Err	df (Between Groups)	df (within groups)	F	p
SVAP	7.583	0.398	0.141	1	14	*17.109*	*0.001*
WQ	8.792	0.724	0.256				

The figure below visually presents the mean scores obtained from SVAP and WQ methods with corresponding standard error bars.

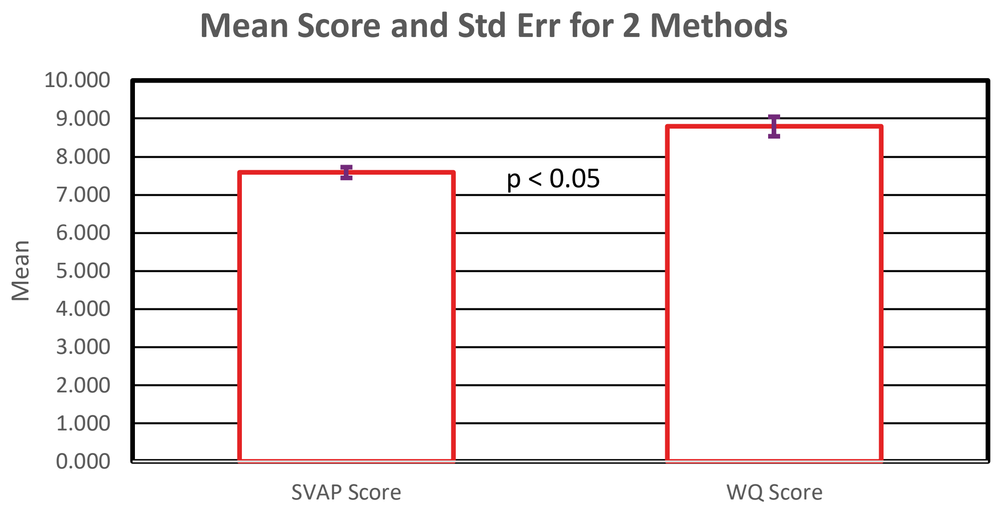

Figure 6. Mean scores with standard error bars obtained using SVAP and WQ methods for 8 sampling locations at Galatea and Kananaskis Bridge Site. ANOVA: Single Factor results shown.

Single Factor ANOVA shows a $p < 0.05$. This result suggests sufficient evidence, at 95% confidence, to reject the null hypothesis. Rejecting the null hypothesis imply that there are significant differences with the means obtained using SVAP (M = 7.583, SD = 0.398) and WQ (M = 8.792, SD = 0.724) methods on the conditions $F\ (1,\ 14) = 17.109$ and $p = 0.001$. Using observations from Figure 6, it can be concluded that WQ values are significantly higher than values obtained using SVAP. It must be noted that the sample size used in the design of this study was arbitrarily chosen and thus statistical power and effect size was assumed to be sufficient to make valid conclusions.

Since evidence show that there are significant differences between the scores obtained using SVAP and WQ and it can be visually observed from Figure 6 that WQ have a higher mean compared to SVAP, it is inferred that WQ is significantly higher than SVAP. This result follows the conclusion in the t-test that when quantifying stream health WQ will yield a significantly higher score compared to SVAP.

5. Regression and Correlation

5.1. Parameters

The parameters considered for this regression analysis are SVAP scores and dissolved oxygen (DO) values. The SVAP is the predictor while DO is considered the predicted. The aim of this regression analysis is to determine if SVAP analysis is a viable predictor of DO values of a section of a river being observed.

5.2. Parameter Table

The table below summarizes the x and y axes used to establish a scatter plot and conduct a regression analysis to observe relationships between SVAP score and DO.

Sample	Predictor	Predicted
	SVAP Score	DO, mg/L
Galatea SP-1	7.32	8.087
Galatea SP-2	7.71	8.113
Galatea SP-3	7.84	8.197
Galatea SP-4	8.05	8.200
Ribbon SP-1	6.96	8.087
Ribbon SP-2	7.94	8.147
Ribbon SP-3	7.13	8.160
Ribbon SP-4	7.72	8.150

5.3. Regression Analysis

Regression analysis was conducted using Excel’s Regression Data Analysis Tool. The table below presents the analysis results obtained.

Table 6. Summary tables of Regression result for SVAP Score and DO of 8 sampling locations at Galatea and Kananaskis Bridge Site

Summary Output

Regression Statistics	
Multiple R	*0.659007989*
R Square	*0.434291529*
Adjusted R Square	*0.340006784*
Standard Error	0.035947529
Observations	8

ANOVA

	df	SS	MS	F	Significance F
Regression	1	0.005952	0.005952	*4.606*	*0.076*
Residual	6	0.007753	0.001292		
Total	7	0.013706			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	7.588	0.259	29.310	1.045E-07	6.954	8.221	6.954	8.221
SVAP Score	0.073	0.034	2.146	0.076	-0.010	0.157	-0.010	0.157

Residual Output

Observation	Predicted DO	Residuals	Standard Residuals
1	8.123	-0.036	-1.092
2	8.152	-0.038	-1.155
3	8.161	0.036	1.070
4	8.177	0.023	0.694
5	8.097	-0.010	-0.307
6	8.169	-0.022	-0.660

7	8.109	0.051	1.521
8	8.152	-0.002	-0.070

5.4. Summary

Regression analysis was conducted to assess the relationship between SVAP score and DO readings for 8 sampling locations at the Galatea and Kananaskis Bridge Site streams. Summary of results are shown in the table below.

Table 7. Summary table for regression analysis result for SVAP vs DO

Model	r	r-critical	R^2	Adjusted R^2	p
SVAP vs DO	*0.659*	*0.707*	*0.434*	*0.34*	*0.076*

The results suggest a moderate positive correlation $(r = 0.659)$ between the variables and there is no statistically significant correlation $(p > 0.05, |r| < r_{critical})$ between them with 43.43% of the variation in DO is accounted for in the variation in SVAP score and when sample size is adjusted 34% of variation is still accounted for.

The figure below visually presents the relationship between SVAP and DO through a scatter plot. It can be observed from the graph that there is indeed an increasing trend for DO as SVAP increases, indicating that SVAP can moderately predict DO values.

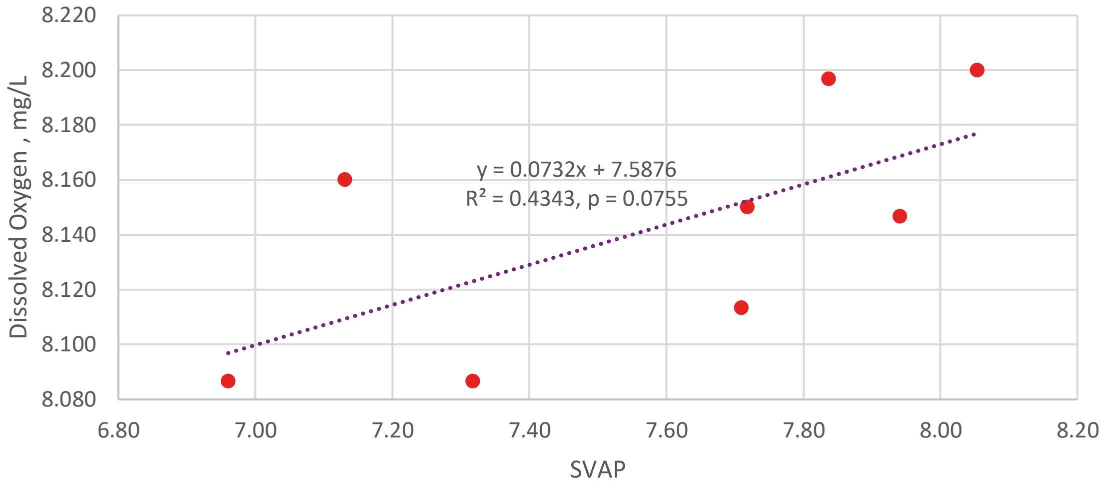

Figure 7. SVAP vs DO scatter plot with best fit line, coefficient of determination and regression analysis result

6. Overall Summary

Inferential statistics were conducted for the data gathered for the self-directed study and showed the following results:

- Parameter estimate of the population mean scores for SVAP and WQ were obtained to be (7.250, 7.916) and (8.187, 9.397) respectively. Values were obtained with 95% confidence and assumed a Student's t distribution of the sample means.
- Hypothesis test using paired samples t-test shows sufficient evidence to reject the null hypothesis $(p < 0.05)$, thus suggesting that there are significant differences in the means obtained using SVAP and WQ.
- Hypothesis test using One-way ANOVA shows sufficient evidence to reject the null hypothesis $(p < 0.05)$ indicating that significant differences can also be observed from the scores obtained using SVAP and WQ.
- Regression analysis to assess the relationship between SVAP Score and DO levels showed moderate positive correlation between SVAP and DO $(r = 0.659)$ however evidence also show that correlation between the variables is not statistically significant $(p > 0.05, |r| < r_{critical})$.

As expounded previously, statistical results are based on an arbitrarily decided sample size thus statistical power and effect size were only assumed to be sufficient. Result may vary and conclusions may be different when appropriate sample size is used for the analysis.

7. Conclusions

The purpose of this study is to determine if SVAP analysis is a viable method of quantifying stream and river health. Water quality (WQ) analysis was used as the baseline metric for the stream and river health and SVAP was compared with this parameter to assess if there are significant difference in the results of the two methods. It was observed that WQ score showed a higher average score compared to SVAP however the distribution of WQ also showed a higher degree of spread. This indicates that WQ provides a higher but more inconsistent quantification of stream and river health compared to SVAP. Statistical tests, paired samples t-test and one-way ANOVA, determined that there are significant differences in the scores obtained using SVAP and WQ. Although, it must be noted that both methods presented similar qualitative results that the streams and rivers are of good (score = (7.0, 8.9)) to excellent (score = (9.0,10.0)) health, quantitative values are significantly different from each other which could suggest presence of extraneous variable in which further investigations are recommended. Regression analysis to evaluate relationships between SVAP score and DO suggests that high SVAP scores also corresponds to a high DO level and low SVAP score can indicate low DO levels. However, SVAP scores alone is not sufficient to provide a quantitative value for DO levels. It is recommended that the sample size is increased to account for statistical power and effect size and to further investigate other water quality parameters to include in the WQ score, the method for establishing the WQ score and adjustments to SVAP elements to be a better evaluator of the designated population.

8. References

[1] U.S. Department of Agriculture, Natural Resources Conservation Service, *Stream Visual Assessment Protocol Version 2*. Washington, DC: USDA, 2009. Accessed: Nov. 13, 2024. [Online]. Available: https://www.nrcs.usda.gov/sites/default/files/2022-10/Stream-Visual-Assessment-Protocol-Version-2.pdf