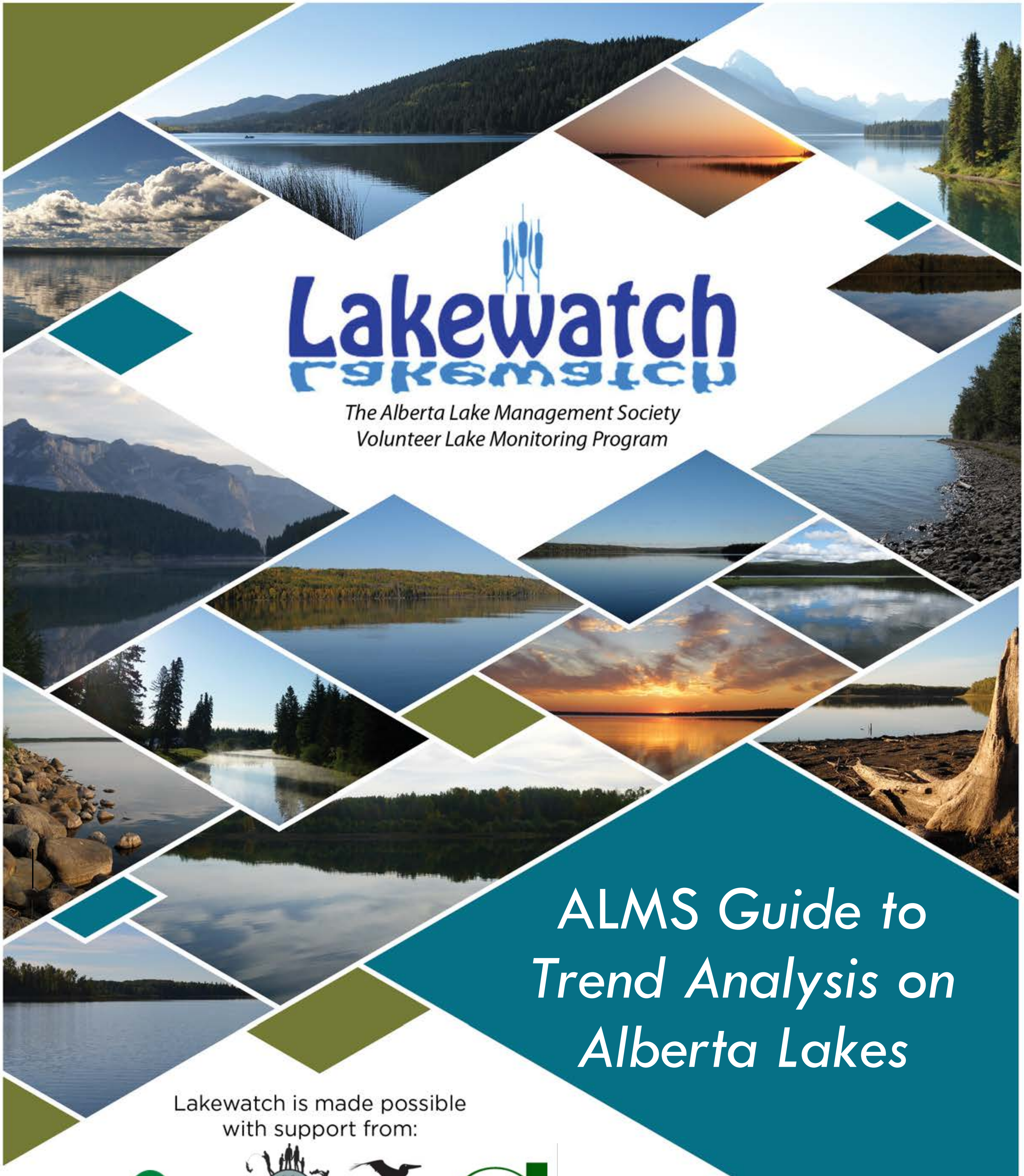


Lakewatch

The Alberta Lake Management Society
Volunteer Lake Monitoring Program

ALMS Guide to Trend Analysis on Alberta Lakes

Lakewatch is made possible
with support from:

Cargill Helping the world thrive
Jackfish Lake
Pigeon Lake Watershed Association
LICA Lakeland Industry & Community Association
Red Deer River Watershed Alliance
Beaver River Watershed Alliance
Alberta Environment and Parks

INTRODUCTION

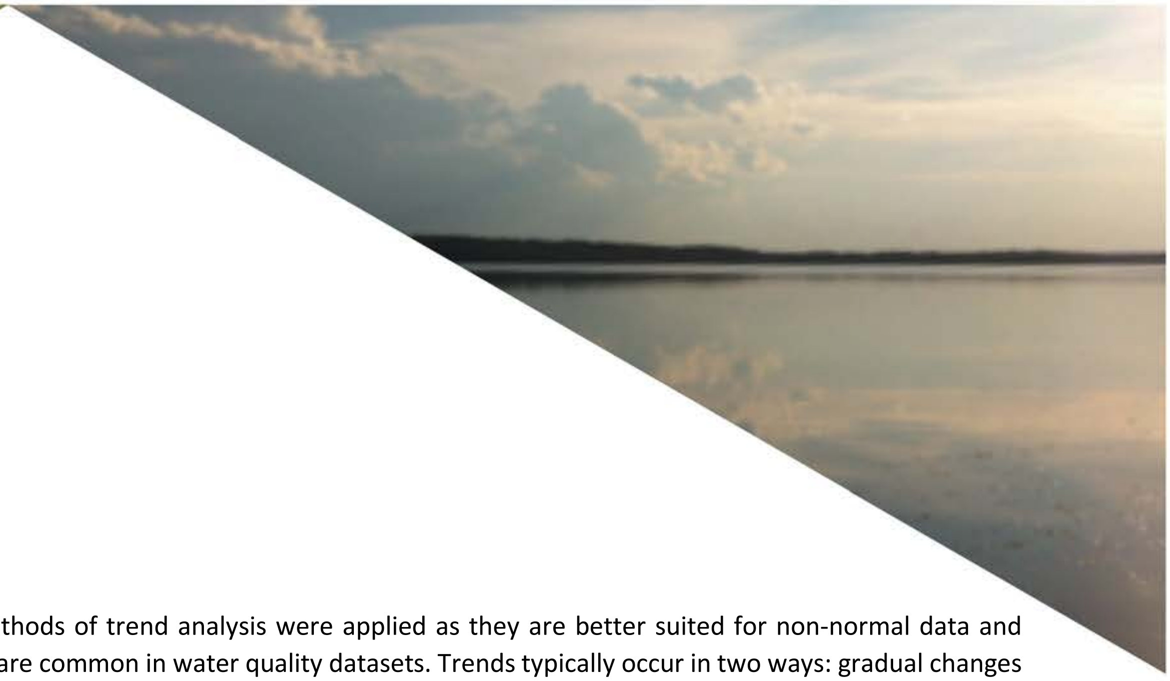
This report summarizes the methods used by the Alberta Lake Management Society (ALMS) for detecting trends in water quality data. In 2018, ALMS began running trend analysis on lakes with 10 or more years of data for the parameters chlorophyll-*a*, total phosphorus (TP), total dissolved solids (TDS) and Secchi depth.

While trends in Alberta Lakes have been assessed on various occasions, there is no standardized method for determining which test is appropriate for the data. This report will summarize methods of trend analysis using results from Pigeon Lake as a case study (Table 1) and propose a methodological way of choosing a trend analysis.

Version last updated April 18, 2018.

ACKNOWLEDGEMENTS

This report was prepared by ALMS Program Coordinator Laura Redmond with guidance from Executive Director Bradley Peter. The ALMS 2018 Technical Advisory Committee (Dorte Koster, Mike Christensen, Rebecca Tasko, Jana Tondu and Francine Forrest) provided valuable feedback for the development of this report. Thank-you to Richard Casey and Arin MacFarlane-Dyer for their comments and discussion.



METHODS

Non-parametric methods of trend analysis were applied as they are better suited for non-normal data and missing data which are common in water quality datasets. Trends typically occur in two ways: gradual changes in a single direction (monotonic) or a shift at a point in time (step-trend). In cases where there is a temporal break in data that is greater than one third of the total data, then a non-parametric step trend such as the Wilcoxon Rank Sum test should be used before and after the data gap (Figure 4, Helsel and Hirsch 2002). Additionally, step-trend analysis can be useful in studies where changes before and after an impact are being assessed.

To test for gradual changes, Mann Kendall tests are used with non-normal data to assess monotonic trends in a dataset (a non-parametric linear regression). This test is robust to missing values which often occur in water quality data (Helsel and Hirsch 2002). The Tau value is an output of the test and is a measure of strength of the monotonic relationship between the parameter and time on a scale from -1 to 1, with -1 being a strong negative trend, and 1 being a strong positive trend. A Tau value close to 0 represents only a weak trend. The p-value indicates the significance of the trend to 90 or 95% (p = 0.10 or 0.05, respectively).

Trend analysis must take into account other variation that may be contributing to the actual trend. Time series data with high frequency sampling often exhibit autocorrelation or dependence in time sequence that violates the assumption of independence. However, when sampling is less regular and spaced further apart in time, the correlation with time is less important (Helsel and Hirsch 2002). A test for autocorrelation must confirm the independence of data before conducting a non-corrected trend analysis. Otherwise, a correction factor must be considered.

Seasonal variation is another measure of noise that may reduce the power of an actual trend. The seasonal Kendall test estimates the presence of monotonic (unidirectional) trends across individual seasons (months) and is summed to give an overall trend over time using the package EnvStats for R (Miller 2013). For lakes that have multiple samplings in a single month, the value closest to the 15th day of the month may be used for analysis (Casey 2011). To check for seasonality of data, the Kruskal-Wallis Test can be used. If there is no difference among months, then a non-seasonal trend method may be more appropriate. Additionally, if monotonic trends differ among seasons (months) then a seasonal Kendall test may not be appropriate. To test for heterogeneity of monotonic trends across seasons we used the EnvStats package. If the p-value for van Belle Hughs heterogeneity of monotonic trends test is significant (< 0.05), then the null hypothesis that seasonal trends are monotonic can be rejected (van Belle and Hughes 1984) and we can assume that monotonic trends are different across seasons. A non intrablock method such as Mann Kendall may be most appropriate if monotonic trends across seasons are different (heterogeneus; Figure 4). No outliers were removed unless values were incorrect due to laboratory or technical error (i.e. sample exceeded hold time).

CASE STUDY: PIGEON LAKE

Differences in temporal sampling and sampling effort across years can greatly skew trend results. Changes in sampling methods or timing may cause shifts in data that could be incorrectly interpreted as trends. It is therefore important to address this variation to ensure consistent methods and sampling effort over time.

Example: Pigeon Lake Monthly Sampling Variation

In Pigeon Lake, monitoring has occurred in months between May and October. However, sampling in May and October only occurred prior to 1990. Since sampling in May and October was not consistent in recent years, we removed these months because they may create a false trend. The Mann Kendall result became non-significant once May and October values were remove for chlorophyll-a (Table 1).

As well, different sampling efforts across years can greatly affect the results of a trend analysis. For example, in 1988 and 2013, Pigeon Lake was monitored intensively multiple times per month (Figure 1). One way to remove this variability is to compare only one monthly value from June to September. In our analysis, the value closest to the 15^{th} of the month was chosen to represent the monthly value. In Pigeon Lake data, when we reduced sampling effort to include one sample per month from June to July (Figure 2), the Tau value changed from 0.07 to 0.19 (Table 1). This method also reduces the effect of autocorrelation across the data, because the sampling dates are less closely spaced thereby reducing the dependence of data points (Helsel and Hirsch 2002).

In Pigeon Lake, the Mann Kendall test was used because although the data demonstrated seasonality, there was heterogeneity in the monotonic trends across months. Therefore, a Seasonal Kendall would not be an appropriate fit for this data because the trends across months could cancel out. Given that yearly sampling effort was not equal, we reduced the number of samples to include one sample per month (Table 1). For more details on results, see the 2017 Pigeon Lake LakeWatch report at: alms.ca/reports

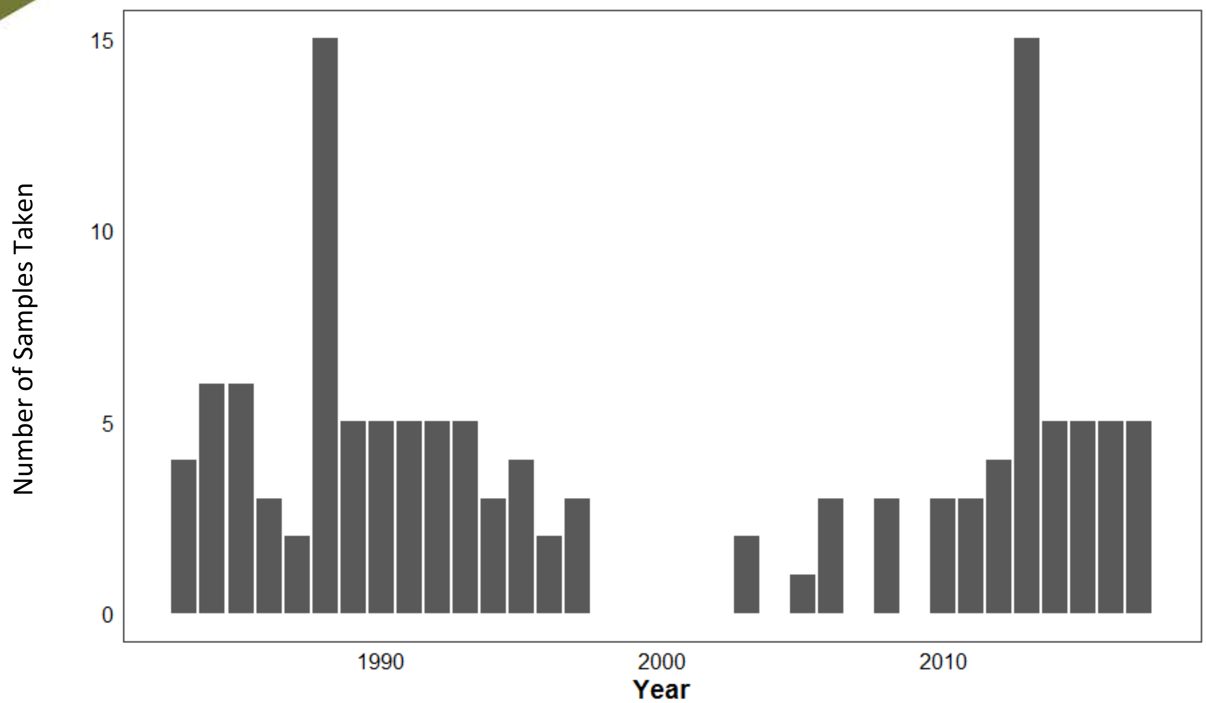


Figure 1: Histogram of sampling effort with raw Pigeon lake data. Includes May-Oct and all sampling dates from 1983-2017.

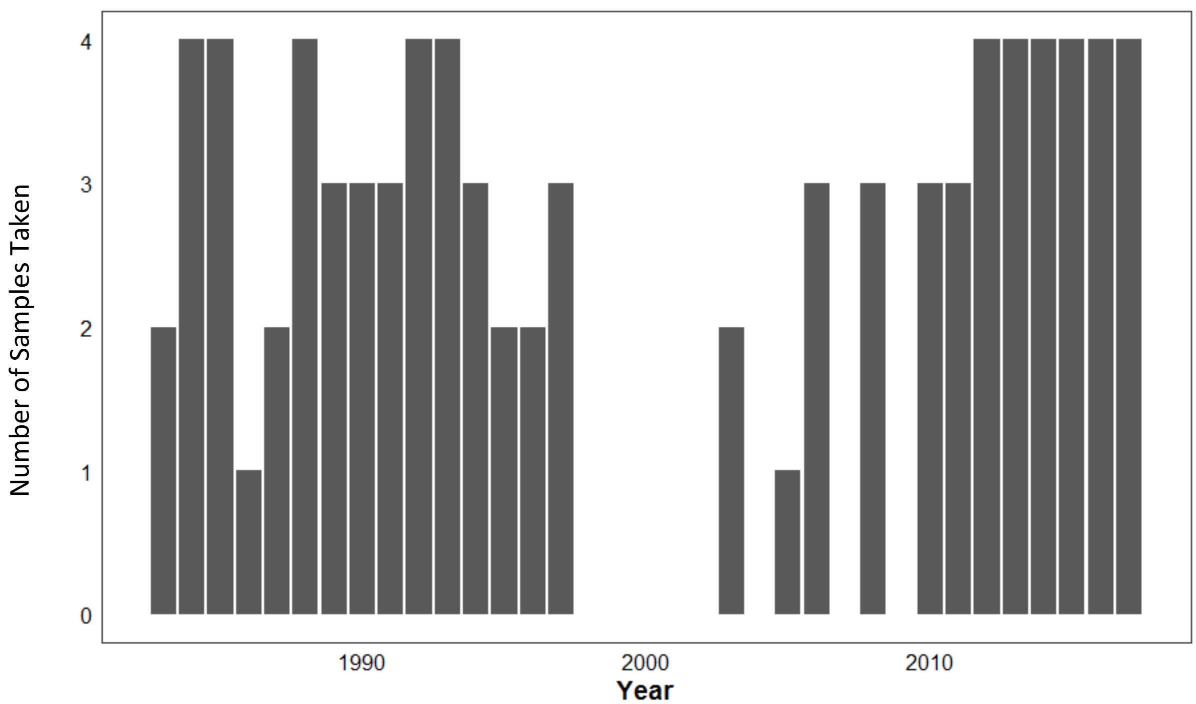


Figure 2: Histogram of sampling effort with one sampling per month from June-September between 1983 and 2017.

Table 1: Trend analysis methods comparison with all data versus monthly data which corrects for different sampling efforts across years. The value closest to the 15th day of the month is chosen to represent the monthly value. The two variations of Mann Kendall results are also compared to Seasonal Kendall results. Results are displayed for June-Sep and May-Oct to account for May and October values being biased towards older data (pre-2000). The column in blue is the selected method of trend analysis for Pigeon Lake. All other columns are for comparison of methods.

	Mann Kendall				**Seasonal Kendall**	
Data	**All Data Points**		**One Sample Per Month**		**Monthly**	
Months	***Jun-Sep***	***May-Oct***	***Jun-Sep***	***May-Oct***	***Jun-Sep***	***May-Oct***
Chlorophyll-a	*Tau=0.073*	*Tau=0.11*	*Tau=0.12*	*Tau=0.16*	*Tau= 0.10*	*Tau= 0.031*
	Slope=0.00040	*Slope=0.00054*	*Slope=0.00058*	*Slope=0.00079*	*Slope= 0.15*	*Slope=0.12*
	N=112	*N=127*	*N=83*	*N=98*	*N=83*	*N=98*
	Z=1.1	*Z=1.8*	*Z=1.6*	*Z=2.4*	*Z= 1.5*	*Z=1.3*
	P=0.257	***P=0.0711****	***P=0.114***	***P=0.0189*****	*P (z Trend)= 0.132*	*P(Z trend)= 0.203*
TP	*Tau=0.042*	*Tau=0.080*	*Tau=0.12*	*Tau=0.17*	*Tau=0.083*	*Tau= 0.089*
	Slope=0.00020	*Slope=0.00044*	*Slope=0.00068*	*Slope=0.00098*	*Slope=0.20*	*Slope= 0.21*
	N=112	*N=127*	*N=83*	*N=98*	*N=83*	*N=98*
	Z=0.6	*Z=1.3*	*Z=1.6*	*Z=2.4*	*Z=1.5*	*Z=1.5*
	P=0.518	*P=0.185*	***P=0.116***	***P=0.0161*****	*P(Z trend)= 0.147*	*P(Z trend)= 0.125*

* Significant $p < 0.1$ **Significant $p < 0.05$

OTHER METHODS

Visualizing the data is important to finding trends but variability may obscure an obvious trend. A smoothing line such as LOWESS (Locally Weighted Scatterplot Smoothing) can reveal monotonic trends in data plotted as a boxplot.

Trend analysis may not be appropriate for validating the observations of lake residents or for describing the behaviour of the data. For example, in some cases, variability may change significantly over time without the presence of monotonic trends. In these situations, further statistical analysis may be required to describe the data. To test for heterogeneity of variance across two groups of dates (i.e. Pigeon Lake pre and post-2003), a Levene Test is used. A Levene Test assesses the equality of variances for two or more groups. If the data plotted as a boxplot seems to show a change in variance, a Levene Test can be used to show if the variances are unequal (heteroscedastic). For example, Pigeon lake chlorophyll-*a* shows an increase in variance post-2003, and a Levene test shows this trend is significant (p = 0.0085; Figure 3).

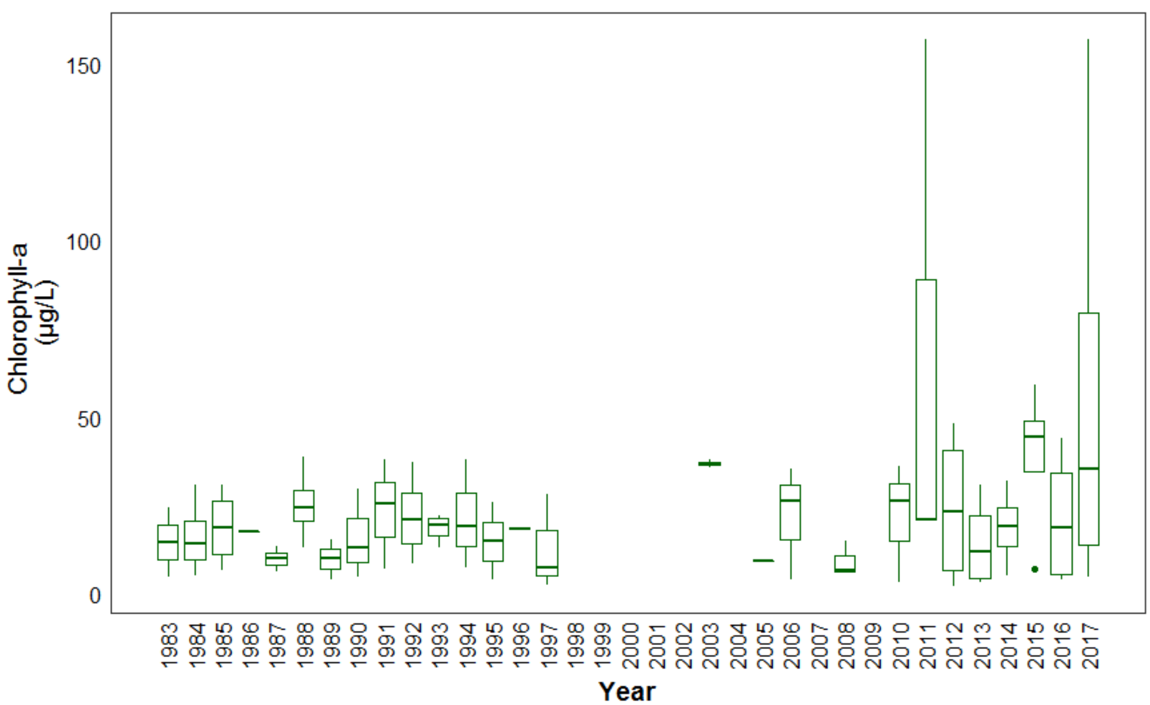
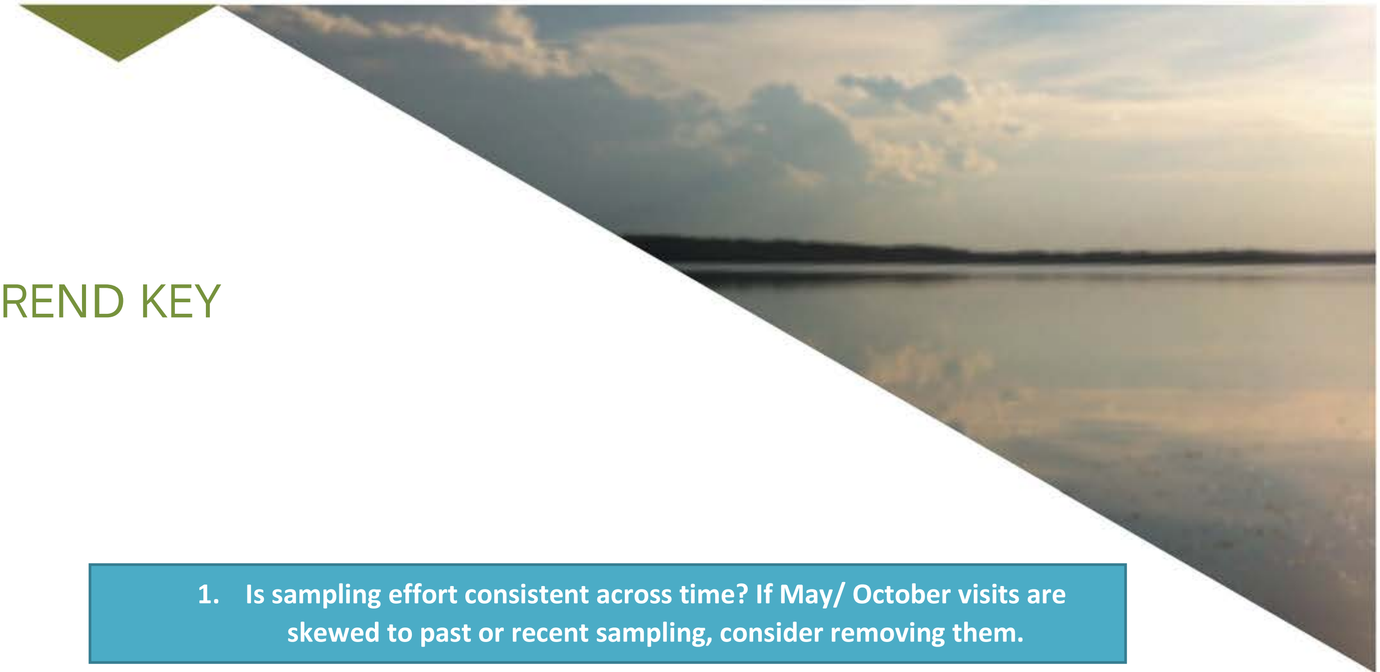


Figure 3: Monthly chlorophyll-*a* concentrations measured between June and September over the long term sampling dates between 1983 and 2017 (n = 83). Note the change in variance in more recent sampling years.



TREND KEY

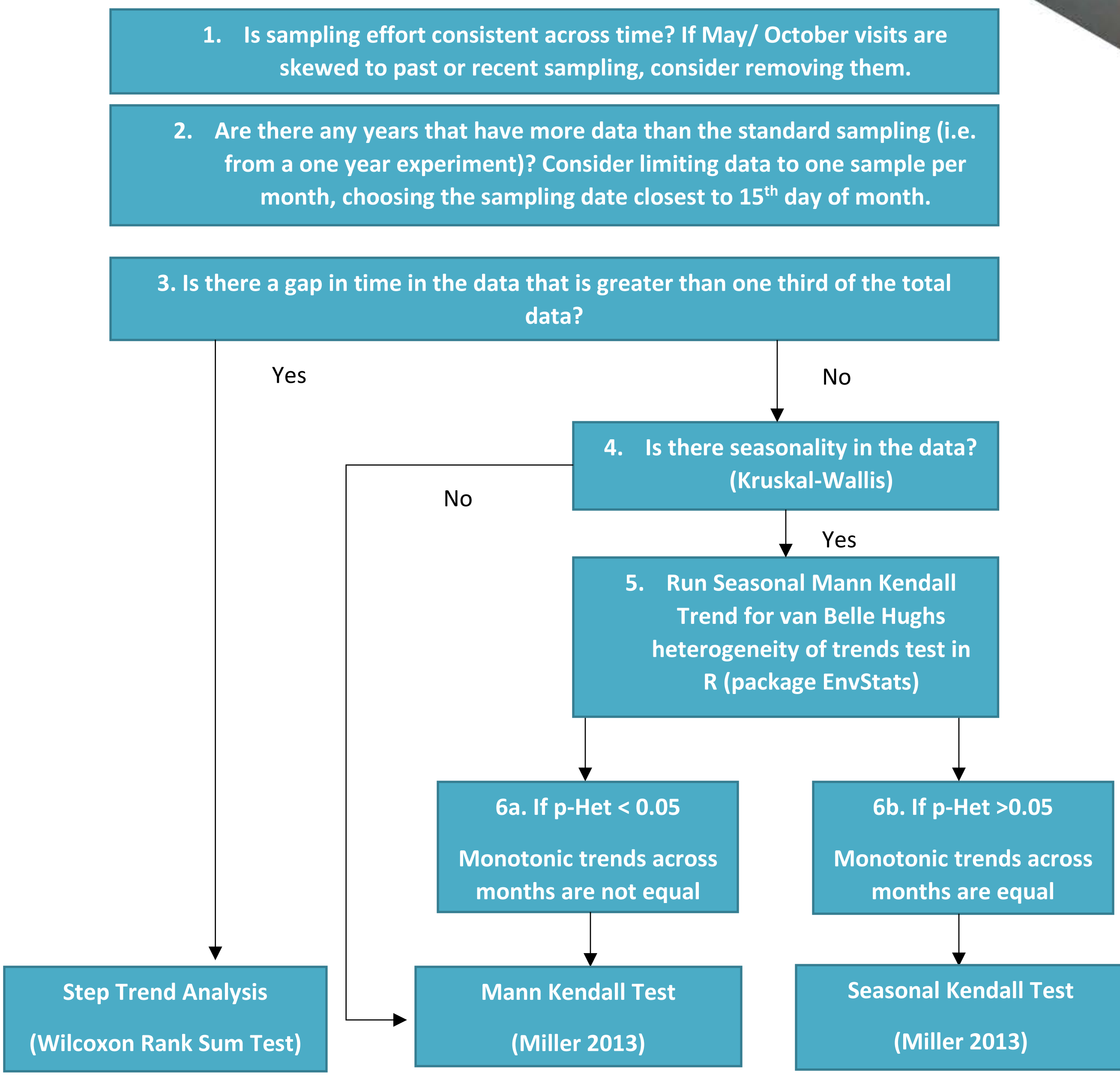


Figure 4: Decision tree for trend analysis selection using ALMS methods.

REFERENCES

Casey R (2011). Water quality conditions and long-term trends in Alberta lakes. Alberta Environment and Water, Edmonton.

Helsel DR and RM Hirsch (2002). Statistical Methods in Water Resources Techniques of Water Resources Investigations, Book 4, chapter A3. U.S. Geological Survey.

Miller SP (2013). EnvStats: An R package for environmental statistics. Springer, New York. ISBN 978-1-4614-8455-4, <URL: http://springer.com>

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

van Belle G and JP Hughes (1984). Nonparametric tests for trend in water quality: Water Resources Research 20: 127-136.