

Standard Operating Procedure (SOP) #35

Data Analysis and Reporting

Version 1.02 (June 3, 2024)

Change History

New Version #	Revision Date	Author	Changes Made	Reasons for Change	Previous Version #
1.01	5/5/2020	Kim Weisenborn	Updated "References" header to "Literature Cited", updated figure and SOP references	Making SOPs consistent with updated content	1.0
1.02	6/3/2024	Kim Weisenborn	Renumbered SOP from #34 to #35. Updated SOP references.	A new SOP has been added to the protocol therefore some SOP reference #s have changed.	1.01

Only changes in this specific SOP will be logged here. Version numbers increase incrementally by hundredths (e.g., version 1.01, version 1.02) for minor changes. Major revisions should be designated with the next whole number (e.g., version 2.0, 3.0, 4.0). Record the previous version number, date of revision, author of the revision, changes made, and reason for the change along with the new version number.

Purpose

This SOP describes methods for data analysis and reporting of stream macrofauna and habitat surveys. Data analyses for the abundance and size of fish, shrimp, and snails, as well as geomorphologic habitat characteristics, are outlined with examples to show data structure, statistical tests, and graphical plots. This SOP outlines two initial steps of data analysis at each of the station, stream, and trend assessment levels. These steps summarize and determine the range-of-variation for some parameters. A third step which integrates habitat characteristics with biological data is also included at the trend assessment level. Report types are identified according to the categories in the PACN monitoring plan, with most reporting occurring on an annual basis. Specifically, station level and stream level size, abundance, and habitat characteristic summarization and range of variation data is reported on an annual basis.

Statistical trend analyses are conducted on abundance data on an annual basis (after five years of data have been collected). Trend analysis of size and habitat data, as well as integration of biological data with habitat characteristics, is included in the five-year reports. Additional, multi-vital sign integration, synthesis, analysis, and reporting is not addressed in this SOP at present. The use of products identified in this protocol does not imply endorsement, effectiveness, or warranty by NPS.

A master equipment list for the entire Pacific Islands Stream Monitoring Protocol can be found in [SOP #2 Before the Field Season](#). The master equipment list should be updated simultaneously if any SOP requiring an equipment list is revised.

Analysis Overview

Data analysis is defined here as the steps by which observations of the environment are processed for interpretation and synthesis into meaningful information that is accessible to managers. Data analysis also includes quality assurance and control efforts that occur in the field, during exploratory analysis, in summarization, during interpretation, and when drawing conclusions.

Analytical Approach

Two basic initial steps are identified in data analysis for all Pacific Islands Stream monitoring data: summarization and establishing the range of variation. These steps are encompassed in the larger construct of data management and data stewardship which is discussed in [SOPs 26-32](#). We identify four basic levels of analytical methods for our monitoring data: station level, stream level, trend assessment level, and synthesis (Table SOP 35.1). Quadrant (Hawaii and American Samoa) or segment (Guam) level macrofauna abundance and size data is used to calculate summarization and range of variation data (mean and standard deviation) for each station. Similarly, transect data is used to calculate summarization and range of variation data for habitat characteristics (e.g. velocity, substrate characteristics, etc.). The station level data includes mean and standard deviation abundance, size, and habitat characteristic data at a given station over time. The stream level data includes mean and standard deviation abundance, size, and habitat characteristic data at all stations in a stream at a given time. The trend assessment level integrates station or stream level abundance, size, and habitat characteristic summarization and range of variation data over time to detect change. For example, some form of regression may be used to identify the slope or trend. In addition, a third step is included in the trend assessment level that uses multivariate statistics to integrate station and stream level macrofauna abundance and size data with habitat characteristics. Finally, synthesis examines relationships between temporal and spatial trends in stream macrofauna size and abundance and water quality data. Station level, stream level, and trend assessment level are addressed here. Synthesis will be addressed within and across multiple Vital Signs and is therefore left for network level or broader scientific consideration in the future.

Table SOP 35.1. Three approaches for analyzing Pacific Islands Stream Vital Sign monitoring data.

Analysis	Description	Responsible Party
Station Level	Quality assurance and control routines and calculation of statistics from monitoring data (abundance, size, habitat):	Biological technician with oversight by PACN Aquatic Ecologist and with assistance from park leads
	Step 1 (Summarization): Measures of mean, median, variation, and other basic statistics. Include graphical presentation of data.	
	Step 2 (Range of Variation): Establish historical or expected range of values, relation to relevant regulatory levels, confidence estimates.	
Stream Level	Quality assurance and control routines and calculation of statistics from monitoring data (abundance, size, habitat):	Biological technician with oversight by PACN Aquatic Ecologist and with assistance from park leads
	Step 1 (Summarization): Measures of mean, median, variation, and other basic statistics. Include graphical presentation of data.	
	Step 2 (Range of Variation): Establish historical or expected range of values, relation to relevant regulatory levels, confidence estimates.	
Trend Assessment	Step 1: Integration of station and stream level summarization over time using some form of regression analysis.	PACN Aquatic Ecologist and researchers from other institutions (e.g., USGS)
	Step 2: Integration of station and stream level variation results over time using some form of regression analysis. Includes establishing a direction and rate of change of variation that may be used to provide early warnings of trends in resources condition. Confidence levels of documenting trend will be established.	
	Step 3: Integration of station and stream level abundance and size data with habitat characteristics using a multivariate approach.	

These analytical approaches will be applied to each of the components of Vital Signs (e.g. size and abundance of fish, shrimp, and snails and habitat characteristics).

Analysis Procedures

Analysis procedures are to be documented, for each stream on an annual basis as well as on a five-year basis, in a compiled document (such as a Microsoft® Word file [.docx]), referred to hereafter as an analysis log file. This file would essentially be a ‘log’ of all the quantitative and qualitative steps taken, such as various transformations tested, and screen shots of data visualizations. A check list for the steps is provided in Appendix 14: Analysis Log File Checklist. Each year, seven such files (one for each stream that is slated for initial implementation of the Pacific Islands Stream monitoring protocol) will be generated. On Tutuila (American Samoa) it is possible that additional streams will be selected for the yearly random sampling. These analysis log files are internal working documents, that while not subject to explicit peer-review for most reporting other than protocol and programmatic reviews, will serve as the documentation and foundation for analyses for reports which are peer-reviewed. Table SOP 35.2 summarizes these analytical procedures and frequency of analysis (annually or every five years) for the parameters measured in each stream.

Table SOP 35.2. Summary of analytical procedures including parameters measured at the stream level.

Level of Analysis	Data Analysis Approach	Frequency of Analysis	Responsible Party
Station	Summarization and range-of-variation for annual macrofauna size, abundance, and habitat characteristics at a given station over time	Annually	PACN Aquatic Ecologist
Stream	Summarization and range-of-variation for annual macrofauna size, abundance, and habitat characteristics at a given time over stream longitudinal distribution.	Annually	PACN Aquatic Ecologist
Trend Assessment	Multi-year change in macrofauna abundance for both station and stream level data.	Annually (after five years of data have been collected)	PACN Aquatic Ecologist and researchers from other institutions (e.g., USGS)
	Multi-year change in macrofauna size and habitat characteristics for both station and stream level data.	Every five years	PACN Aquatic Ecologist and researchers from other institutions (e.g., USGS)
	Integration of station and stream level abundance and size data with habitat characteristics using a multivariate approach.	Every five years	PACN Aquatic Ecologist and researchers from other institutions (e.g., USGS)

Station and Stream Level Analyses: These analyses are conducted annually for macrofauna size, abundance, and habitat characteristic summarization and range of variation data. The data sets are analyzed for normality of the distribution. Text and graphic plots of mean and standard deviation at each station over time (station level analysis) and at each time over stream longitudinal distribution (stream level analysis) are created. If multiple observers have been collecting data, the summaries should be grouped by observer to evaluate potential observer bias.

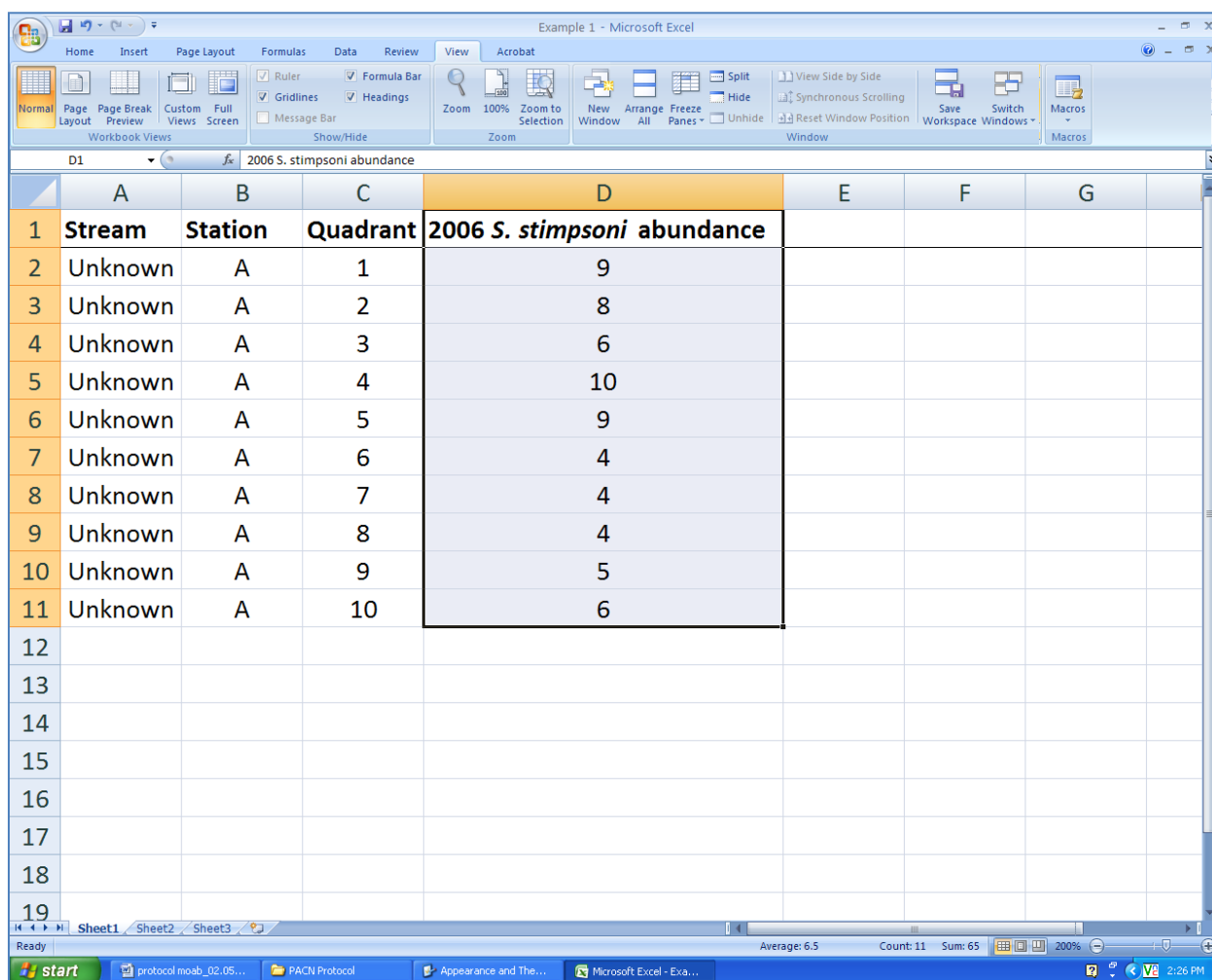
Trend Assessment Analyses: These analyses are conducted annually for macrofauna abundance data after the first five years of data have been collected. These analyses are conducted on a five year recurrence interval for macrofauna size data and habitat characteristic data. However, if an unusual change in abundance is detected, trend analyses for size and habitat characterization should be conducted annually. Text and graphic plots of mean and standard deviation of abundance, size, or habitat characteristics at each station over time (station level analysis) and at each time over stream longitudinal distribution (stream level analysis) are created. Change detection over time can be analyzed using a variety of statistical techniques, including repeated measures ANOVA or regression. Integration of stream and station level abundance and size data with habitat characteristics will be accomplished every five years using canonical correspondence analysis (CCA).

Example Procedure for Stream Macrofauna Abundance

It should be noted that the following analytical procedures should only be conducted by an individual with advanced statistical training. In addition, many of the procedures described below may change over time due to advancements and updates in statistical tests, perspectives, techniques and software. Consequently, this example merely serves to illustrate the desirable sequence of steps necessary to examine and present the data. The data presented in these sections is used for example purposes only and should not be used as part of actual data analysis activities.

Example #1: Checking for normality of the distribution.

Step 1: Copy annual quadrant (or segment in Guam) data for a station from a Microsoft Access query into Microsoft Excel. Keep in mind that data copied from Access is treated as text and must usually be converted into value format. Programs like Excel can do the value conversion. In this example, abundance of *Sicyopterus stimpsoni* in Unknown Stream from ten quadrants at Station A in 2006 is pasted into the Example 1 Excel file (.xlsx) (Figure SOP 35.1).



	A	B	C	D	E	F	G
1	Stream	Station	Quadrant	2006 <i>S. stimpsoni</i> abundance			
2	Unknown	A	1	9			
3	Unknown	A	2	8			
4	Unknown	A	3	6			
5	Unknown	A	4	10			
6	Unknown	A	5	9			
7	Unknown	A	6	4			
8	Unknown	A	7	4			
9	Unknown	A	8	4			
10	Unknown	A	9	5			
11	Unknown	A	10	6			
12							
13							
14							
15							
16							
17							
18							
19							

Figure SOP 35.1. Abundance data structure at the station level in Microsoft Excel.

Step 2: Check the normality of the distribution of the abundance data by selecting “Data Analysis” under the “Data” tab. Choose “Descriptive Statistics” and define the abundance data as the input range. Select a blank cell as the output range and check the “Summary statistics” box. If the absolute value of the skewness statistic is greater than two standard errors of skewness you can assume that the distribution is significantly skewed (non-normal). The standard error of skewness can be estimated as $\sqrt{\frac{6}{N}}$ (Tabachnick and Fidell, 1996). In this example, the absolute value of skewness is 0.3, which is less than the standard error of skewness (0.77). Therefore, the data can be considered normally distributed.

Step 3: Repeat this procedure for each species in each segment in each stream for the macrofauna abundance and size data as well as the habitat characteristics data each year.

Example #2: Station level analysis of fish abundance data.

Step 1: Copy annual quadrant (or segment in Guam) data for a station from a Microsoft Access query into Microsoft Excel.

Step 2: Using quadrant (or segment in Guam) data calculate the average abundance and standard deviation in a segment on a given date for each species. In this example, the 2007 annual average abundance and standard deviation data of *S. stimpsoni* for Unknown Stream at Station D in Unknown Stream is pasted into the Example 2 Excel file (.xlsx) which already contains data from 2000-2006 (Figure SOP 35.2).

The screenshot shows a Microsoft Excel spreadsheet titled "Example 2 - Microsoft Excel". The spreadsheet contains a table with the following data:

	A	B	C	D	E	F	G	H
	Stream	Station	Year	<i>S. stimpsoni</i> Station Average Abundance	<i>S. stimpsoni</i> Station Standard Deviation			
1	Unknown	D	2000	10.3	5.40			
2	Unknown	D	2001	16.3	4.50			
3	Unknown	D	2002	9.8	4.32			
4	Unknown	D	2003	6.7	2.75			
5	Unknown	D	2004	4.2	3.26			
6	Unknown	D	2005	1.6	1.90			
7	Unknown	D	2006	3	1.94			
8	Unknown	D	2007	1.7	1.25			
9								
10								
11								
12								
13								
14								
15								
16								

Figure SOP 35.2. Average abundance and standard deviation data structure at the station level in Microsoft Excel.

Step 3: Create a plot that shows the change in average station abundance and standard deviation over time (Figure SOP 35.3). This plot represents station level summarization and range of variation data for *S. stimpsoni* at Station D in Unknown Stream.

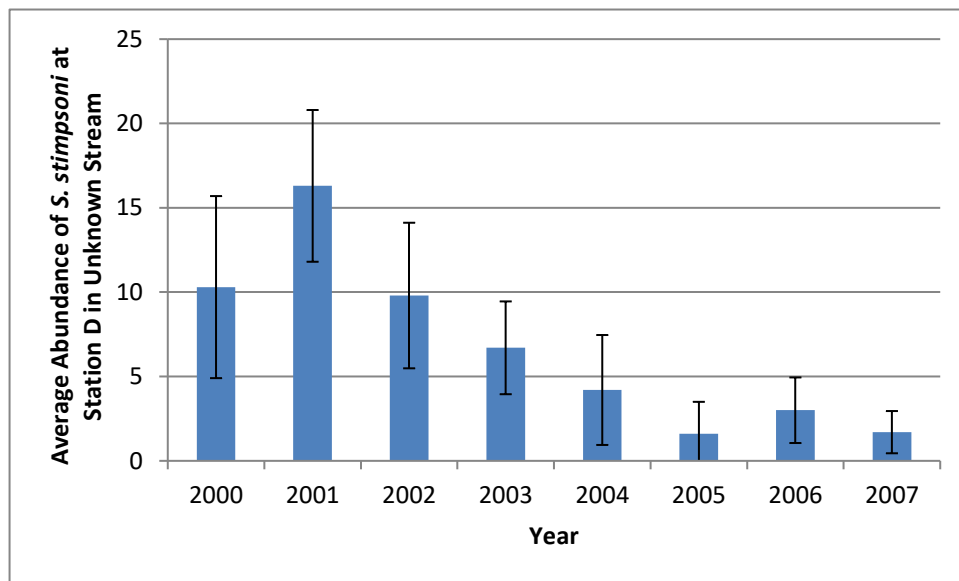


Figure SOP 35.3. Average station level abundance $\pm$ 1 standard deviation of *S. stimpsoni* at Station D in Unknown Stream.

Step 4: Repeat this procedure for each species in each segment of each stream for the macrofauna size, abundance, and habitat characteristics data on an annual basis.

Example #3: Stream level analysis of fish abundance data.

Step 1: Copy annual quadrant (or segment in Guam) data for a station from a Microsoft Access query into Microsoft Excel.

Step 2: Using quadrant (or segment in Guam) data calculate the average abundance and standard deviation in each segment on a given date for each species. In this example, the 2000 annual average abundance and standard deviation data of *S. stimpsoni* for Unknown Stream at Stations A-W in Unknown Stream is pasted into the Example 3.xls (Figure SOP 35.4).

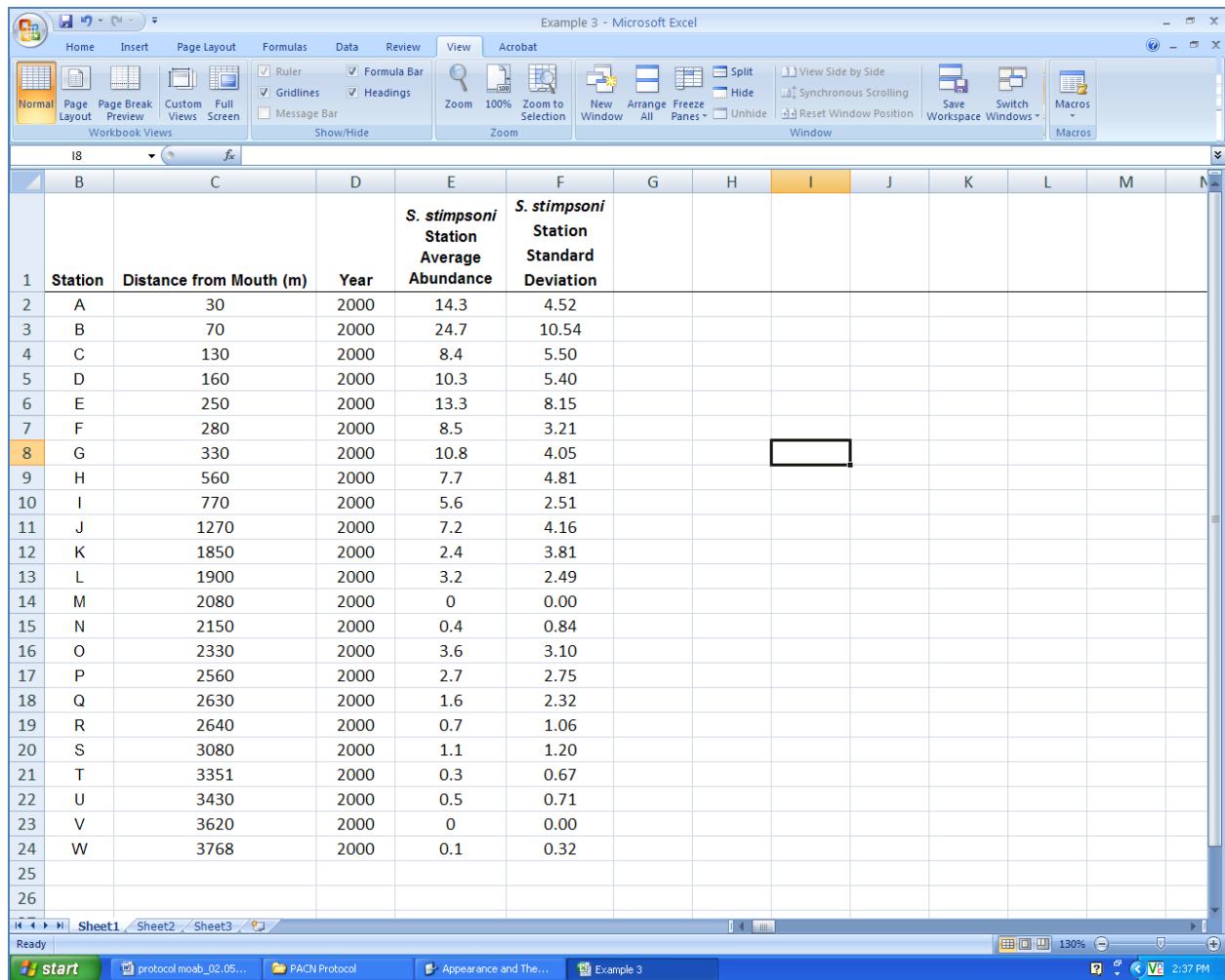


Figure SOP 35.4. Average abundance and standard deviation data structure at the stream level in Microsoft Excel.

Step 3: Create a plot that shows the change in average station abundance and standard deviation over longitudinal distance from the mouth of the stream (Figure SOP 35.5). This plot represents stream level summarization and range of variation data for *S. stimpsoni* at stations A-W in Unknown Stream in 2000.

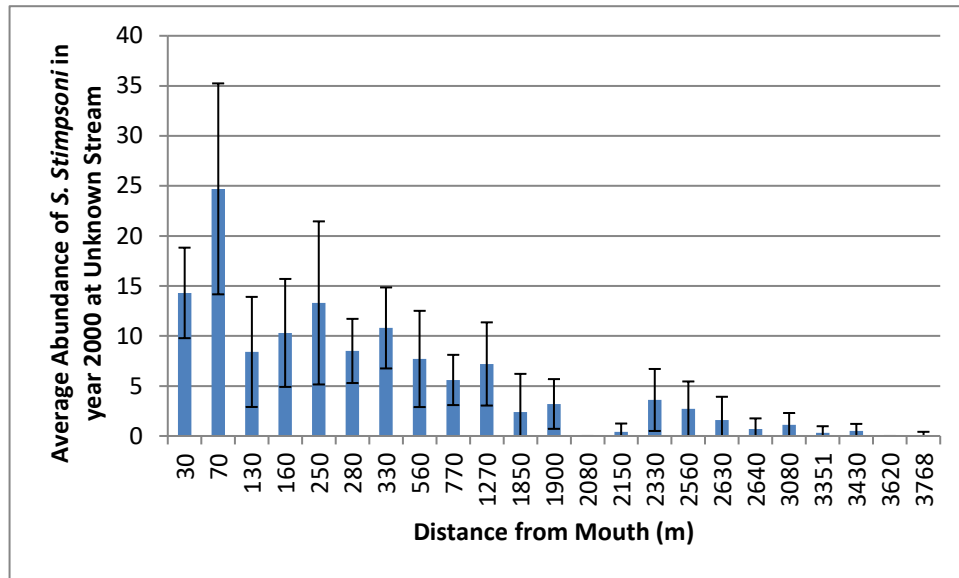


Figure SOP 35.5. Average stream level abundance $\pm$ 1 standard deviation of *S. stimpsoni* at in year 2000 at Unknown Stream.

Step 4: Repeat this procedure for each species in each stream for the macrofauna size, abundance, and habitat characteristics data on an annual basis.

Reporting Overview

Report Types

Unlike the analysis log files which will generate individual files for each park and each of the three analysis strategies (station, stream, and trend assessment levels), reporting documents will integrate both parks and analysis strategies. Table SOP 35.3 identifies product types, purposes, targeted audiences, responsible parties, production frequency, and review processes. We have identified a cohesive suite of seven product categories: (1) program and protocol reviews, (2) monitoring protocol and project reports, (3) status and trends reports, (4) scientific writing and presentations, (5) management briefings, (6) website communication, and (7) interpretation and outreach.

Table SOP 35.3. Summary of anticipated products grouped by type and frequency.

Type of Report	Purpose of Report	Targeted Audience	Initiated by	Frequency of Reporting	Review Process
Program and Protocol Reviews					
Protocol Review Reports	Document where actual procedures fall short or exceed expectations, recommend necessary changes, document changes since last protocol review report; document overall quality of protocol – particularly in terms of protocol objectives and implementation, effectiveness, and data management.	Superintendents, park resource staff, I&M staff, service-wide program managers, external scientists, partners	PACN Aquatic Ecologist	Within 1-3 years of implementation, 5-year intervals thereafter	Peer review at network and regional level
Monitoring Protocol and Project Reports					
Vital Signs Monitoring Protocol Reports	Document and archive annual monitoring activities and data, describe current resource condition and core analysis results, document related data management activities, document changes in monitoring protocol, communicate monitoring efforts to resource managers. During protocol development stages, will emphasize progress made and challenges encountered.	Park resource staff, PACN staff, external scientists, partners	PACN Aquatic Ecologist	Annually, compiled each March	Peer review at network level
Summary of Vital Sign Monitoring Protocol Reports	Same as annual "Vital Sign Monitoring Protocol Reports" above, but highlights key points for non-technical audiences	Superintendents, NPS interpreters, public, partners	PACN Aquatic Ecologist	Annually, compiled each March	Peer review at network level
Pilot Projects and Monitoring Research Reports	Provide background and methods of monitoring protocol development and other methods related investigations	Park resource managers, PACN staff, external scientists, partners	PACN Aquatic Ecologist	Variable with annual status reporting, as necessary	Peer review at network level plus review appropriate to final product
Status and Trends Reports					

Type of Report	Purpose of Report	Targeted Audience	Initiated by	Frequency of Reporting	Review Process
Trend Analysis and Synthesis Reports	Describe and interpret patterns/trends of monitored resources, identify new characteristics of resources and correlations among monitored resources, identify relationships between drivers/stressors and responses, recommend changes to management of resources (adaptive management feedback).	Park resource managers, PACN staff, external scientists, partners	PACN Aquatic Ecologist, park leads	3-5 year intervals. Abbreviated annual edition, as necessary for PACN Vital Signs Program	Peer review at the network and regional level
Summary of: Trend Analysis and Synthesis Report	Executive summary of "Trend Analysis and Synthesis Report" above with key points on one page for non-technical audiences. Usually this is a bulleted list.	Superintendents, NPS interpreters, public, partners	PACN Aquatic Ecologist	Commensurate with reporting activity of "Trend Analysis and Synthesis Report"	Peer review at the network level
PACN contribution to NPS-wide "State of the Parks" Report	Describes current conditions of park resources, reports interesting trends and highlights of monitoring activities, identifies resource issues of concern, explores future issues and directions	Congress, budget office, NPS leadership, superintendents, general public	Compiled by the Washington Support Office (WASO) from data provided by networks	Annual	Peer review at national level
Scientific Writing and Presentations					
PACN Vital Signs Monitoring Conference	Review and summarize information on this Vital Sign, help identify emerging issues and generate new ideas	Park resource staff, network staff, external scientists, partners	PACN Aquatic Ecologist	Biennial (around time PACN "Status and Trends Report" is published)	Peer review at national level
Scientific journal articles and book chapters	Document and communicate advances in knowledge, provides a broader perspective on quality assurance and peer review	External scientists, park resource managers, and professional staff	PACN Aquatic Ecologist, others	Variable	Peer review according to journal or book standards
Other symposia, conferences and workshops	Review and summarize information on this Vital Sign, help identify emerging issues and generate new ideas	External scientists, professional staff, park resource managers, and other resource managers.	PACN Aquatic Ecologist, others	Variable (e.g., Hawaii Conservation Conference, George Wright Society)	Peer review at network level; for papers may also be peer reviewed
Management Briefings					

Type of Report	Purpose of Report	Targeted Audience	Initiated by	Frequency of Reporting	Review Process
Protected area managers briefing	Communicate highlights and potential management action items, with 1-2 page briefing statements for each protocol	Park resource staff, Network staff, agency, academic scientists, other Federal, State, and Territorial Protected Area managers, discipline specialists, interpretive staff	Network program manager, PACN Aquatic Ecologist	Annually, likely in conjunction with Board of Directors administrative meetings	Peer review by network, PICRP, and monitoring staff
Executive briefings	Update superintendents and other VIPs on park-specific findings and potential resource issues; suggest action items where appropriate	Individual Superintendents and other VIPs	PACN Aquatic Ecologist, network program manager, park leads	As needed	Peer review by network, PICRP and monitoring staff
Website Posting					
Web-based media	Centralized repository of all final reports to ensure products are easily accessible in commonly-used electronic formats; other synthesized information on the PACN	Superintendents, park resource staff, PACN staff, service-wide program managers, external scientists, partners, students, public	Variable, typically network webmaster	As media is completed	Peer review at network level to NPS web standards as finalized, reviewed products
Interpretation and Outreach					
Science Days	Communicate main monitoring findings as well as underlying data; discuss potential significance for management, further monitoring, potential additional research needs, and for outreach	Superintendents, park resource staff, PACN staff, protocol managers, partners, public	PACN Aquatic Ecologist, biological technician, and others as needed	Variable by park, annual when possible	Meeting / presentation itself is a form of review
Interpretive Conversations	Interactive conversations with park interpretive staff to discuss main monitoring findings as well as underlying data; discuss potential significance for management, further monitoring, potential additional research needs, and for outreach	Park interpretive staff, environmental educators, PACN staff	PACN Aquatic Ecologist, biological technician, and others as needed	Variable by park, at least annual when possible	Meeting / presentation itself is a form of review

Type of Report	Purpose of Report	Targeted Audience	Initiated by	Frequency of Reporting	Review Process
Park Interpretive / outreach sessions	Review and summarize information on PACN Vital Signs; engage and involve greater participation in monitoring efforts	Park staff, public, partners	PACN Aquatic Ecologist, biological technician, and others as needed	Variable	Peer review by network, PICRP staff
Park staff meetings (results synthesis)	Communicate results to non-technical audiences, discuss potential significance for management, receive feedback on resource and monitoring issues in park operations	All park staff, volunteers, and partners, especially those not typically encountered in I&M program, RM, science operations	PACN Aquatic Ecologist, biological technician, and others as needed	Annually for each network park	Peer review by network, PICRP staff



Figure SOP 35.6. Image example of one site of actual field data collection upon which reports should be based.

Report Preparation Process

Reports identified in Table SOP 35.3 are to be based on collected data (Figure SOP 35.6) and other science-based documentation. An example summary report of vital signs data that could be presented to park superintendents is included in Appendix 15: Pacific Islands Stream Monitoring Report: Example Summary of Vital Signs Data. These reports will provide context and explanations for the results presented, as well as include any additional analyses that may result from initial investigations. The ‘analysis log’ files will serve as the basis for most of the reporting results. The biological technician for this Vital Sign is identified as the primary individual responsible for updating the analysis log files and will also be pivotal in preparation of subsequent reports. It is anticipated that the PACN Aquatic Ecologist or another statistically trained individuals will be responsible for the statistical analyses.

Drafts of the reporting products will be organized according to the categories outlined above, with key versions and review comments archived, and final versions clearly communicated to PACN network staff to ensure distribution via websites and other channels.

Literature Cited

Tabachnick BG, and Fidell LS. 1996. *Using multivariate statistics (3rd ed.)*. New York: Harper Collins, New York, New York, USA.