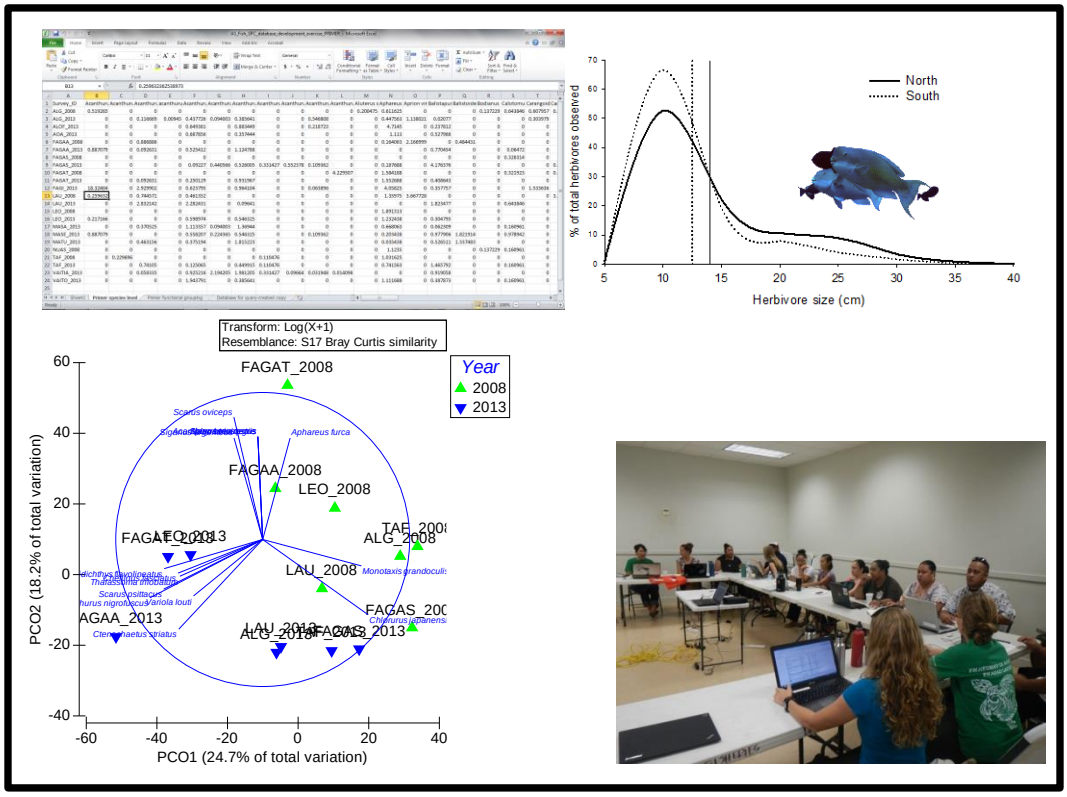
Workshop Report: Improving Data Collection, Storage, Handling, Visualization, and Analyses for Coral Reef Monitoring Programs

Dr. Peter Houk

In partnership with:

The Nature Conservancy
American Samoa Coral Reef Advisory Program



February 2 - 6, 2015

Pago Pago, American Samoa

NOAA CORAL REEF CONSERVATION PROGRAM

The Nature Conservancy
Protecting nature. Preserving life.

Acknowledgements:

Thanks to the American Samoa Division of Marine and Wildlife Resources for assisting with the workshop logistics, especially Alice Lawrence. Thanks also to Cherie Wagner, The Nature Conservancy, for logistical support with the workshop and for administering the evaluations. Conversations between numerous local biologists, TNC, and the author were instrumental in guiding the workshop agenda. This report was prepared by the author for The Nature Conservancy under award #NA13NOS4820145 from the National Oceanic and Atmospheric Administration's (NOAA) Coral Reef Conservation Program, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of NOAA, the NOAA Coral Reef Conservation Program, or the U.S. Department of Commerce.

Table of Contents

Workshop background and summary:

A workshop was conducted to improve the analytical expertise within resource monitoring and management programs in American Samoa. Key topics covered were database development, data querying, univariate and multivariate visualization, statistical power assessments, and assessments of overall monitoring program designs. The workshop represented a logical progression of previous analytical themes that Dr. Peter Houk (instructor) has introduced to local programs during 2008. The previous workshop provided an introduction to database designs, data querying, and basic visualization of data queries using exemplary data from coral-reef monitoring programs *outside* of American Samoa. The present workshop utilized two modern sets of data that were made available to resource managers and of direct relevance to their reef ecosystems: 1) decadal datasets pertaining to coral, benthic substrate, and fish assemblages collected by Dr. Houk in partnership with American Samoa Environmental Protection Agency (ASEPA) since 2003, and 2) the National Oceanic and Atmospheric Administration (NOAA) Coral Reef Ecosystem Division (CRED) fish datasets collected from Tutuila in both 2010 and 2012 graciously provided by NOAA-CRED scientists. The workshop began with a review of previous concepts, and then advanced several key topics: localized database querying for user-defined needs, assessments of statistical power and sampling designs, building of compelling and professional graphics to display data, and multivariate approaches towards data assessments (Appendix 1). The format of the workshop began with hands on presentations and exercises conducted simultaneously by both Dr. Houk and participants. However, each day also included independent and group exercises for participants to showcase their skills and present back to the group. Beyond analytical skillsets, key outcomes also included a deeper understanding of coral-reef and fish assemblage dynamics across Tutuila since 2003 developed from the perspective of two different datasets. This represented a strong foundation for current biologists to base their evolving question-based monitoring designs. Last, the workshop was presented in a generalized format with relevance to both terrestrial and marine scientists and managers, as keen interested during the workshop clearly existed across resource managers (Appendix 2 and 3).

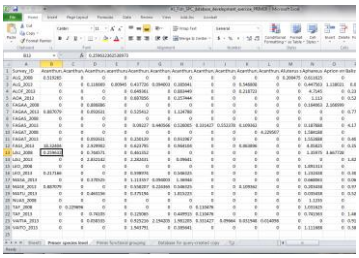
Daily Activities:

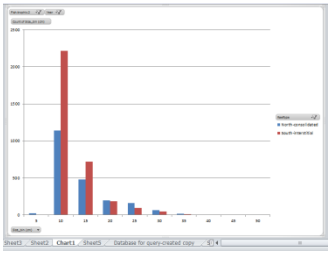
- Day 1 activities included a review and catch-up of database development themes introduced several years back. However, participants spent time in developing their own localized database to fit coral-reef fish monitoring needs in American Samoa. We next began the process of data querying using Microsoft Excel Pivot tables. All participants were given sets of questions that they answered independently and presented back to the group. Useful discussion and question/answer followed.

- Day 2 focused on advanced use of Pivot Tables to export data from databases into several other formats required for advanced graphing, power analyses, and multivariate analyses. Similarly, participants were required to conduct their own examples after some instruction, and discuss findings with the group (Figure 1).

- Day 3 focused upon using power analyses tools to evaluate question-based monitoring programs. We introduced the R-program platform, and generated our own scripts (i.e., code) to assess statistical power from numerous aspects of both ASEPA and CRED monitoring programs (Figure 2). This exercise led into discussions on how to improve power, and specifically focus upon accounting for natural variability by stratification versus increasing sampling sizes. This also led to discussions on how to best design localized monitoring efforts.

Figure 1. *Progression of the data query process from Pivot Tables, to query charts, and finally to compelling graphics describing the trend in parrotfish size distributions between the north and south side of Tutuila.*





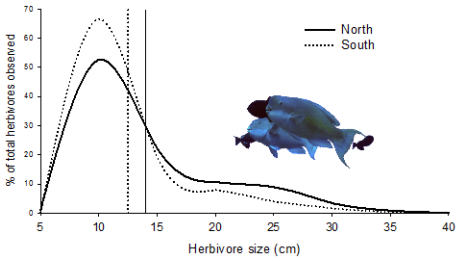
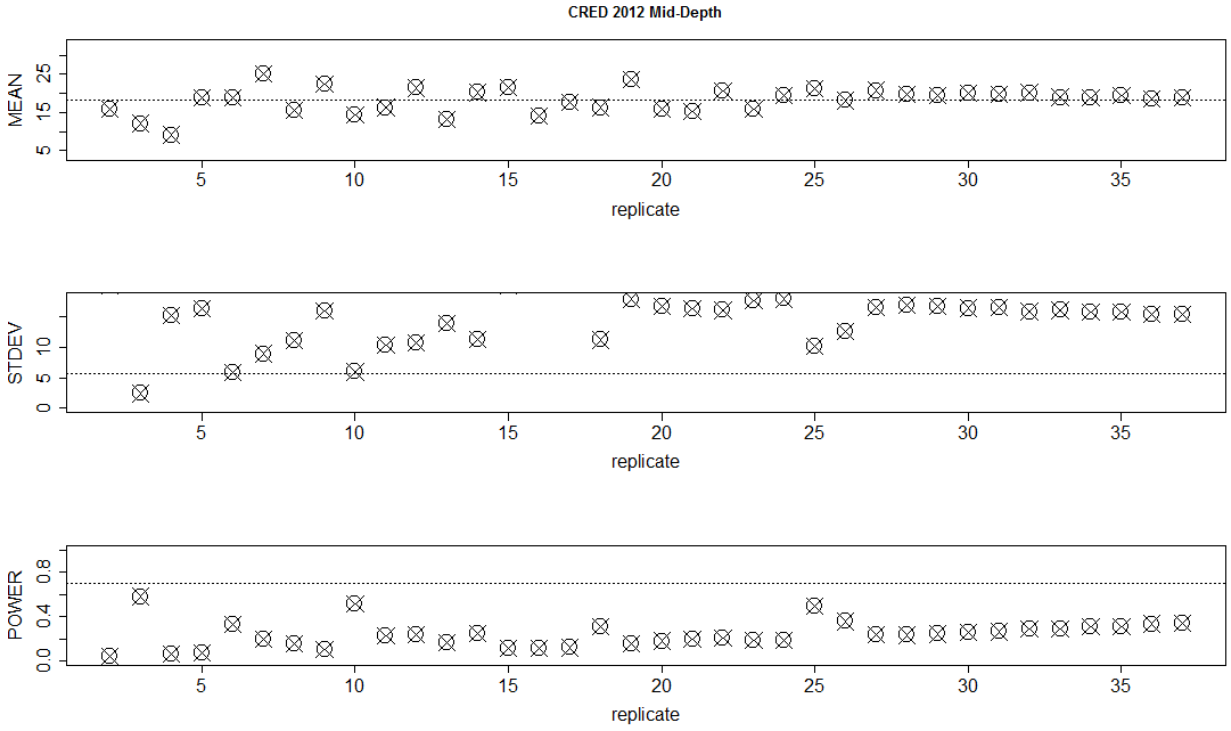


Figure 2. *R-plot showing the accumulation of statistical power based upon CRED 2012 data for total fish biomass collected at the 'mid' depth range. Here, criteria were set to assess the power to detect a relative 30% change in overall fish biomass as sampling intensity increased.*



- Day 4-5 focused upon introduction and advanced use of multivariate species abundance datasets. We used the PRIMER/PERMANOVA software platform for these needs because it maximized user understanding. These analyses provided participants a greater appreciation for how to understand complex, large species abundance datasets. Individual exercises examined the changes in fish assemblages depicted by multivariate analyses between 2008 and 2013 from ASEPA monitoring data (Figure 3).

Figure 3a-b. *Examples of multivariate analyses examining the potential change in fish assemblages around Tutuila between surveys conducted in 2008 and 2013 at the same set of 7 sites around the island. This example was selected because challenges existed. First, data from different years were collected at varying intensities as 5 replicates-per-site were conducted in 2008 versus 12 in 2013. In addition the observers were both well-experienced scientists, but we were careful not to assume taxonomic expertise was the same among observers. Data were first normalized by sampling intensity, then compared. Species-level data depicted a shift in fish assemblages between the years that was mainly driven by a reduction in larger parrotfishes, some secondary consumers, and a few tertiary snappers (Fig. 3a). Next, in order to account for potential biases due to taxonomy, fish were first aggregated at the 'functional group' level, and a similar analysis was conducted. Functional grouping also depicted difference between the two survey years, with one notable outlier (Fig. 3b). The functional groups of fish driving these trends were similar, yet distinctions were less pronounced. This example provided a means for participants to perform similar analyses with the numerous other datasets that exist (i.e., CRED, DMWR), and provided a null hypothesis regarding fish assemblages that can be furthered with additional data collection/analysis.*

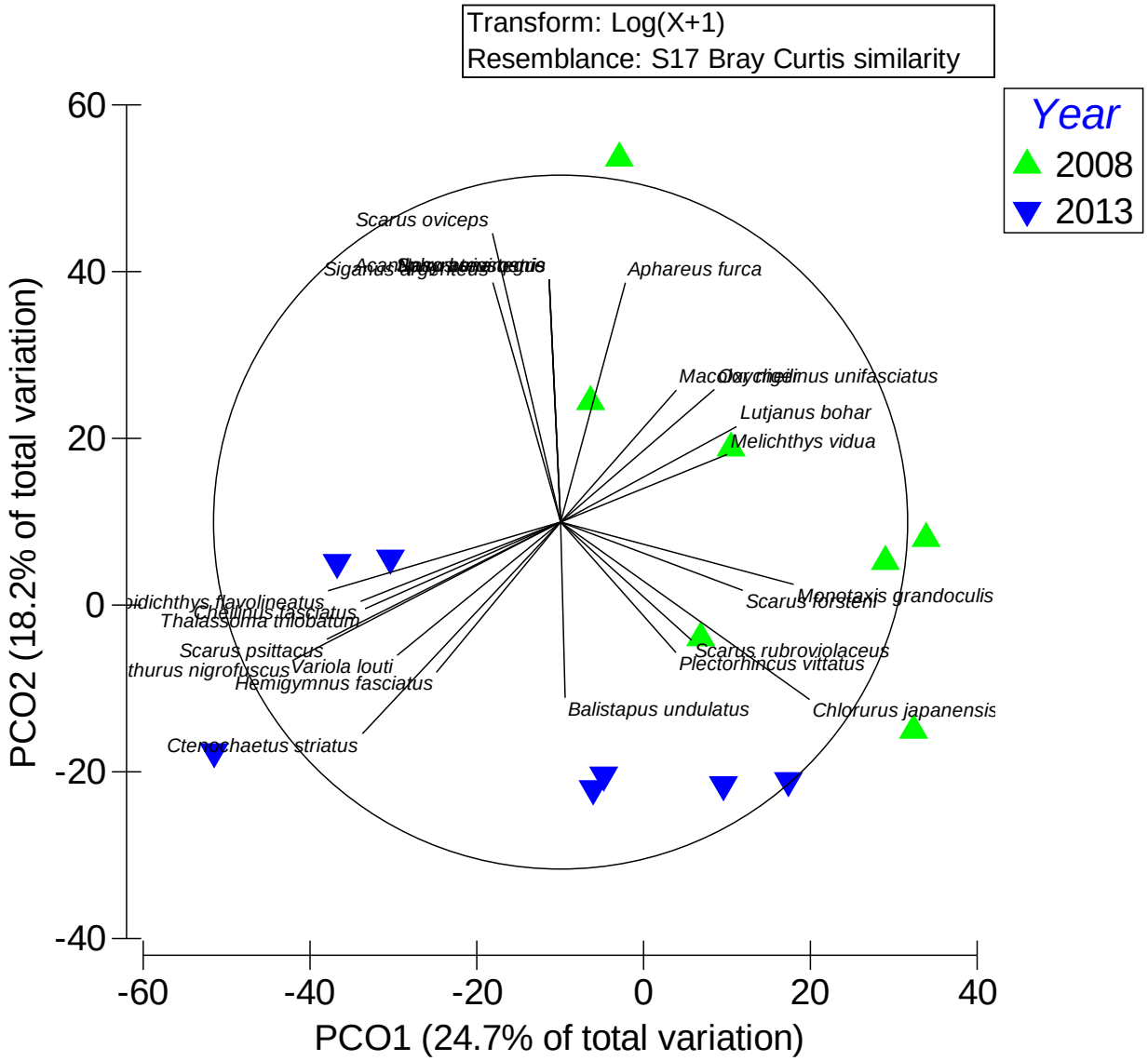


Fig. 3a.

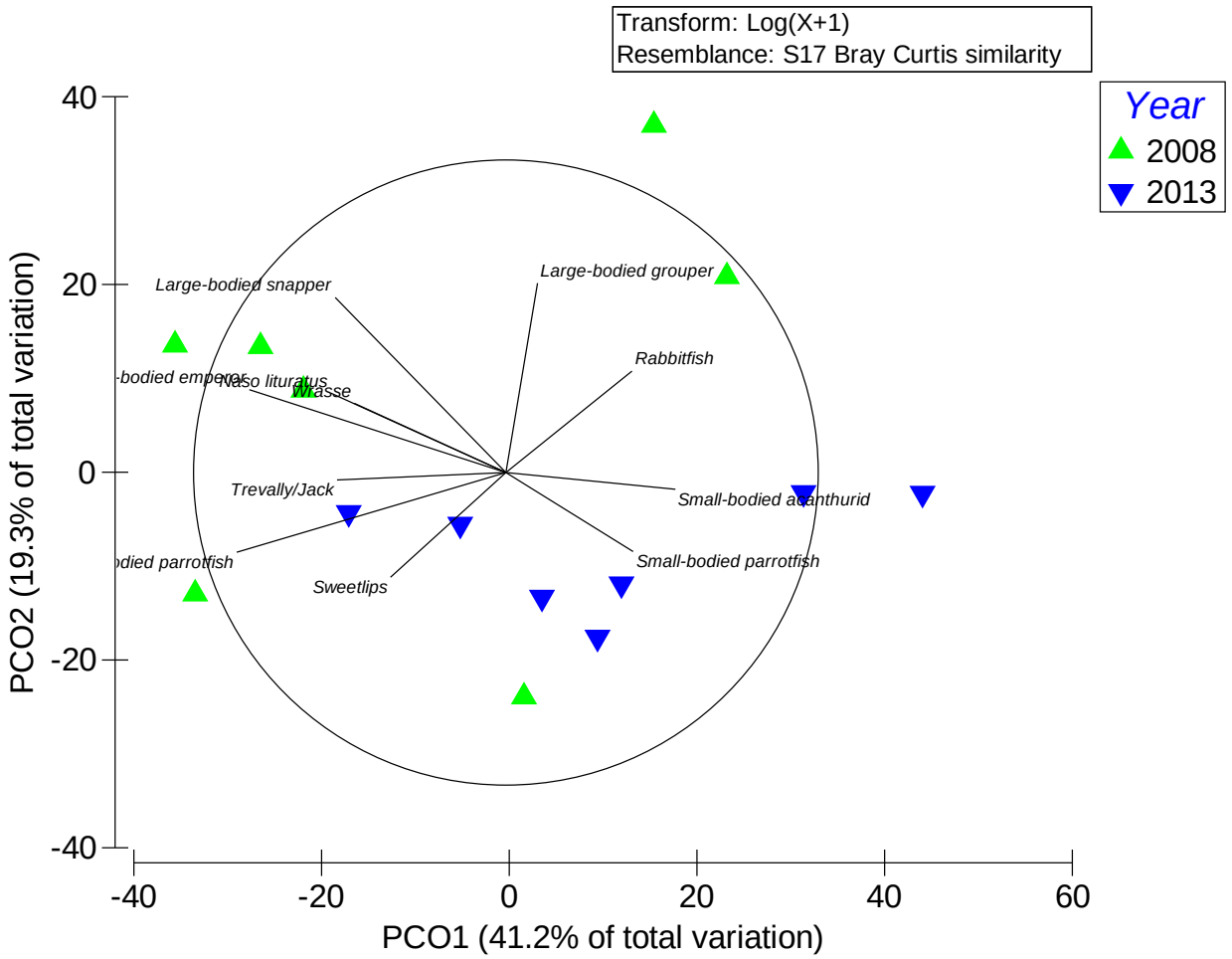


Fig. 3b.

Summary and Future Directions:

Anonymous workshop evaluations were provided to all participants and compiled by TNC (Appendix 3). Evaluations confirm that one of the most difficult aspects of local resource monitoring and management programs is to have access to all the emerging tools that aid in data understanding and visualization. Beyond having access to these tools, working through specific examples together was a desirable outcome for transferring both knowledge contained within datasets and skills needed to extract this knowledge into the future. Building upon our successful collaboration, participants expressed an interest in further data development partnerships. Several avenues were discussed and are presented hierarchically, in accordance with the participants desires. Prior to workshops, it would be requested that American Samoa resource managers meet and come up with priority questions that overlap agencies. This would then help maximize efficiency for future workshops. Key themes were:

1. Participants expressed an interest for a future workshop where they bring their own data to the table, and over the course of the workshop they would (i) clean the databases, (ii) examine the statistical confidence in their survey designs, (iii) query the databases and discuss findings, and (iv) develop compelling, professional graphics that depict answers to priority questions that the agencies have.

2. A second suggestion was made to focus more deeply upon (1) major set of data that are central to several pressing questions that exist across the resource management programs. Then, the group would generate testable questions together, query/analyze data, produce compelling graphics, and prepare a technical report/peer-review publication on behalf of the group. In sum, the group identifies, or collects, a dataset together that is central to many pressing questions, we then analyze and professional report upon the findings together.
3. Finally, a suggestion was made for programs to undertake the above step(s) independently, then convene a group workshop to aid in writing/revising publication-quality summaries of pertinent findings, and creating summary powerpoints for dissemination. This of course hinges upon the successful completion of the above steps, and also with the expertise of individuals. Notably, our workshop had a range of skill and expertise.

In conclusion, the present workshop bolstered the analytical skillsets of resource managers in American Samoa. There was an appreciation for this line of guidance, and specifically for improving the ability to interpret survey designs, efficiently interpret key findings from expansive datasets, and professionally presenting results. The feedback and insight provided by the participants to the instructor was also greatly appreciated, and will benefit monitoring programs across Micronesia with whom the instructor has longstanding relationships with.

Appendices

Appendix 1. Workshop agenda.

Meeting Agenda

Improving Data Collection, Storage, Handling, Visualization, and Analyses for Coral Reef Monitoring Programs

Division of Marine and Wildlife Resources, American Samoa

February 2 - 6, 2015

Day 1. General introduction to the diversity of monitoring programs and question-based designs. Introduce 2-3 examples of regional/local coral-reef monitoring programs, designs, and pressing questions that can (and can't) be answered based upon survey designs. We will formally examine statistical power to understand these concepts, but first, we will spend 1-2 days on building monitoring databases. These databases are required prior to understanding and analyzing statistical power.

Begin building basic relational databases using MS Excel for American Samoa food fishes (i.e., understand the framework for future use). Work on each step required to build a complete database. Using a completed database associated with the ASEPA coral-reef monitoring, begin exploring Pivot Tables for quality control checks and general querying.

Day 2. Continue with using Pivot Tables to generate data summaries. Begin the process of exporting data from Pivot Tables to produce high-quality graphics. We will either use Sigmaplot or R to produce graphics, I will have both software platforms. Basic graphs contrasting fish assemblages inside and outside of MPAs, gradients of fish biomass along the south and north coasts of Tutuila, and other key features of datasets will be produced together.

Begin basic examinations of the data taking a multivariate approach. We will use PRIMER software package for this, which will be provided for those that don't already have it. This will only be a basic introduction/review of multivariate approaches and what they offer. A deeper understanding would take more time than we have available, but we can discuss relevance and future desires that may exist.

Day 3. Finish up any activities from Day 1 and 2. Use Pivot Tables to generate data metrics needed for understanding statistical power. Examine data in greater detail with respect to statistical power and confidence using R. Develop R-scripts for your own personal use with data. This will provide participants with tools to evaluate statistical power on their own. Exercise on your own, apply R-script to some dataset you have, present back to group. Time/interest permitting, we will also evaluate statistical confidence taking a novel approach that incorporates multivariate, species abundance datasets.

Day 4 and 5. Use ASEPA dataset to examine reef 'health' or 'condition' scores. This will be hands-on. We will look at how to define individual metrics of interest, how to combine these, and how to come up with overall ecosystem-condition scores. We will then generate two useful proxies. First, the proxy to fishing pressure (distance from main boat launching areas, population, and wave exposure). Second, the proxy to land-based sources of pollution (watershed GIS layers). Once completed, examine the extent to which our proxies can (or can't) predict the gradient in coral-reef ecosystem condition across the south coast of Tutuila using regression modeling in R. Build your own R-scripts for future use with regressions. In small groups, build a similar regression model to examine trends for the same regression analysis across northern Tutuila sites. Present back to group.

Time permitting: 1) intro to population modeling, or 2) deeper introduction to the wave exposure GIS tool as a predictor of ecological assemblages.

Contact. Dr. Peter Houk, University of Guam Marine Lab, peterhouk@gmail.com

Participant List

Name	Agency
Alice Lawrence	Department of Marine and Wildlife Resources (DMWR)
Meagan Curtis	Coral Reef Advisory Group (CRAG)
Sean Felise	Department of Marine and Wildlife Resources (DMWR)
Peter Taliua'a	Department of Marine and Wildlife Resources (DMWR)
Josh Tuaunu	Department of Marine and Wildlife Resources (DMWR)
Christina Mata'afa	Department of Marine and Wildlife Resources (DMWR)
Salefu Tuvalu	Department of Marine and Wildlife Resources (DMWR)
Jewel Tuiasosopo	AS Environmental Protection Agency (AS-EPA)
Mia Comeros	AS Environmental Protection Agency (AS-EPA)
Siumu Faaiuaso	AS Environmental Protection Agency (AS-EPA)
Wendy Cover	NOAA National Marine Sanctuaries (NMSAS)
Adam Miles	Department of Marine and Wildlife Resources (DMWR)
Faleselau Tuilgai	Department of Marine and Wildlife Resources (DMWR)
Tepora Lavatai	Department of Marine and Wildlife Resources (DMWR)
Mareike Sudek	Department of Marine and Wildlife Resources (DMWR)
Kim Kayano	Department of Marine and Wildlife Resources (DMWR)
Jeremy Raynal	Coral Reef Advisory Group (CRAG)
Tim Clark	National Park Service (NPS)

Workshop Evaluation Summary

Overall participants strongly agreed that after attending the workshop they have a better understanding of how to handle, interpret, and visualize coral reef monitoring datasets. They also found the workshop beneficial in building their analytical skills, helping them to better understand and translate the data they collect. And a majority agreed that the information presented would held them with their work. Some comments about the workshop include:

- Awesome. Everyone stayed for the whole week for a data workshop! Unheard of!! Instructor very positive and supportive—helpful and motivating!
- This is the most valuable workshop I have ever attended. Very clear, excellent way of presenting difficult material so even those with minimal experience could grasp.
- This has really been a most helpful and beneficial workshop that will help our organization with looking at our environmental data to answer pertinent questions required not just to meet funding goals, but also to answer relevant questions to manage local resources.
- Learning how to use software like Sigma Plot and Primer (Permanova) were useful because I will be able to use it for our data analysis. During the workshop, I was able to pick up new tricks in Excel too.
- I believe understanding the basics ranging from pivot tables, to Sigma Plot, has ease the task of managing our data. Also the different programs that are included have made data management more useful.
- Data analysis will be done for our program and this has opened the door to producing technical reports, finally!
- I hope to use existing datasets (like CRED) to learn more about our sites.
- I intend to analyze about 10 years of data collected by the Department of Marine and Wildlife Resources (DMWR) that has never been investigated.
- Most of all our data has been managed, sorted, analyzed and so much more manually. In addition, individuals within different divisions are not well prepared. With this new skill of data handling, we are able to collect data efficiently, analyze correctly including detecting population change. Thanks to these new programs and Peter for teaching us.
- I will use the skills to analyze coral reef monitoring data for the territory and relate the results to management action. I will also use the information gained from the workshop to assist with developing a coral reef health index for the Territory.