

NOAA Fisheries Pacific Islands Fisheries Science Center
External Review Report
20-22 July 2010
Mark Fornwall, Chair

Review Panel: Dave Colpo (Pacific States Marine Fisheries Commission), Mark Fornwall (United States Geological Survey), Ted Habermann (NOAA, National Geophysical Data Center), Jim Sargent – NOAA, National Marine Fisheries Service), Steven Smith, (University of Miami)

Introduction and Background

A review of the Pacific Islands Fisheries Science Center (PIFSC) scientific data management systems and capabilities was held on 20-22 July 2010, at the East West Center, Honolulu, Hawaii. The principal objective of the review was to evaluate the PIFSC approach to management of its scientific data. The review included high-level background information, detailed presentations on various data projects/programs and follow-up discussions for clarification with selected individuals as well as exit meetings with division chiefs and PIFSC Directors. All sessions on days one and two were open to PIFSC staff and stakeholders with one exception, the stakeholder session did not include PIFSC staff. (See Attached agenda)

Report Organization

This report represents the consensus of the review panel, but is the chairman's report alone. The report is based upon the substance of all presentations as well as individual interaction with PIFSC members (formal and informal) and stakeholders. The report is organized around the principal questions posed with the review objectives (Appendix 2) and presents a summary of the findings. More detailed information and recommendations are included in the appendices 3 and 4.

General Observations

A key comment made by the Director was that the PIFSC mission was to: *Provide high-quality, timely applied scientific information to support the conservation and management of marine fisheries, protected species, and marine habitats on the central and western Pacific Ocean.* This theme seemed to flow through all presentations and conversations during the review. The PIFSC was established from an earlier organization in 2003 and so while relatively new, administers some well-established mature scientific programs. Presentations by staff exhibited considerable depth of understanding and competency regarding PIFSC programs and the issues regarding data and information management. The legacy of these programs has benefitted the PIFSC in accomplishing its historical mission; however, as noted above the mission has now expanded to include data integration and dissemination to a variety of stakeholders and the general public. This legacy may in part be a hindrance to meeting this new mission. The stand-alone data management practices of existing programs seem to be an impediment to cross-center sharing of data and evolution of technical capabilities.

Review Team Comments

The panel recognizes the considerable data management expertise distributed throughout PIFSC. The underlying theme of our comments is that many of the data and information

management challenges can be dealt with by addressing three critical areas.

- There is a need for stronger leadership for data and information management. From the outset the Director's support was evident; however, there was no clearly identified individual responsible for leadership within the Center. This was evident from the gaps in Center data policies to the disparate ways that the data are being managed.
- Communication seemed to be an issue. There was confusion about roles and responsibilities among staff and line organizations assigned a data or information management role and an apparent lack of cooperation between organizations regarding data and information management activities.
- The Center seems to have considerable expertise that could be leveraged to improve management both within and across the divisions. Collaboration should to be encouraged.

Stakeholder Session

The stakeholder session was well attended with nearly 20 stakeholder representatives present. The attendees included representatives from Western Pacific Regional Fishery Management Council, National Ocean Data Center, NOAA's Pacific Island Regional Office and the Papahānaumokuākea National Marine Monument among others. The overall impression from the attendees was that the PIFSC had considerable expertise and data critical to meeting their needs; however it is difficult at times to gain access to the data in a timely fashion. Furthermore, it is not always clear what data may be accessed and the accessibility of the data seems to vary depending on who is contacted. The current process is too dependent upon the requirement to know the center contact for each data set and the availability that individual to respond to each request. Some specific examples as well as specific comments about data follow:

- Current requests for data are met by providing spreadsheets, this solution is acceptable but seems far too reliant on personal knowledge of who to contact;
- It would be helpful to have access to Pacific Island recreational fish data directly rather than having to go through headquarters;
- Dependence on contractors with a relatively high rate of turnover has been a challenge;
- Data on pelagic species was excellent, non-pelagic and recreational data were weaker; and
- Access to detailed catch statistics would be appreciated.

Most who spoke out requested digital access to data rather than having to make a specific request to individual(s) and they would like to have one central point of contact for assistance when necessary.

Relationship of current and planned data management activities to Center mandates and requirements – is the Center doing the right things?

The center is doing many things right including:

- Reflecting the key role for PIFSC in sharing science data and information as a part of the mission;
- Emphasizing data documentation as a key enabler to that mission;
- Moving toward a center-wide database and enterprise data management; and
- Collaborating with NMFS-wide committees (e.g. Fisheries Information Management Advisory Committee (*FIMAC*), Fisheries Information System (*FIS*) and Geographic Information Systems (*GIS*)).

However, there is a need to articulate a new framework to guide the evolution of the PIFSC data and information enterprise. This framework will help PIFSC conform to NOAA mandates, optimize resource use and provide critical support to scientific staff and external customers.

Recommendations:

Establish the position of a chief data officer or CDO (Band 5) – This would be a management position responsible for defining and implementing the enterprise-wide data management framework. The role includes developing strategic priorities for the Center in the area of data systems and representing data management as a critical asset within the Center’s executive body. The scope would include the full life-cycle of information beginning with data documentation, through processing and to archive.

As an interim step, the Center should quickly complete and issue policy documents that have been underdevelopment for some time and assign unambiguous responsibilities across divisions and staff.

Opportunities - are there opportunities that the Center should be pursuing in data management systems, including shared approaches with partners?

There are many opportunities for leveraging data management expertise within PIFSC. In addition there are benefits to building partnerships with other NOAA offices and beyond. The panel recognizes a critical need to improve collaboration across PIFSC programs and suggests this as a short-term emphasis.

Recommendation:

PIFSC should establish a data management working group for identifying existing expertise and facilitating convergence towards an enterprise data management approach. The working group should be chaired by the CDO who should also select the members.

Once the working group begins operations, efforts should be made to expand the already good relationships within the agency and with local partners to address common goals and leverage opportunities.

The panel also strongly recommends that this group be an action oriented group given enough latitude to implement programs and projects to improve data management throughout its life cycle.

Scientific/technical approach – are the Center’s data management objectives adequate, and is the Center using the best suite of techniques and approaches to meet those objectives?

Center-wide data management objectives were not elucidated during the review. Based on the information provided, the panel concluded that the objectives are program centric and focus is on serving needs of individual scientists/projects or organizations. This “single purpose” approach to data management has resulted in a collection of disparate data sets that are difficult to manage or combine with other data.

Recommendations:

The center needs to develop clear enterprise data management objectives, priorities and processes. These priorities should include the development of data management plans as a part of new project proposals and accountability measures must be put in place to ensure that once projects are funded the data management component is implemented according to the plan. Furthermore, legacy database tools being used at PIFSC (FoxPro and dBase) are high risk and should be replaced as soon as possible.

Last, many of the staff suggested or expressed distrust of the emerging Oracle-based data system. This suggests a critical need for educating center staff on the value of an enterprise approach to data management including: 1) benefits to their individual efforts; and 2) the potential for addressing the more complex scientific questions (e.g. ecosystem based management) through improved interoperability of data. It is also critical to provide staff training in the technical tools necessary to fully exploit the capabilities of the Oracle database system. (See Appendix 3)

Organization and priorities – is the Center’s data management system properly organized to meet its mandates and is the allocation of resources among programs appropriate?

The division of responsibilities between the two center-wide data management entities, Information and Technology Services division (ITS) and the Scientific Information

Services division (SIS), is unclear to the review panel and many PISFC staff. In addition, the science divisions have independent data management groups with little or no collaboration across divisions or with ITS or SIS. (See Appendix 3)

Recommendations:

New data management tasks in the Science Divisions should be designed in coordination with the CDO and the data management working group to further the enterprise data management system. Science division staff should be specifically tasked with building bridges to enterprise systems rather than creating more stovepipes.

The Systems Design Team and GIS staff currently within ITS should report to the CDO. This would provide a simplified and clearer emphasis on data services and support services and a cleaner separation from information technology and security services.

Scientific conduct – are the Center’s data management programs being conducted properly (integrity, peer review, transparency, confidentiality, PII, etc.)?

There was considerable concern with data security and integrity, and with protection of confidential data. Transparency and peer review of data management programs across the center need to increase significantly. There was a concern expressed about performing data and metadata management inconsistently across divisions. Additionally, the Center Director observed the potential value of applying scientific strategies to data management. The review panel recognizes the value of data documented properly and consistently across the divisions. As an example, one National Marine Fisheries Service Science Center has established a metadata management performance element (that requires consistent, centralized data documentation), which is included in performance plans for science staff.

Recommendations:

The Center should establish a data documentation performance element for all relevant staff. Additionally, a policy should be established requiring all supervisors to address this element during performance reviews and recommendations for increases in pay for performance. This directive approach should complement the recommendation for educating staff on the benefits of proper data management both to individual projects and to the Center and training staff to use data management tools that are articulated on page 4.

The CDO, and data information management working group should develop and maintain policies and procedures for the management and sharing of data within and across divisions. These should reflect best business practices and be consistent with Center and NOAA policies for data management and access. One critical goal should be

to provide the broadest access possible to data to meet the needs of stakeholders and the public. Division heads should be held accountable to ensure that their staff implement these policies.

Concluding Comments

PIFSC impressed the panel with the depth and breathe of its mission and the number of staff that are dependent upon data to support their work or respond to stakeholder and public needs. It was also evident that there was considerable variation in staff expertise and support for data management activities. The overriding gap seems to be the lack of a policy and operational framework for data management within the PIFSC to guide data management throughout the Center. Recommendations are provided to strengthen data management by identifying a senior individual responsible for data management as well as providing critical support staff to aid in identifying and implementing policies and solutions.

Appendices

1. Review Agenda
2. Review Prospectus
3. Suggested Tools for Data Management
4. Considerations for Fisheries Data Management

**PIFSC External Program Review Agenda
Scientific Data Management Systems and Capabilities
July 20-22, 2010 Honolulu, Hawaii
East West Center - Imin Conference Center - Koi Room
Day I - Tuesday, July 20th**

- 8:00 am** **Hotel Lobby for pickup**
- 8:30 am** **Continental Breakfast and informal introductions**
- 9:00 am Welcome, introductions and purpose of review (Pooley)
Overview of data policies/requirements (Pooley/Shoji)
- 9:30 am PIFSC Overview (Seki)
- 10:30 am** **Break**
- 10:45 am Division/Program Overview
- Information and Technology Systems - ITS (Tieman)
Scientific Information Services - SIS (Wetherall)
DO Socio-economics (Malloy)
Coral Reef Ecosystem Division - CRED (Wong)
- Noon** **Lunch brought in (Paradise Café) ...** slideshow of photos from CRED
field research
- 1:00 pm Division/Program Overview continued ...
- Ecosystems and Oceanography Division - EOD (Polovina)
Fishery Biology and Stock Assessment Division - FBSAD (Boggs)
Fisheries Monitoring Division - FMD (Hamm)
Protected Species Division - PSD (Baker)
- 2:45 pm Panel discussion with Stakeholders without PIFSC staff
- Focus on “What works for you; what doesn’t work; constructive
suggestions for improvement?”
- 3:15 pm** **Break**
- 3:30 pm Stakeholder session continued ...
- 4:50 pm Walk to Dole St for reception
- 5:00 pm** **Dole Street Reception**
(all program leaders and data management POCs invited)

**PIFSC External Program Review Agenda
 Scientific Data Management Systems and Capabilities
 July 20-22, 2010 Honolulu, Hawaii
 East West Center - Imin Conference Center - Koi Room
 Day II - Wednesday, July 21st**

7:30am	<u>Hotel Lobby for pickup</u>	
8:00 am	Coffee service	
8:30 am	Data Catalog Services (Tokita)	(45m)
	Scientific Information Services Data Projects (Miyamoto)	
9:15 am	PIFSC Systems Design Team (Sender/Pappas)	(60m)
10:15 am	Break	
10:30 am	Bottomfish Integrated Product (Parke)	(15m)
10:45 am	Kona Integrated Ecosystem Assessment (IEA) Project (Howell)	(15m)
11:00 am	Fishing Ecosystem Analysis Tool (FEAT) (Richmond)	(15m)
11:15 am	OceanWatch – Central Pacific (Moxey)	(15m)
11:30 am	International Data Coordination and Reporting by NOAA Fisheries (Ito/Tagami)	(20m)
11:50 am	Lunch brought in (Kaka’ako Kitchen)	
1:20 pm	Economic Data Collection (Pan)	(20m)
1:40 pm	Life History and Bio-Sampling (Humphreys)	(25m)
2:05 pm	Fisheries Monitoring Division: An Overview of Data Management (Lowe)	(30m)
2:35 pm	Oceanographic Applications (Polovina/Domokos)	(25m)
3:00 pm	Break	
3:20 pm	Marine Turtle Research Program (Hargrove)	(20m)
3:40 pm	Monk Seal & Cetacean Programs (Khurana)	(20m)
4:00 pm	Managing Data from an Integrated Ecosystem Monitoring Program (DesRochers/Kanemura)	(45m)
4:45 pm	PIFSC wrap-up: Lessons learned, points to re-emphasize (Pooley)	
5:00 pm	Wrap-up of Day II – finalize tomorrow’s PIFSC program site visits for data system applications	
6:30 pm	Group Dinner - Reservations at Chai’s Island Bistro	

PIFSC External Program Review Agenda
Scientific Data Management Systems and Capabilities
July 20-22, 2010 Honolulu, Hawaii
East West Center - Imin Conference Center - Koi Room
Day III - Thursday, July 22nd

- 6:00 am** **Hotel Lobby for Fish Auction pickup**
- 6:30 am Optional trip to Honolulu Fish Auction – Breakfast at Nico’s
(Wear closed-toed shoes! Bring spare shoes/shirt you will smell slightly fishy)
- 8:30 am PIFSC program site visits (as determined by review panel)/Informal discussions between external review panel members and PIFSC staff
- 12:30 pm** **Working lunch at Dole Street (Natsunoya Tea House)**
- 2:30 pm private discussion amongst external review panel members
External review panel members’ comments and discussion with Division Chiefs
- 4:00 pm Private meeting between external review panel and PIFSC Directors
- 5:00 pm** **Review Completed**

**PIFSC External Program Review Prospectus
Scientific Data Management Systems and Capabilities
East West Center - Imin Conference Center - Koi Room
1777 East-West Road, Honolulu, HI 96828
July 20-22, 2010 Honolulu, Hawaii**

Objectives

Review and evaluate the Center's current scientific data management regime; from soup to nuts:

- data collection systems
- database/archival systems
- data accessibility/usability
- data integration

Provide advice to the Center on the direction and quality of these programs

Questions

- Relationship of current and planned data management activities to Center mandates and requirements – are we doing the right things?
- Opportunities – are there opportunities that the Center should be pursuing in data management systems, including shared approaches with partners?
- Scientific/technical approach – are the Center's data management objectives adequate, and is the Center using the best suite of techniques and approaches to meet those objectives?
- Organization and priorities – is the Center's data management system properly organized to meet its mandates and is the allocation of resources among program appropriate?
- Scientific conduct – are the Center's data management programs being conducted properly (integrity, peer review, transparency, confidentialityPII, etc.)?

Format

Scientific review panel: A panel of external scientists and data managers who will be responsible for the formal advice to the Center on the nature of our data management programs and systems.

Stakeholders: A broad variety of stakeholders have been invited to participate as observers including a first day discussion with the review panel. The stakeholder panel will provide the review panel and PIFSC with comments on how we are meeting their priorities related to information products and data sharing, and what issues they foresee.

Participants

Scientific review panel:

Mr. Dave Colpo – Pacific States Marine Fisheries Commission (PSMFC)
Dr. Mark Fornwall (Panel Chair) – Biological Resources Division, Director, USGS
Center for Biological Informatics
Dr. Ted Habermann - NOAA's National Geophysical Data Center
Mr. Jim Sargent – NOAA Fisheries Information Architect
Dr. Steven Smith, Research Scientist, University of Miami

Stakeholders (invited with the expectation they will choose their data managers and/or primary information and data recipients to participate):

Dr. Ned Cyr, Director, NMFS Office of Science and Technology
Dr. Steven Murawski, NMFS Director of Scientific Programs and Chief Science Advisor
Mr. Mike Tosatto, Acting Regional Administrator, NMFS Pacific Islands Regional Office
Allen Tom, Superintendent, ONMS Pacific Region
Naomi McIntosh, Superintendent, Hawaii Humpback Whale National Marine Sanctuary (Mr. Paul Wong)
Aulani Wilhelm, Superintendent, PMNM (Dr. Randy Kosaki, Deputy)
Dr. Kaylene Keller, PMNM Data Information Coordinator
Ms. Eileen Shea, Director, NOAA IDEA Center
Ms. Kristina Kekuewa, Acting Director, NOAA Pacific Services Center
Dr. Jo-Ann Leong, Director, UH Hawaii Institute of Marine Biology (Erik Franklin)
Dr. Jim Potemera, UH Faculty/Assistant Manager, Asia-Pacific Data-Research Center/PacIOOS Data Management Chair
Ms. Kitty Simonds, Executive Director, Western Pacific Fishery Management Council
Mr. Pat Caldwell, NODC Pacific Islands Regional Science Officer
Mr. Jeff Campbell, NODC CRCP
Ms. Sarah O'Connor, NODC CoRIS

TECHNOLOGY CHOICES THAT SUPPORT INTEGRATION

Ted Habermann

The ITS Division at PIFSC includes two groups that form the core of the center-wide data management services: the System Design Team (SDT) and the GIS Services (GIS). These groups have existed for some time and each has a number of well-known and significant achievements. At the same time, each group has limited resources and tool or approach preferences that have limited the breadth of their impact.

The resource issues can be ameliorated by engaging expertise from the Science Divisions in the collaborative task of building shared data management tools. At present, expertise in the Center is going in different directions, focusing on work-arounds that avoid the Center wide tools instead of making contributions that address specific needs while improving the foundation for everyone. In order to facilitate these contributions, the focus of ITS needs to shift away from convincing the Center to use tools that SDT and GIS have expertise in (i.e., Oracle Forms and ESRI Spatial Database Engine) towards using tools that the Center can use or is already using.

1. Excel and Access are ubiquitous “data management” tools in the Center. The scientific staff is clearly comfortable using these tools. There are a variety of ways to connect Excel spreadsheets and Access databases to Oracle. These need to be understood and exploited to create connections between desktop tools being used throughout the Center and the enterprise database. A second option for making these connections is Oracle Heterogeneous Services which allows users to access data stored outside Oracle as if they were in a local Oracle table. This approach would allow web interfaces to be developed to Access databases using the same tools being used for creating those tools on top of Oracle.
2. It is clear that Excel and Access can be used in ways that make integration with Oracle difficult. SDT must continue to encourage good data management and design practices throughout the Center with examples and training. The benefits of these practices will become more apparent as more connections are created and exploited to share data.
3. The Center is using ESRI’s Spatial Database Engine to implement spatial capabilities on top of Oracle. This choice makes sense for access to the database using ESRI clients like ArcMap and ArcGIS Server, but it creates a significant obstacle to data integration using spatial capabilities in the underlying database. Oracle Locator or Oracle Spatial can be used together with ArcSDE to create a broad and powerful foundation (see <https://www.ngdc.noaa.gov/wiki/index.php?title=SpatialCookbook>)

that supports access with a variety of commercial and open mapping and analysis clients.

4. Getting data out of the Oracle is a challenge for many PIFSC staff. The tools that are being used (Oracle Forms) and developed (Data Trawler) by SDT are powerful tools for experts, but intimidating barriers for most PIFSC staff. Oracle Application Express (Oracle APEX) is a rapid web application development tool for Oracle. It makes it possible to create fast and secure web applications using only a web browser and limited programming experience. It is a fully supported, no cost option of the Oracle database which would revolutionize Oracle usage at PIFSC by engaging the entire community in building interfaces to their own data.

The approaches described above can fundamentally change the data management and integration equation at PIFSC by connecting capabilities throughout the Center with an enterprise data management foundation. Work-arounds would become contributions to a collection of capabilities supported by the entire community and bottlenecks to data sharing would disappear.

CONSIDERATIONS FOR FISHERIES DATA MANAGEMENT

Steven Smith

Overview

Two divisions, Fishery Monitoring and Fishery Biology/Stock Assessment, form the core of PIFSC's mission to provide scientific information on the sustainable use of fisheries resources to managers and policymakers. Fisheries under the purview of PIFSC include:

- highly migratory species (tunas, billfishes)
- bottomfish (deepwater snappers and groupers)
- shallow-water reef-fishes

Together, the two fisheries divisions collect/assimilate, organize, maintain, summarize, and analyze the fundamental data (e.g., catch, effort, size composition, life history tissues) for evaluating the sustainability of these fish stocks. They also distribute data summaries and analysis results to regional fishery regulatory agencies/entities.

At present, data on the various fish stocks are primarily fishery-dependent, i.e., the data are collected during the course of fishing operations, and come from a wide variety of sources including logbooks and reports provided by fishers, scientific observers placed on fishing vessels, seafood dealers, dockside intercept sampling, etc. Some of this information is directly generated by PIFSC personnel, and the remainder is provided by partner agencies; however, data assimilation and management mostly occurs within PIFSC. PIFSC scientists also conduct fishery-independent monitoring studies for certain fish stocks, e.g., bottomfish. These data are exclusively maintained at PIFSC.

General Inefficiencies

There are some general inefficiencies with the current system for fisheries data management at PIFSC. These occur at three hierarchies or levels of information:

1. Elemental Databases - these are the fundamental, record-level databases containing information such as species-specific catches by license holder, date, location, gear, etc., from which all data products and analyses are derived. For the major types of fishery-dependent information (e.g., longline observer data, fisher-reported catch-effort), the elemental databases are fairly well-organized but are mostly maintained as separate entities and thus not part of a centralized, center-wide enterprise database system. A key source of frustration to PIFSC scientists responsible for analyzing these data is the difficulty of accessing the various types of fishery-dependent

information. For example, there may be one resident expert for a particular type of data who is responsible for extracting and providing information to analysts within PIFSC, rather than a system being in place for analysts to retrieve data for themselves. While great strides have been made in recent years to catalogue all of the extant fishery-dependent data at PIFSC, knowing that it exists is a far cry from being able to access it in a timely manner.

These problems are much worse for fishery-independent information, which are mostly housed and maintained by an individual researcher on a particular computer. Thus, in addition to the inaccessibility of this information to other scientists, there is a risk that the data may be lost when the scientist retires or leaves PIFSC for another job, or when a hard drive crashes.

2. Analysis-Ready Datasets - these are datasets derived from the elemental databases described above. In many cases, there is a large amount of work involved (weeks to months) after extracting data from the elemental databases in preparing an input dataset for producing various types of summaries, e.g., quarterly catch-effort by species and gear for highly migratory species, or for conducting stock assessments. These processing tasks go well beyond SQL queries, and usually involve procedures for standardizing units of catch (numbers, weight) or effort (trip-day, hook-hour, etc.), streamlining numerous specific gear types into major categories, identifying valid zero-catch trips to be included in fishing effort computations, and so on. For the case of stock assessments, fisheries data from a variety of sources are typically integrated to produce analysis-ready datasets, e.g., combining catch-effort and length-composition data to produce an abundance-at-length time-series. Many of these procedures require a substantial amount of analysis in their own right before they are finalized, usually involving specialized data processing software (e.g., SAS).

At present, the vast majority of analysis-ready datasets at PIFSC used in producing data summaries and stock assessment analyses reside with an individual investigator, and are not well documented in terms of either the variables, formats, units, etc., comprising the dataset, or the processing procedures used to create the dataset from the various elemental databases. This lack of accessibility and documentation can lead to some major problems:

1. Due to the complexity of the processing and analysis procedures needed to create analysis-ready datasets for inputs to stock assessments, it is likely impossible to exactly reproduce a particular dataset by anyone other than the analyst who originally created it.

2. Following from (i), stock assessment duties may transfer to another scientist at some point, and this new analyst will need to ‘re-invent the wheel’ with respect to creating analysis-ready datasets.
3. Also following from (i), stock assessments often come under review by independent scientists or are the subject of legal challenges. The ‘black box’ of the origin of the input data to mathematical/statistical assessment models usually turns out to be the biggest problem in these scientific reviews and court cases.
4. Data Summaries and Analysis Products - these are produced from either elemental databases or analysis-ready datasets, and are the major types of fisheries scientific information provided by PIFSC to scientific and regulatory stakeholders, including the State of Hawaii and the Western Pacific Fisheries Management Council. At present, stakeholder requests for this type of information occur at a frequency of several per week, and each request is handled individually by a specific data manager. The lack of direct access to summarized fisheries data and stock assessment analyses is perhaps the number one frustration of PIFSC’s stakeholders.

It is likely that a special summary database system could be developed that would satisfy the majority of stakeholder requests, greatly reducing the staff time currently spent on dealing with these requests. The access and documentation problems described above for data levels (1) and (2) are the biggest roadblocks in developing such a summary database system.

Looking Towards the Future

At present, the responsibilities for maintaining and managing the primary fishery-dependent databases for highly migratory species and bottomfish take up nearly all the time of the present fishery data management staff. Looming on the near horizon, however, are a host of new data management responsibilities pertaining to the shallow-water reef fish complex. These responsibilities include setting up data collection systems in remote places (e.g., American Samoa, Guam, Northern Mariana Islands). The inefficiencies outlined above will make it very difficult for the current overstretched staff to meet the new data management needs for shallow-water reef fisheries.

Looking just beyond the horizon, it is likely that fishery-independent information will play a much more important role in assessment and management of certain fisheries, especially bottomfish and shallow-water reef fishes. Recent regulatory actions involving seasonal and spatial closures of fishing activities have resulted/will result in large portions of these fish stocks being off-limits to fishers; consequently, fishery-dependent data will no longer provide fully representative information on the exploitation status of

certain fish stocks. There will thus be a growing need to conduct fishery-independent surveys to provide this missing information. For the case of shallow-water reef fishes, these surveys will undoubtedly involve scientists from the Coral Reef Ecosystem Division as well as from the two fisheries divisions. PIFSC needs to begin preparing for these future increasing data management responsibilities for fishery-independent information.

Looking out a little further, management concerns of fisheries in general are now expanding beyond sustainable rates of exploitation for target species to include the impacts of fishing on ecosystem trophic structure and food-web dynamics. In many coastal fisheries, there is also high concern over non-fishing human threats to the productivity of fish stocks from habitat and water quality alterations. The potential impacts of global climate change on fisheries productivity will also need to be addressed in the coming years. This new ecosystem-oriented perspective is spurring the evolution from traditional, single-species assessment approaches towards incorporation of predator-prey dynamics and environmental features into mathematical models for evaluating the productivity and sustainability of multiple stocks occupying multiple trophic levels and subject to a variety of meteorological-oceanographic conditions. Development of these new assessment approaches and models will require collaborative teams of scientists (e.g., fisheries biologists, mathematical demographers, statisticians, physical oceanographers) with the ability to obtain, share, and integrate large amounts of disparate data on fishes, fisheries, and the physical environment. Without question, the present data management capabilities needed to efficiently handle fishery-dependent and -independent information pale in comparison to the information systems that will be required to face this brave new world of fisheries ecosystem assessment and management.

Solution Ideas and Priorities

Some ideas for improving the efficiency of fisheries data management at PIFSC are given in priority order.

1. Primary Fishery-Dependent Databases - The first priority are the primary fishery-dependent databases used to conduct stock assessments and provide data summaries and assessment results to stakeholders responsible for management policies. Major tasks include: (i) transitioning the elemental databases into the center-wide enterprise database system, along with the necessary tools for easy input and extraction of data into and from the database; (ii) creating tools to facilitate production of standard data summaries and assimilation into a summary database system with easy access for relevant stakeholders; (iii) clearly documenting the various databases and

intermediate processing steps used in creating summary products to make the origins of any dataset or data product transparent to users.

This will be an enormous undertaking; however, the central importance of these fishery-dependent databases to the core mission of PIFSC cannot be overlooked. The present fisheries data management staff will need considerable assistance with these tasks due to limitations in both time and expertise. In the short term, the near-full attention of the current SDT and future Chief Data Officer is warranted to address these needs. One possibility for additional help is to recruit 2-3 graduate students with technical data management expertise at the University of Hawaii to work under the SDT to help with the transition. The principal data users (e.g., stock assessment analysts, management council staff) need to be involved at various stages of the transition process to ensure ease of access to databases and data summary products.

2. Analysis-Ready Datasets for Assessment - The transition process for fishery-dependent data described above includes documenting the procedures for producing standard types of data summaries. A parallel activity needs to begin that documents the creation of analysis-ready datasets used to conduct stock assessments, including data processing code and description of intermediate datasets and analyses. The documentation should be thorough enough so that an independent analyst can easily reproduce a given analysis-ready dataset from the original source databases, and it should be clear enough so that the main processing steps/decisions are transparent to independent scientific reviewers. Emphasis should be given to ongoing stock assessments for highly migratory species and bottomfish, and any future stock assessments. While this task will add time to the current process for conducting stock assessments, it will save immeasurable amounts of time in the future and greatly improve the transparency and credibility of the entire assessment process. The final analysis-ready datasets should be added into the center-wide enterprise database system.
3. Fishery-Independent Databases - Once the transition process for fishery-dependent information is well underway, a similar project needs to be undertaken with fishery-independent databases for ongoing monitoring surveys and those in the planning stages. Many of the tasks and tools already developed for fishery-dependent data will likely be applicable to fishery-independent monitoring data. This will need to be carried out primarily by the SDT-grad student database experts in conjunction with the key scientists responsible for the various fishery-independent monitoring programs at PIFSC.

4. Secondary/Historical Fisheries Databases - As the primary fishery-dependent and fishery-independent databases are transitioned into the center-wide enterprise data system, attention of the SDT-grad student group and the fisheries data managers can be turned to some of the smaller, secondary fishery-dependent datasets and historical fishery-independent monitoring data that are needed to support ongoing/future stock assessments. A working group of analysts should make the judgment as to which secondary/historical databases are truly relevant to the current or near-future assessment needs of PIFSC.