

SOP 11: Five-Year Report (FQ_H Creation)

Version 1

Revision History Log:

Prev. Version #	Revision Date	Updated By	Changes Made	New Version #
-	Sept 2011	C. Sergeant	Initial version	1

Contents

	Page
1.0 Summary	118
2.0 Five-Year Report Format and Content	118
3.0 Contents of Results Section	119
3.1 Summary Figures	120
3.2 Time-Series Trend and Parameter Correlation Analysis	125
3.3 Comparison to Other Data Sets	127
3.4 Analytical Software	127
4.0 Monitoring Program Recommendations	128
5.0 Tips for Improving Readability of Results	128
6.0 Submission, Validation, and Dissemination of the Deliverable	128
7.0 Literature Cited	130

1.0 Summary

Beginning in 2016 and every five years thereafter, the Project Leader will conduct a detailed analysis of long-term status and trends in water quality data from the three network parks (SITK, GLBA, and KLGO). Additional regional datasets may be integrated with SEAN water quality monitoring to form comparisons over broader temporal and spatial scales. If the Project Leader requires additional expertise in order to complete the report, SEAN will acquire external assistance.

Important questions to be examined in five-year reports include:

- What is the seasonal variation of each water quality parameter?
- Do long-term positive or negative trends in water quality parameters exist?
- Are long-term water quality trends consistent across parks?
- What are reference conditions for water quality parameters in each park?
- Does long-term data suggest an emerging environmental shift from past conditions?
- Do water quality trends correlate to streamflow and weather patterns?

2.0 Five-Year Report Format and Content

The five-year report will be formatted and published as a NPS Natural Resource Technical Report (<http://www.nature.nps.gov/publications/NRPM/>). The goal of the report is to expand upon annual results by describing and interpreting long-term monitoring status and trends in a broader temporal and spatial context. Generally, the Introduction section will be similar to annual reports, while the Methods and Results sections will be supplemented with explanations of long-term and regional trend analyses commonly used with water quality datasets. The five-year report will include similar tables and figures from annual reports (SOP 10), thus, this SOP describes additional reporting beyond the scope of annual reports.

1. **Executive Summary:** Intended for park superintendents and managers; review program purpose and design; summarize Results and Discussion
2. **Introduction:** Present the monitoring program context, goals, and objectives
3. **Methods:** Describe study areas, field and analytical techniques, annual deviations from sampling prescribed in protocols, and the sources for external regional datasets
4. **Results:** Present tables and figures of annual and historical results, summarize long- and short-term regional or local trends, and describe anomalous or notable system-specific observations
5. **Discussion:** Generalize results in the context of monitoring program objectives and long-term questions, provide interpretation for long- and short-term trends or anomalies, briefly describe expected future sampling, what has gone well and should continue, and present recommendations and future considerations for new monitoring work

Acknowledgement of field crew, logistical supporters, document reviewers, contracted reporting assistance, etc. is included before the introduction.

3.0 Contents of Results Section

In the SEAN, four to five water quality parameters are measured hourly for five to six months each sampling season, producing thousands of lines of water quality data each year that must be summarized in an informative and concise manner. To achieve this goal and maintain long-term relevance, the results of five-year reports will be directly linked to the objectives described in the protocol narrative and the questions listed in Section 1.0 of this SOP. Below is the list of essential data reporting that must accompany each objective and the related, long-term questions. The proceeding sub-sections describe each of these report components in more detail. Each year, if the Project Leader discovers additional informative results through other exploratory analyses, reporting will expand beyond the essential components listed below.

- **Objectives 1 and 2:** Track the status and trends of each core water quality parameter; describe the timing and magnitude of seasonal and annual variation of each core water quality parameter; as data accumulates over time, place short-term variation in the context of observed long-term trends
 - **Related long-term questions:** What is the seasonal variation of each water quality parameter? Do long-term positive or negative trends in water quality parameters exist?

Report components: Temperature and turbidity mean/minimum/maximum summary (Figure SOP 11.1), annual mean temperature trends (Figure SOP 11.2), monthly box plots for conductivity, dissolved oxygen, and pH (Figure SOP 11.3), scatter plot matrix of pairwise parameter comparisons (Figure SOP 11.4), and Mann-Kendall tests for trend (Figure SOP 11.5).
- **Objective 3:** Evaluate whether state and/or federal water quality standards are met or exceeded (in the context of the five-year report, this objective will be interpreted more broadly, considering whether any monitored rivers are trending toward water quality regulatory thresholds or if there is an observed shift in water quality that signifies potential future exceedance)
 - **Related long-term questions:** What are reference conditions for water quality parameters in each park? Does long-term data suggest an emerging environmental shift from past conditions?
Report component: period of record and 90% confidence interval (Table SOP 11.1), comparison of state water quality criteria to measured parameter means and ranges (if no standards were exceeded, no comparison table is required)

Most results will be summarized in graphical form. Brief narratives summarizing general trends or notable results will support tables and figures. Graphic summaries and analyses presented in this SOP combine ideas from SEAN staff, other NPS scientists, Helsel and Frans (2006), Helsel and Hirsch (2002), and Hirsch et al. (1982).

Tables will be formatted in Excel and inserted into Word documents using Paste Special/Picture (Windows Metafile) commands; figures are produced and formatted in several common software

packages (Excel, SigmaPlot, or R). When possible, results from all three rivers (Salmon, Taiya, and Indian; potentially more depending on future monitoring program expansion) will be consolidated into a single panel.

3.1 Summary Figures

Temperature and turbidity are expected to display the highest degree of variability throughout the sampling season and will undergo more detailed initial analyses than specific conductance, dissolved oxygen concentration, and pH. If notable variability or seasonality exists in the latter three parameters, further analyses will be included in the five-year report.

3.1.1 Time Series

Average daily time-series data are plotted for water temperature at all sampling stations and for turbidity in the Taiya River (Figure SOP 11.1). To summarize seasonal trends of water temperature and turbidity, each time series includes the five-year daily mean, mean minimum, and mean maximum lines. Depending on the breadth of data, certain anomalous years or shorter time periods of interest are graphed separately from the average summaries, described in the Results text, or included in a detailed appendix.

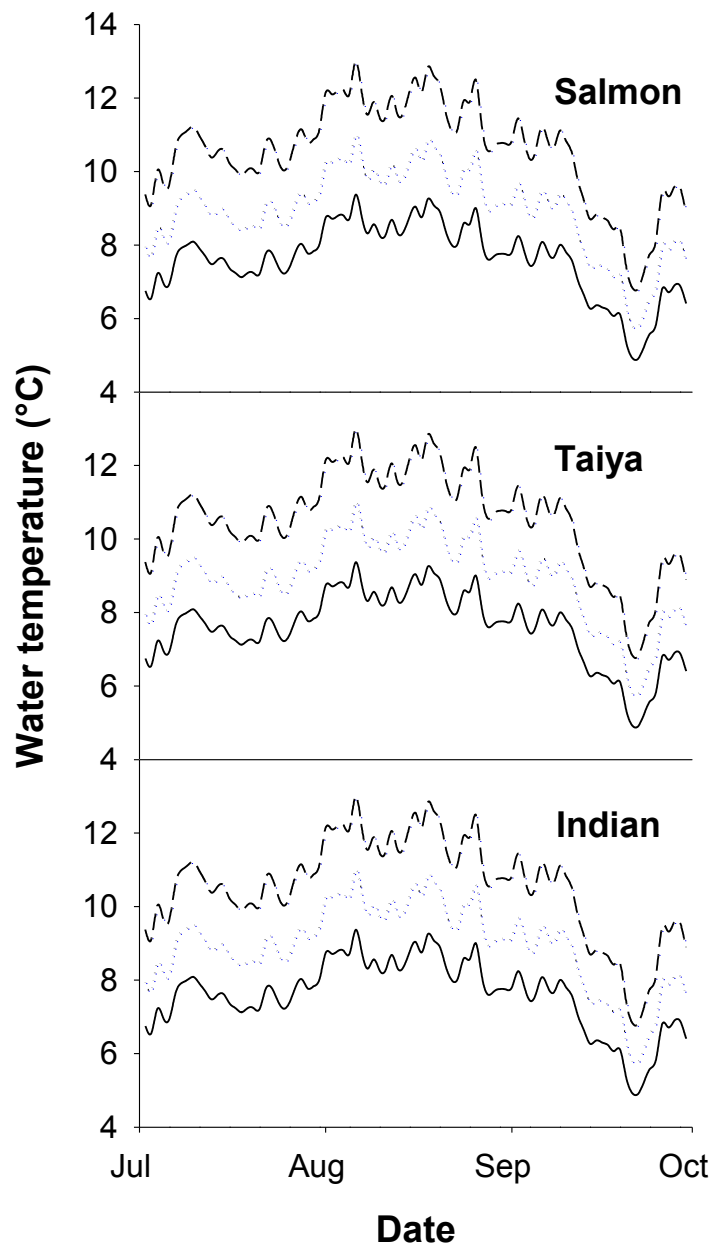


Figure SOP 11.1. Five-year daily mean (dotted), mean minimum (solid), and mean maximum (dashed) water temperatures for artificial data in the Salmon, Taiya, and Indian rivers.

Annual water temperature trends are summarized by scatter plots, where each point represents an annual mean temperature calculated from average daily mean values (Figure SOP 11.2). These graphs illustrate warming, cooling, or stable trends in annual water temperature over time. Indices such as these are important general indicators of climate shift and annual growth and survival conditions for all levels of organisms from bacteria to fish. If significant trends exist based on correlation coefficients (Kendall's tau; Section 3.2.2), regression analyses are an appropriate addition to this figure. When excessive data gaps exist, the Project Leader must justify the inclusion or exclusion of a seasonal mean. If considerable gaps occur through five years, monthly averages may be a more appropriate summary method.

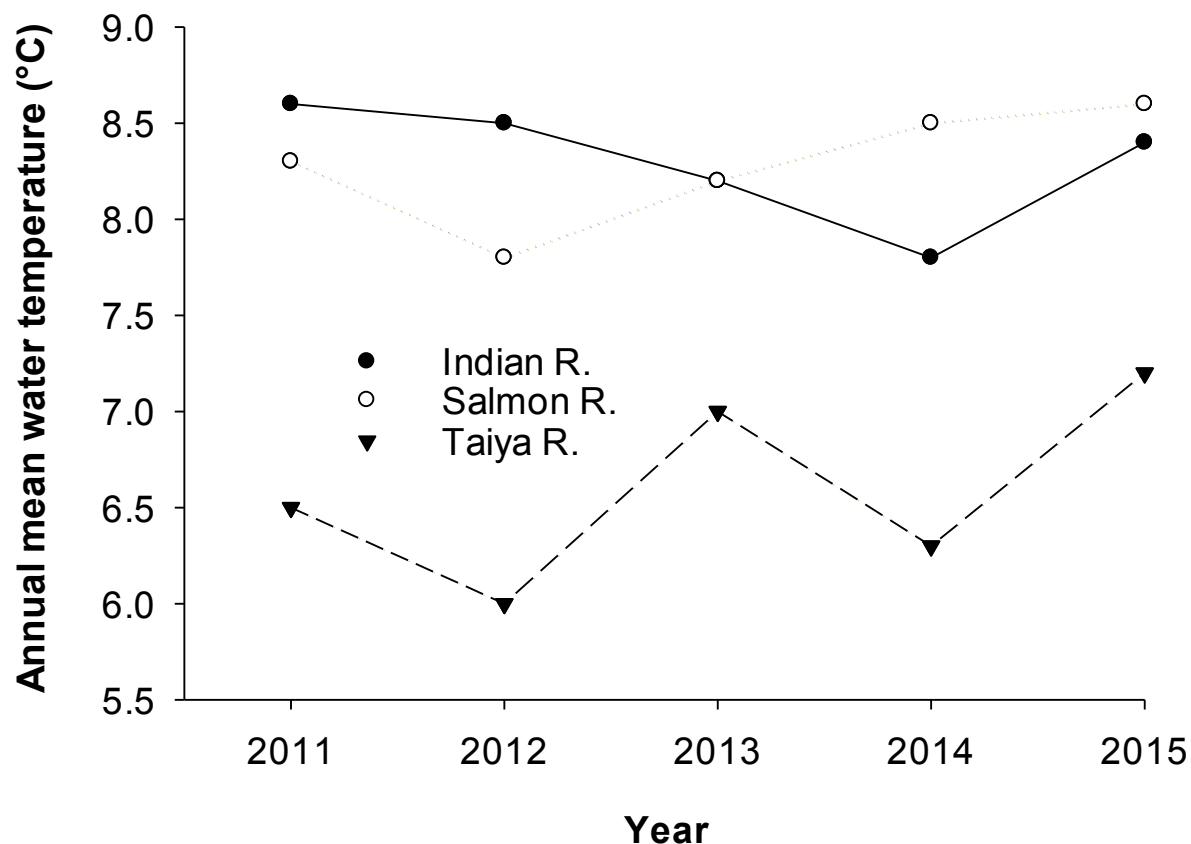


Figure SOP 11.2. Annual mean temperature using artificial data for the Indian, Salmon, and Taiya rivers.

To further describe long-term warming and cooling trends, the date of annual peak water temperature and turbidity are graphed by year as a scatter plot in a manner similar to figure SOP 11.2. In this case, the X-axis will again be the year, while the Y-axis will increase by day of year. If displaying results for three rivers becomes too cluttered for interpretation, a separate panel for each river is displayed.

Monthly summaries of each water quality parameter are described by box plots (Figure SOP 11.3). Monthly values will be averaged over all years combined.

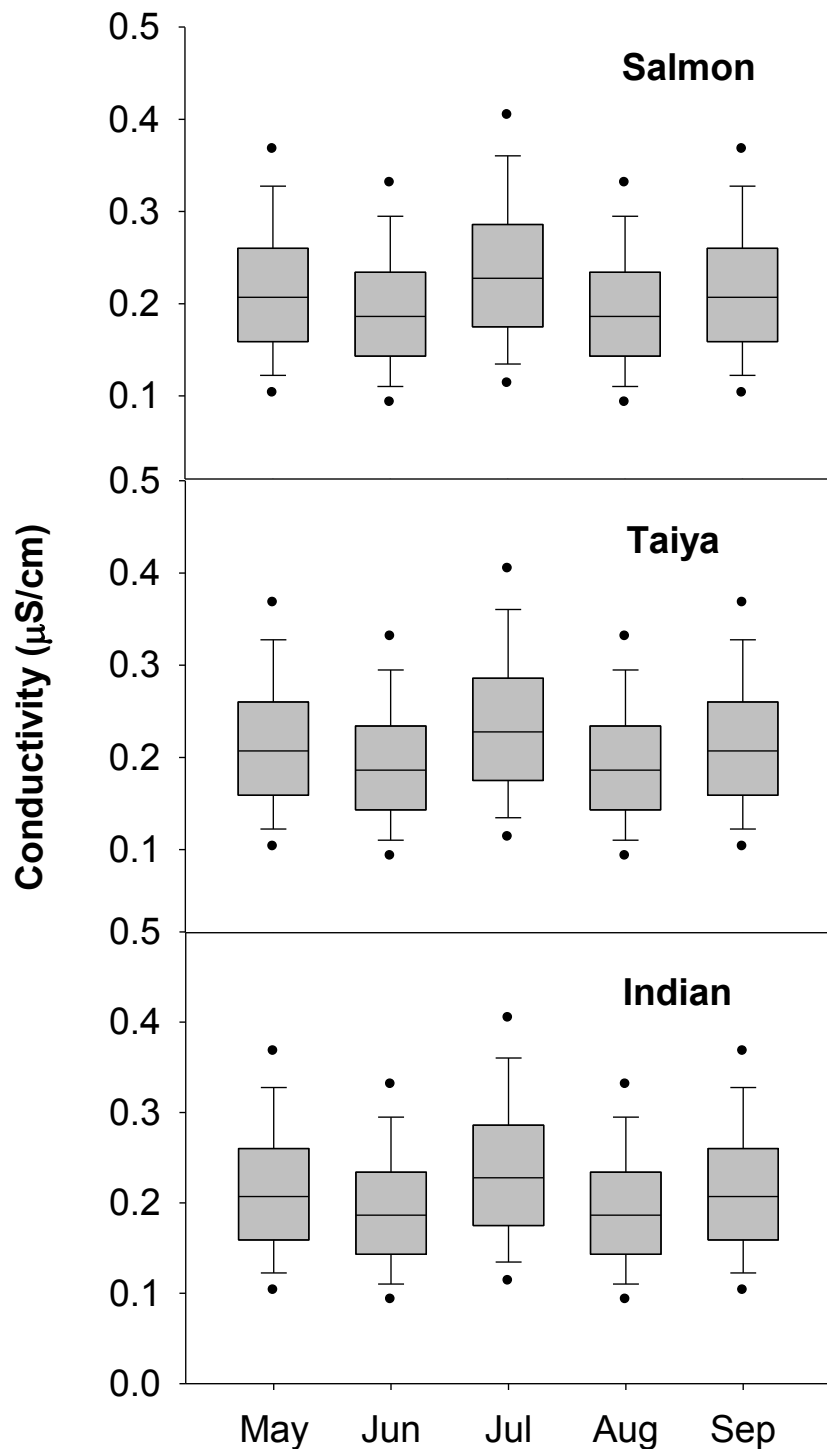


Figure SOP 11.3. Box plot example summarizing monthly water quality over a five-year period using artificial data for the Indian, Salmon, and Taiya rivers. Lines within the boxes represent median values, lines bounding the lower and upper portion of the boxes represent 25th and 75th percentiles, lower and upper whiskers represent 10th and 90th percentiles, and single points represent 5th and 95th percentiles.

3.1.2 Reference Conditions

Reference conditions for each water quality parameter are established by computing 90% confidence intervals around the mean value of each parameter. Regardless of the shape of the data distribution, a symmetric confidence interval around the mean based on a normal data distribution is appropriate due to the high number of observations from hourly sampling over five seasons (Central Limit Theorem). This presentation is especially useful when comparing data to established state water quality regulations or discussing the typical range of water quality conditions at each monitoring station (Table SOP 11.1).

Table SOP 11.1. Period of record and five-year 90% confidence intervals for each river and water quality parameter.

River	Parameter	Period of Record	No. observations	Confidence interval summary			
				Lower 90% confidence limit	Mean	Upper 90% confidence limit	Confidence interval range
Salmon	Conductivity ($\mu\text{S}/\text{cm}$)	1 May 2012 to 30 September 2016	683	0.08	0.25	0.42	0.34
	Dissolved Oxygen (mg/L)						
	Dissolved Oxygen (% Sat)						
	pH						
	Temperature ($^{\circ}\text{C}$)						
Taiya	Conductivity ($\mu\text{S}/\text{cm}$)						
	Dissolved Oxygen (mg/L)						
	Dissolved Oxygen (% Sat)						
	pH						
	Temperature ($^{\circ}\text{C}$)						
Indian	Conductivity ($\mu\text{S}/\text{cm}$)						
	Dissolved Oxygen (mg/L)						
	Dissolved Oxygen (% Sat)						
	pH						
	Temperature ($^{\circ}\text{C}$)						
	Turbidity (NTUs)						

3.1.3 Correlation between Two Parameters

Often, two water quality parameters can be correlated, but the magnitude or pattern of correlation differs across systems depending on the local dynamics of environmental forces such as glacial input, geology, and climate. Scatter plot matrices for each river will be a useful tool for quickly displaying and diagnosing potential correlation between any two water quality parameters (Figure SOP 11.4).

Several documented examples illustrate the importance of pair-wise comparisons for elucidating the mechanisms underlying seasonal trends in water quality parameters. Among other environmental factors, dissolved oxygen concentration can change with water temperature. In less productive river systems, water temperature may be the main factor affecting DO concentration, while no response is seen in relation to turbidity or pH. The relationship between pH and conductivity can be influenced by glacial or snow meltwater and streambed geology. As glacially-influenced rivers such as the Salmon experience spring run-off, turbidity will be expected to increase while conductivity decreases. This relationship can be confirmed with a pair-wise comparison. Some relationships may be seasonal. The Project Leader will examine relationships during discrete time periods to determine whether pair-wise comparisons should be summarized on a seasonal or annual basis. In general, observing relationships such as the examples listed here will lead to an increased understanding of how the unique environments of each river system affect the observed values of water quality parameters.

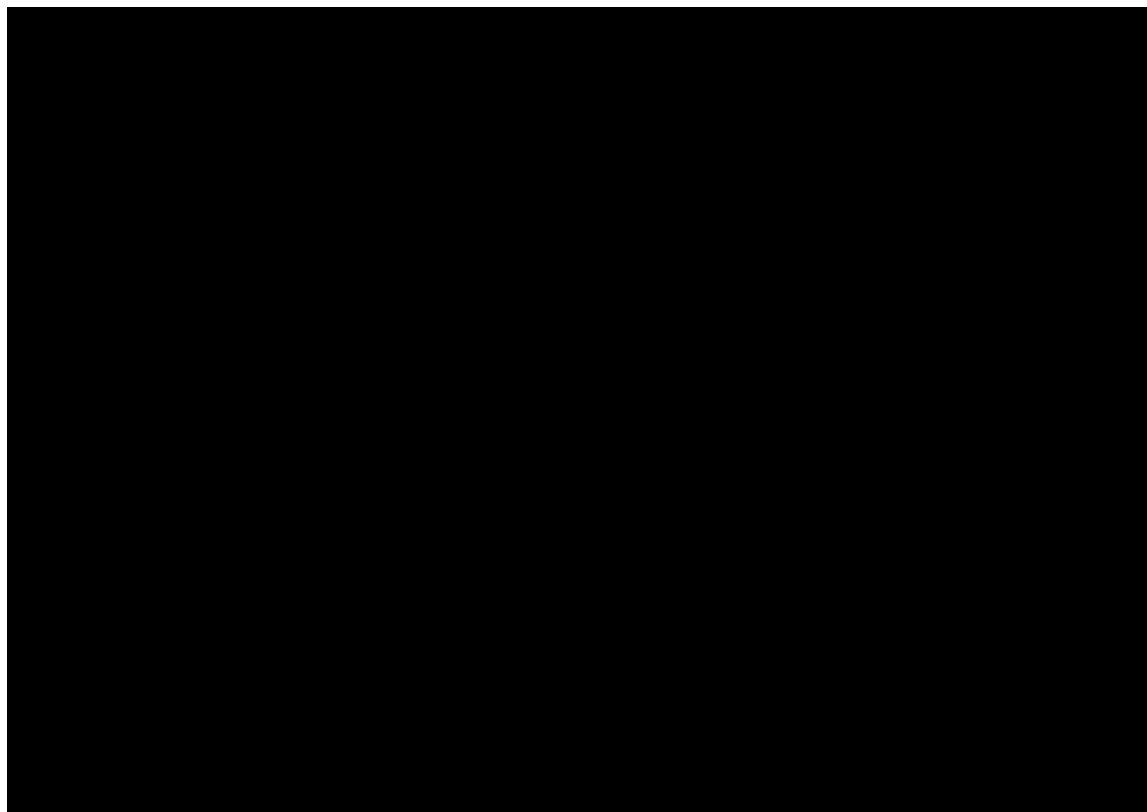


Figure SOP 11.4. Example scatter plot matrix for water quality data from Helsel and Hirsch (2002). Each point represents an individual measurement. The bottom row uses histograms to illustrate the distribution of values for each water quality parameter.

3.2 Time-Series Trend and Parameter Correlation Analysis

In addition to the summary figures described above, the suite of Mann-Kendall tests, associated correlation coefficient, and slope estimator are used to quantify trends. Five-year report authors may need to implement further analyses based on preliminary results and visually inspecting summary figures. In all cases, though, final report components are limited to those that support the clearest interpretations and do not move outside the scope of the monitoring objectives set forth in section 1.0 in this SOP and section 1.6 of the protocol narrative.

3.2.1 Time-Series Trends

One of the most commonly accepted and informative water quality analyses is the Mann-Kendall test for trend. The non-parametric Mann-Kendall test evaluates whether a dependent Y-axis variable (e.g., temperature) changes over time in a positive or negative direction regardless of the function's shape. The test is robust for water quality data that are typically not normally distributed.

A subset of the Man-Kendall test is the seasonal Kendall test, in which monthly medians for a selected parameter are calculated for each monitoring year (equating to six medians per year [May-October] for the freshwater water quality program). The test separates the effects of seasonality from the overall annual trend to determine if there is a single direction (monotonic) trend over time. If a significant trend is detected, the seasonal Kendall slope estimator (also

known as the Sen or Theil-Sen slope) describes the magnitude of the trend. Trends can be linear or non-linear over the time series. One caution for test result interpretation is that cyclical trends cannot be discerned. Mann-Kendall tests are discussed in great detail in Helsel and Hirsch (2002), Helsel and Frans (2006), and Hirsch et al. (1982).

Serial correlation, typically positive autocorrelation, may be an issue with water quality time-series data, and p-values resulting from Mann-Kendall tests can be corrected for autocorrelation. However, time series under 10 years in length are not usually corrected for autocorrelation (Helsel and Frans 2006). The presence of autocorrelation is tested by calculating Kendall's tau on the residuals of adjacent data points. Significant values ($p < 0.10$) represent autocorrelation. This process is described in Helsel and Frans (2006; Section 9.5.4).

Trend test results from Mann-Kendall (within the same month over multiple years) or seasonal Kendall (over multiple years) tests will be summarized in a single figure (Figure SOP 11.5).

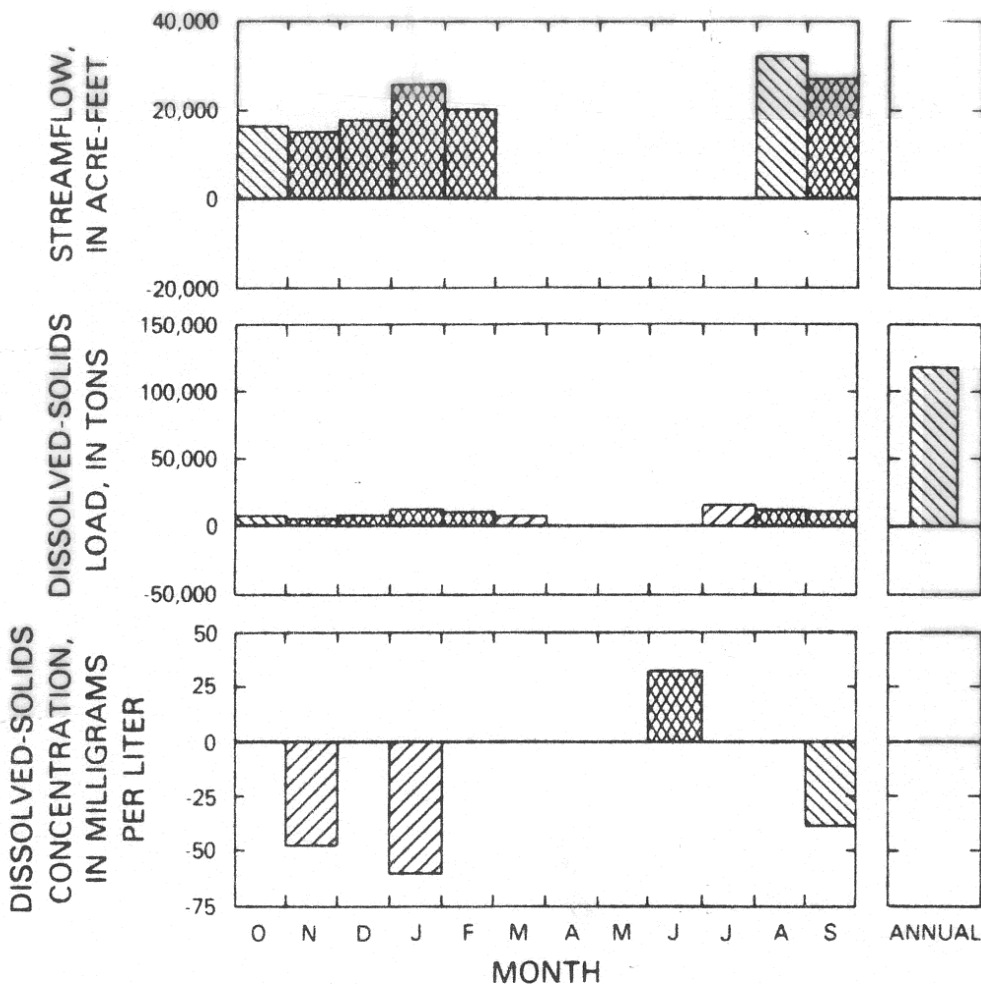


Figure SOP 11.5. An example of graphically summarizing results from monthly Mann-Kendall and seasonal Kendall tests from Helsel and Hirsch (2002). Bars represent the magnitude of positive or negative trends for each water quality parameter.

3.2.2 Scatter Plot Correlation

For each scatter plot in figures SOP 11.2 and 11.4, the correlation coefficient Kendall's tau is calculated with a significance level of $\alpha = 0.10$. Non-parametric correlation coefficients such as Kendall's tau are preferable over parametric correlation coefficients for describing water quality relationships because these data are usually not normally distributed and tau is not sensitive to outliers. A tau p-value less than 0.10 signifies a measurable positive or negative relationship between two variables, regardless of the function's shape. Linear or local regression will aid in understanding the shapes of relationships between monitored water quality parameters in each river, but time-consuming regression procedures are only necessary when visually obvious correlation patterns are present and supported by Kendall's tau values.

3.3 Comparison to Other Data Sets

Often, water quality measurements are closely linked to other environmental parameters such as weather and streamflow. At a minimum, when available, streamflow in each monitored river will be compared to core water quality parameter values. Parameters such as snow pack and air temperature will likely affect the observed patterns in water quality data, as well, and will be included in the five-year report when significant correlations are found. As more parameters are compared to water quality data, the Project Leader is encouraged to explore multivariate modeling and multiple regression techniques to help determine complex patterns among several climate and water quality variables (e.g., effect of snow pack and precipitation on conductivity).

Streamflow data sets:

- Salmon River does not currently have a stream gage, but flow peaks and valleys can be discerned from the Montana Creek gage (USGS 15052800) in Juneau. While Montana Creek has a smaller drainage area, its long-term data set (1965 to the present) and relative proximity to the Salmon River makes it the nearest available and relevant gage for comparison. If available, Falls Creek discharge data collected by the Gustavus Hydroelectric Company is a more accurate proxy for Salmon River discharge.
- Taiya River gage near Skagway (USGS 15056210; partially funded by NPS)
- Indian River gage in Sitka (currently maintained and operated by NPS)

Weather station data sets:

- GLBA fuel dock
- NPS weather station data in KLGO (Taiya River Bridge, Klondike Highway Summit, Sheep Camp, Chilkoot Pass, Moore Creek, Gold Lake)
- FAA airport data from SGY, GST, and SIT
- NOAA/NWS Cooperative Observer 15-minute Precipitation Network
- USDA-NRCS Snowpack Telemetry (SNOTEL) Network (including Moore Creek north of Skagway)

3.4 Analytical Software

The R package 'wq' will be used for calculating Kendall's tau, slope estimators (mannKen), and seasonal Kendall tests for trend (SeaKen). R can efficiently produce scatter plot matrices (Figure SOP 11.4). Data will be formatted in Excel spreadsheets before importing into R. Over time, the

Project Leader will develop an appendix of useful R code and Excel data formatting tips to improve the efficiency of future reporting.

4.0 Monitoring Program Recommendations

In addition to interpreting results, the Discussion section of each five-year report will include descriptions of the pros and cons of the current monitoring approach, recommended corrective actions for the long-term program, and future considerations for expanding the program or conducting alternate analyses.

5.0 Tips for Improving Readability of Results

1. **Monitoring program context:** To retain continuity across annual reports, structure results and narrative in the context of the monitoring program's objectives. Results that may be interesting but are unrelated to program objectives are placed in discrete sub-sections within the Results section but separate from the main results. If many interesting results occur, or a separate research or monitoring program needs to be discussed in relation to water quality results, an appendix will be added.
2. **Parallel structure:** When listing results in graphs, tables, or narrative, consistently refer to rivers in alphabetical order by park: Salmon (GLBA), Taiya (KLGO), and Indian (SITK). When referring to combined water quality figures, list conductivity, dissolved oxygen, and pH in alphabetical order. Repeating these patterns throughout the annual report will allow the reader to develop a reading pattern and comprehend more quickly.
3. **Black and white:** Use only black and white or grayscale figures. This avoids problems with color-blindness or non-color printers.

6.0 Submission, Validation, and Dissemination of the Deliverable

The final version of FQ_H is reviewed by the SEAN Program Manager and generated as a PDF file by the Project Leader, who delivers it to the Data Manager for the validation checks defined in Appendix D. After the valid report is certified by the Project Leader, the Data Manager installs the final PDF in the auxiliary repository, creates a web link for it, updates the Deliverable Product Tracking grid, and forwards the final FQ_H to the NPS IRMA Data Store.

1. *Program Manager Tasks*
 1. Review the final Word report document.
 2. Apply any essential corrections.
 3. Forward to Project Leader.
2. *Project Leader Tasks*
 1. Generate a PDF file from the original Word Document.
 2. Submit both the PDF and DOCX files via email attachment to the Data Manager for validation. Specify in the message body it is deliverable FQ_H, as defined in protocol FQ-2013.1.

3. Data Manager Tasks

1. On receipt of the submission, assign the next formal Submission Number to this file, as found in the master Submission_Log table.
 1. Use the “Update Submission Log” web tool at the data management site:
http://165.83.57.239/0_submission_update.aspx.
 2. Complete Submission_Log details up through the Submission_Date column.
2. Save the attached files to the staging area for validation at:
\\inpglbafs03\data\SEAN_Data\Staging\FQ\FQ_H\nnnn where “nnnn” is the assigned submission number.
3. Rename the files to FQ_H-yyyy.PDF and FQ_H-yyyy.DOCX, where “yyyy” is the year of report publication.
4. Validate the two files according to current criteria.
5. Record validation summary data in the Submission_Log using the web tool.
6. If submission fails mandatory criteria, reply with a “failure email” that includes:
 1. The submission number assigned.
 2. The deliverable ID.
 3. The protocol ID.
 4. Documentation listing all the specific mandatory criteria failed.
7. If submission passes mandatory criteria, reply with a “success email” that includes:
 1. The submission number assigned.
 2. The deliverable ID.
 3. The protocol ID.
 4. Documentation listing any specific optional criteria failed.
 5. Request to certify deliverable as complete.

4. Project Leader Tasks

1. On receipt of a failure email:
 1. Make corrections so the deliverable meets mandatory criteria.
 2. Make another submission with corrected deliverable candidate.
2. On receipt of a success email:
 1. After reviewing any optional criteria violations, if the deliverable is acceptable reply with a “certification email” stating the deliverable is certified and may be disseminated.
 2. If it is unacceptable:
 1. Reply with a “withdrawal email,” stating the deliverable is withdrawn.
 2. Take remedial action to obtain a corrected deliverable.
 3. Restart the process from the beginning.

5. Data Manager Tasks

1. On receipt of a withdrawal email:
 1. Mark the withdrawal in the Submission_Log’s Status column using the web tool.

2. *File the email in the Data Manager's email folder named "SEAN Data Management / FQ".*
3. *Terminate the process.*
2. On receipt of a certification email:
 1. *File the email in the Data Manager's email folder named "SEAN Data Management / FQ".*
 2. *Due to the nature of these data, NO sensitive information is in the deliverable.*
 3. *Copy the two submitted files to test environment at:*
[\\inpglbafs03\data\SEAN_Data_Mgmt\Web_Sites\SEAN-Test\AuxRep\FQ\FQ_H.](\\inpglbafs03\data\SEAN_Data_Mgmt\Web_Sites\SEAN-Test\AuxRep\FQ\FQ_H.)
 4. *Create web page link to the new PDF, but not the DOCX*
 1. **Mark the certification in the Status column in Submission_Log using the web tool.**
 2. **Propagate from test to production environment.**
 3. **Update the annual deliverable tracking spreadsheet showing date of completion for FQ_H**
 5. *Verify that the final document is in the NPS IRMA Data Store and is marked for public access.*

7.0 Literature Cited

- Helsel, D. R., and L. M. Frans. 2006. Regional Kendall test for trend. *Environmental Science and Technology* **40**(13):4066–4073.
- Helsel, D. R., and R. M. Hirsch. 2002. Statistical methods in water resources. Chapter A3 in *Techniques of water-resource investigations of the United States Geological Survey, Book 4, Hydrologic analysis and interpretation*. Available from <http://water.usgs.gov/pubs/twri/twri4a3/> (accessed 1 September 2011).
- Hirsch, R. M, J. R. Slack, and R. A. Smith. 1982. Techniques of trend analysis for monthly water quality data. *Water Resources Research* **18**(1):107–121.