Information Management and Quality Assurance for Resource Management Data Collection Efforts

Craig Palmer, Harry Reid Center for Environmental Studies, University of Nevada–Las Vegas, 4505 Maryland Parkway, Box 454009, Las Vegas, Nevada 89154-4009; palmerc@unlv.edu
Mark Sappington, Lake Mead National Recreation Area, 601 Nevada Way, Boulder City, Nevada 89005; mark_sappington@nps.gov

Development of an Improved Data Management System at Lake Mead National Recreation Area

Good data management is achieved, in practice, when data that have been collected and archived are recognized for their high quality, are readily accessible, and contribute to the intended purpose of the project, such as resource management decisions, regulatory processes, scientific research, or interpretive and educational needs. Good data management also fosters recognition by the scientific community that translates into increased research funding and scientific credibility.

In October 1999, Lake Mead National Recreation Area began working with the Harry Reid Center for Environmental Studies at the University of Nevada–Las Vegas to improve the organization and use of Lake Mead's resource management data and information. The first phase in this multi-phase project was an assessment of the current status at Lake Mead and development of a framework for data management. During this initial phase, four specific objectives were identified for any new data management system:

1. The data must be persistent, or long lived.
2. The data must be easy to locate and readily accessible.
3. The data must be of a quality and in a form that is usable, credible, and promotes knowledge to a variety of users both within and outside Lake Mead.
4. The system must accommodate resource managers' requirements for accountability.

The second phase of the project was development of a data management system to address these four objectives. Rather than beginning after data have been collected or a project has been completed, the new system starts with the conception and design of a research or monitoring project and continues until the desired end information product (data, report, map, etc.) is made available to the intended audience. This approach involves six distinct steps from project initiation to distribution of the project's findings: data design, collection, manipulation, analysis, archiving, and access.

Data design. Many potential difficulties in data collection, analysis, archiving, and distribution can be avoided when sufficient thought and effort are given to the data design and management process prior to data collection. With this in mind, the data design process begins with a project proposal detailing the purpose, methodologies, budget, references, and other aspects of the proposed project. This proposal is reviewed by two to three in-house scientists or managers and may be sent out for independent review if appropriate. After it is reviewed, the proposal must be approved by the chief of resource management at Lake Mead before work can proceed. This step addresses the data management objectives of data quality and accountability.

Data collection. In addition to actually collecting data, several practices must be followed during the data collection process. As data are collected, any changes to protocols detailed in the original proposal need to be documented. One of the most important aspects of data collection is ensuring that the data collected are of known and high quality. Consequently, data quality and assurance procedures must be followed during data collection. These procedures are detailed later in

this paper. In addition, data should be backed up as they are collected to avoid loss, and FGDC-compliant metadata (i.e., conforming to the content standard set by the Federal Geographic Data Committee) should be created for the data sets from details in the project proposal. This step addresses the data management objectives of data persistence, accessibility, and quality.

Data manipulation. After data have been collected, they often need to be manipulated before they can be analyzed. This process can include conversion to a different data format, standardization of data fields, organization into databases, and linking to other data. Numerous computer software tools can be used to accomplish these tasks, including databases, spreadsheets, and geographic information systems. This step addresses the data management objectives of data quality and usability.

Data analysis. During data analysis, data are summarized and formatted for delivery to their intended audience. Final products may include maps, reports, data summaries, raw data sets, and databases, among others. In addition, data analysis and product delivery should take place in a timely manner. This step addresses the data management objectives of data accessibility, usability, and resource manager accountability.

Data archiving. Archiving, or proper storage, of data allows potential data users the ability to access data and provides security against loss. Part of the data management procedures is to archive data both locally for internal users and externally for outside users and to ensure data security. Internally, data will be archived on the resource management data server for access by Lake Mead employees. Data on this server are protected onboard with a fault-tolerant hard drive system (i.e., RAID array), and data that frequently change are backed up on a daily basis using a high-capacity tape drive. Data that do not frequently change, such as images, are also archived on current-technology optical discs (e.g., DVD+RW). Externally, data will be archived in web-based databases (when appropriate, such as for non-sensitive GIS data), and tapes and disks will be stored in secure off-site storage. This step addresses the data management objectives of data persistence and accessibility.

Data access. The ability to easily locate data is as important as proper storage of data. For internal users at Lake Mead, data and metadata will be accessible through professional information management software, such as Synthesis. External users will be able to locate metadata for data sets by searching on-line metadata databases, such as those operated by the National Park Service (www3.nature.nps.gov/im/metadata/quicksearch.cfm) and the U.S. Geological Survey (mercury.ornl.gov/nbii/). Using the metadata, external users would then be able to determine if data would be useful to them and be able to request the data. This step addresses the data management objectives of data persistence and accessibility.

Integration with NPS data management tools. To ensure persistence and accessibility of data collected, the new data management system being implemented by Lake Mead is designed to integrate with existing NPS data management tools, such as Dataset Catalog, NPBib, NPSpecies, Database Template, and the GIS Theme Manager.

Legacy data. Since the new data management system is designed to be integrated with all current and new data collection projects, legacy data will have to go through a process of inventorying, prioritizing, re-formatting, cataloguing, and re-archiving to make them compatible with this new system. This process will ensure the persistence and accessibility of legacy data.

The third phase of the project involves implementing the new data management system by conducting and evaluating a pilot test of the system. Currently, pilot test projects are being conducted for project planning and data quality and assurance within each program area at Lake Mead. These projects include desert tortoise monitoring, *Rana onca* habitat studies, aquatic plant surveys, exotic plant management, arid land restoration, and bat monitoring within abandoned mines.

Elements of the Lake Mead National Recreation Area Quality Assurance System

To help achieve the goal of credible, persistent, assessable, and useful natural resource information, a quality assurance (QA) system is being developed for the resource management staff at the Lake Mead. The purpose of this section is to provide some background on the approach we are using in the development of this QA system.

Natural resource management agencies such as NPS often have limited experience with the establishment and implementation of formalized QA programs (Figure 1; Palmer 2003). In contrast, federal regulatory agencies such as the Environmental Protection Agency and the Department of Energy have very detailed and structured QA programs that must be implemented by their staff whenever they collect data. The experience of these agencies has been that QA not only assists in making their data more defensible in court, but also improves the likelihood of high-quality data that have been adequately documented so as to be persistent and assessable.

QA is an overall system of management activities designed to assure the quality of data and information that are generated by a project or program. The two principal components of QA are quality control and quality assessment. Quality control includes those operational techniques and activities that are used to control the data acquisition process. Quality assessment or evaluation includes the application of statistical tools to determine the uncertainty in the data and whether or not they are appropriate to support management decisions. For example, the precision and bias of measurements can be estimated to identify if measurements should be considered quantitative, semi-quantitative, or qualitative.

The approach used to develop the QA system for Lake Mead was to follow the American National Standard (ANSI 1994), which provides specifications and guidelines for quality systems for environmental data collection programs. The value of selecting this approach is that it is based on the extensive experience of a large group of QA professionals and is currently used as the common standard for the development of QA programs by many agencies. This standard requires that a QA program be developed in two main parts. The first part is to specify the quality management elements for the overall program. This is documented in a quality system management plan. The second part is to specify QA ele-

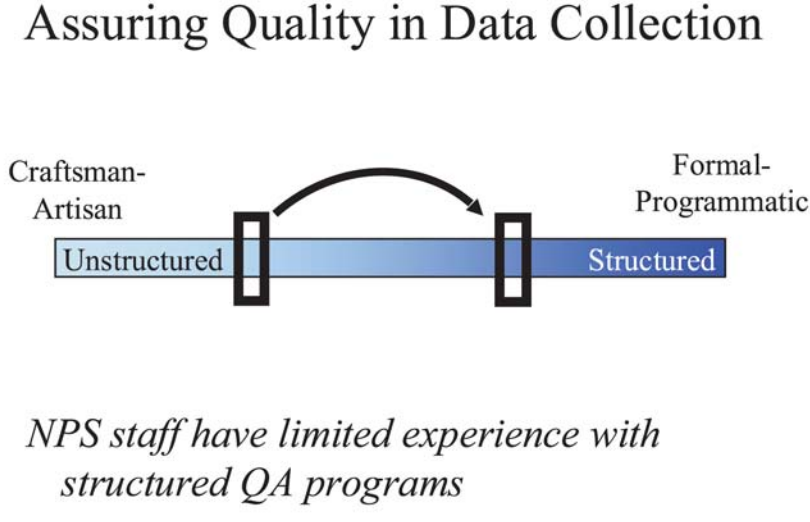

Figure 1. Quality assurance in data collection requires moving from unstructured to structured programs.

ments that should be included in any data collection effort. These QA elements should be included as part of the planning for data collection in any given project.

The development of the QA program for Lake Mead began with the preparation of a draft quality system management plan (Palmer and Landis 2002) for consideration by park staff. In accordance with the guidelines from the American National Standard, ten topics were addressed in this plan: management and organization, quality system description, personnel qualification and training, procurement of items and services, documents and records, computer hardware and software, planning, implementation of work processes, assessment and response, and quality improvement.

The Lake Mead quality system management plan details a QA system to be implemented whenever natural resource information is collected. This QA system is detailed in Table 1. The table is divided into QA activities that should be undertaken during the planning, data collection, assessment, and continual improvement phases of each project. Each of these topics will be considered in more detail in the following paragraphs.

Planning. The primary project planning tool for the quality system is the resource management project plan (RMPP). During the process of preparing a RMPP, a project leader answers a specific list of questions. A unique property of this RMPP is that it includes all the information needed to complete a fully compliant FGDC metadata record.

An important component of the RMPP is the identification of each of the measurements that will be undertaken during the project study period. For each of the measurements, the project leader is asked to specify a measurement quality objective (MQO). For example, an MQO might be set for the measurement of the width of a tortoise shell, frequency of correct identification of plant species, range in acceptable condition codes, or the concentration of a chemical in a water sample. The development of MQOs is a critical QA step as it serves as the basis for evaluating and improving the quality of data over time.

Data collection. Data collection in projects should follow written protocols called *standard operating procedures.* Field crew members should be trained in these proce-

Table 1. Quality assurance activities, tools, and responsibilities for the Lake Mead Resource Management Division quality system

Project Phase	Activity	Tools	Responsibility
Planning	Develop a project plan with a QA section	Resource Management Project Plan (RMPP)	Prepare: Project Leader Review: QA Team
	Select measurement quality objectives (MQOs)	RMPP	Prepare: Project Leader Review: QA Team
Data Collection	Develop detailed methods and data quality objectives	Standard operating procedures (SOPs)	Prepare: Project Leader Review: QA Team
	Conduct training and certification of trainees	Training guide and certification forms	Prepare: Project Leader Review: QA Team
	Collect, record, and control data	Scientific notebooks, field forms, data recorders	Prepare: Project Leader Review: QA Team
	Collect and control samples (if required)	Sample labels and sample handling procedures	Prepare: Project Leader Review: QA Team
	Calibrate and maintain field and laboratory equipment	SOPs	Prepare: Project Leader Review: QA Team
Assessment and Response	Conduct audits	Field audit form	Prepare: Project Leader Review: QA Team
	Remeasurements	Field data collection forms, remeasurement schedule	Prepare: QA Manager Conduct: Auditors, QA remeasurement crew
	Data review, verification, and validation	Data entry checks, illegal data filters, outlier detection, internal consistency checks	Prepare: Project Leader Program: Data Manager Conduct: Project Team
	Assess quality of data	Quality assessment section in project reports	Prepare: Project Leader Review: QA Manager
Continual improvement	Conduct annual reviews of project	Debriefing reports, client interviews; system audits	Prepare: Project Leader Review: QA team

dures and then tested as to their ability to perform them within the limits specified in the MQOs. Data collection should proceed using standardized field data collection forms or portable data recorders with built-in data collection programs. The advantage of using portable data recorders is that they minimize field data collection errors, such as missed fields or the entry of invalid codes. During data collection, all field equipment should be calibrated and maintained frequently.

Assessment and response. During the first few weeks of data collection, audits should be conducted of field crew members to ensure that they are following established protocols and to answer questions that might not have been adequately covered in training sessions. The purpose of conducting the audits early in the field season is to prevent the collection of erroneous or questionable data. During an audit or in a subsequent visit, independent remeasurements need to be taken of a subset of the data being collected by the field crews. When these data are collected during a field audit, they can be used to help identify problems the field crew might be having with the interpretation of field protocols. When they are collected at a different time without knowing the values obtained by the original crew, these remeasurement data can be used to calculate the precision and (in certain situations) bias in the data.

All data that are collected should be reviewed. The first step is to verify whether or not the numbers placed on the field data sheets have been correctly transferred to the project database during computer data entry. This step is called *data verification.* The next step is to evaluate whether or not the data are internally consistent and scientifically sound. This step is called *data validation* and includes evaluation for outliers and comparisons between parameters (Edwards 2000).

Continual improvement. An important component of any quality system is to have in place a process to improve the system over time. The approach recommended in the Lake Mead quality system management plan is to focus on debriefing of field crews at the end of the field season and to conduct annual reviews of on-going projects. The overall quality system management plan should also be reviewed on an annual basis.

Approach to implementation. The approach we have used to implement the quality system at Lake Mead has been to gradually implement the program through training and pilot studies. A day-long training session was used to introduce the staff to quality concepts and the overall approach. Each project leader was asked to select one of his or her projects to act as a pilot for QA during the coming year. Assistance has been provided to the project leaders to help them with the implementation of the quality system components, such as the preparation of RMPPs, the selection of MQOs, and the identification of opportunities for the collection of independent remeasurement data.

Summary

It is our belief that the formal planning of QA and information management systems will improve the likelihood that credible, persistent, accessible, and useful data will be collected by resource management staff in our national parks. This planning should begin with the preparation of a data management plan and a quality system management plan.

References

ANSI [American National Standards Institute]. 1995. *American National Standard: Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs.* Milwaukee: Energy and Environmental Quality Division, Environmental Issues Group.

Edwards, D. 2000. Data quality assurance. In *Ecological Data: Design, Management and Processing.* W.K. Michener and J.W. Brunt, eds. Malden, Mass.: Blackwell Science, 70–91.

Landis, E.B., and C.J. Palmer. 2002. *Lake Mead National Recreation Area Natural Resources Division: Data and Information Management Procedures.* Las Vegas: Harry Reid Center for Environmental Studies, University of Nevada–Las Vegas.

Palmer, C.J. 2003. Approaches to quality assurance and information management for regional ecological monitoring programs. In *Ecological Monitoring of Ecosystem Initiatives.* D. Busch and J. Trexler, eds. Covelo, Calif.: Island Press, 211–225.

Palmer, C.J., and E.B. Landis. 2002. *Lake Mead National Recreation Area Natural Resources Division: Quality System Management Plan for Environmental Data Collection Projects.* Las Vegas: Harry Reid Center for Environmental Studies, University of Nevada–Las Vegas.