National Park Service
U.S. Department of the Interior

National Capital Region Network

DATA MANAGEMENT PLAN FOR NATURAL RESOURCES IN THE NATIONAL CAPITAL REGION NETWORK

September 30, 2005

Submitted by:

Inventory and Monitoring Program
National Capital Region Network
Center for Urban Ecology
4598 MacArthur Blvd. NW
Washington, DC 20007

Editor:

Geoff Sanders, Data Manager, National Capital Region Network (2004 to present)

Preferred citation: National Park Service. 2005. Data Management Plan for Natural Resources in the National Capital Region Network. Inventory and Monitoring Program, Center for Urban Ecology. Washington DC.

Executive Summary

Collecting data on specific natural resource parameters is our first step toward understanding the ecosystems within our national parks. These ecosystems are evolving, as is our knowledge of them and how they work. We use these "raw" data to analyze, synthesize, and model aspects of ecosystems. In turn, we use our results and interpretations to make decisions about the Park's vital natural resources. Thus, *data* collected by researchers and maintained through sound data management practices will become *information* through analyses, syntheses, and modeling. This can only be achieved through the development of a modern information management infrastructure (e.g., staffing, hardware, software) and procedures to ensure that relevant natural resource data collected by NPS staff, cooperators, researchers, and others are entered, quality-checked, analyzed, reported, archived, documented, cataloged, and made available to others for management decision making, research, and education.

Data Management Mission

The mission of the NCRN Data Management Program is to support to the NCRN Inventory and Monitoring Program by promoting:

- Data quality – Ensure that appropriate quality assurance measures are taken during all phases of project development, data acquisition, data handling, summary and analysis, reporting, and archival.
- Security – Our objective is to make certain that both digital and analog forms of source data are maintained and archived in an environment that provides appropriate levels of access to project managers, technicians, decision makers, and others.
- Interpretability – Sufficient documentation should accompany each data set and any reports and summaries derived from it, so as to ensure that users will have an informed appreciation of its applicability and limitations.
- Availability – Our objective is to expand the availability of natural resource information by ensuring that the products of inventory and monitoring efforts are created, documented and maintained in a manner that is transparent to the potential users of these products.
- Longevity – The longevity of a data set can be enhanced by thorough documentation, by maintaining the data in a widely interpretable format, and by appropriate archival measures.

This document summarizes the NCRN data management strategy, which is more fully presented in the NCRN Data Management Plan (DMP; Sanders et al. 2005). The NCRN DMP serves as the overarching strategy for achieving the goals noted above. The plan supports I&M program goals and objectives by ensuring that program data are documented, secure, and remain accessible and useful indefinitely.

NCRN DATA MANAGEMENT STRATEGY

The NCRN data management strategy holds that all data and derived information generated or otherwise used by the program will meet a high level of quality standards. Further, all data and information the NCRN program deems necessary to meet its objectives, and that are not otherwise maintained, will be archived, documented, and made easily available and accessible. Data and information will be managed in a transparent manner such that all components

may be easily compared by location, time, and subject. Data and information will be accompanied by supporting documentation (metadata) that provide context, value, utility, and longevity, thereby facilitating broad understanding of NCRN program output to current and future end users.

Data Defined

The NCRN I&M Program works with data from various sources and defines different data types in this way:

Programmatic Data—*Data developed or acquired directly by the network as a result of inventory, monitoring, or other projects*. This category includes project-related protocols, reports, spatial data, and associated materials such as field notes and photographs. Also included in this category are collaborative efforts between the NCRN and other NPS or non-NPS entities.

Non-Programmatic Data—*Data developed by entities other than the I&M Program.* There are two types of non-programmatic data:

> NPS Data—*Data developed by other NPS entities.* The NCRN utilizes numerous datasets developed by parks (e.g., park based inventories or research, regional programs).
>
> Non-NPS Data—*Data acquired and/or maintained by non-NPS sources.* Datasets developed by other government agencies and/or non-government organizations. Examples of these data sources include air quality data from EPA, water quality data from county agencies, and remote sensing products such as satellite imagery and aerial photography.

DATA MANAGEMENT ROLES AND RESPONSIBILITIES

Collecting, analyzing and maintaining high quality data products require the combined efforts of numerous personnel. Therefore, proper data management standards must be understood and practiced by everyone involved in Network operations including regional staff, park staff and contractors and cooperators. Successful data stewardship requires that all Network personnel work together as a cohesive unit to ensure that data are collected using the appropriate methods and that all data sets are held to the highest quality standards. The table below lists the data stewardship roles and responsibilities of those involved in Network operation and data management.

CATEGORIES OF DATA STEWARDSHIP INVOLVING ALL NETWORK AND PARK PERSONNEL

Stewardship Category	Related Activities	Principal jobs or positions
Note: Each position is listed in only one category according to overriding responsibilities. However, most positions contribute in each category.		
Production	Creating data or information from any original or derived source. This includes recording locations, images, measurements, and observations in the field, digitizing source maps, keying in data from a hardcopy source, converting existing data sources, image processing, and preparing and delivering informative products, such as summary tables, maps, charts, and reports.	Project Crew Member Project Crew Leader Data/GIS Specialist or Technician
Analysis	Using data to predict, qualify, and quantify ecosystem elements, structure, and function as part of the effort to understand these components, address monitoring objectives, and inform park and ecosystem management.	Quantitative Ecologist Resource Specialist
Management	Preparing and executing policies, procedures, and activities that keep data and information resources organized, available, useful, compliant, and safe.	Data Manager Project Leader GIS Manager IT Specialist
End Use	Obtaining and applying available information to develop knowledge that contributes to understanding and managing park resources.	Monitoring Coordinator Park Managers and Superintendents Public and others

Although data management is a combined effort, the Network data manger is responsible for coordinating the Network's data management program to ensure that standards are met and guidelines adhered to in all Network activities. Achieving the high standards of data quality the data manager must work closely with Network project leaders. Project leaders are usually Network ecologists responsible for managing individual projects. They are tasked with ensuring that project data management meets the Network standards. Figure ES-1 illustrates some of the specific data management tasks involved with I&M projects and how the responsibility for those tasks is divided.

NETWORK INFRASTRUCTURE AND DATA MANAGEMENT ARCHITECTURE

The foundation of the NCRN data management program is composed of a combination of computer hardware connected over local and national networks as well as the applications, tools, and repositories run on these computer systems.

- *Infrastructure* – represents the computers and servers connected through computer networking services.
- *Architecture* – the applications, system tools, databases, and repositories that make up the program data management enterprise.

Our digital infrastructure has three main components: park-based local area networks (LAN), network data servers, and servers maintained at the national level. This infrastructure is maintained by park, regional, and national IT specialists, who administer all aspects of system security and backups.

The NCRN I&M Program relies heavily on regional and national IT personnel and resources to maintain the computer infrastructure as well as assistance in developing and augmenting the program's data management architecture. Therefore, communication with regional IT specialists is essential to ensure adequate resources and service continuity.

To achieve an integrated information management system, three of the national-level data management applications (NatureBib, NPSpecies, and NR-GIS Metadata Database) utilize application architecture with both desktop and internet-accessible (master) components (figure ES-2).

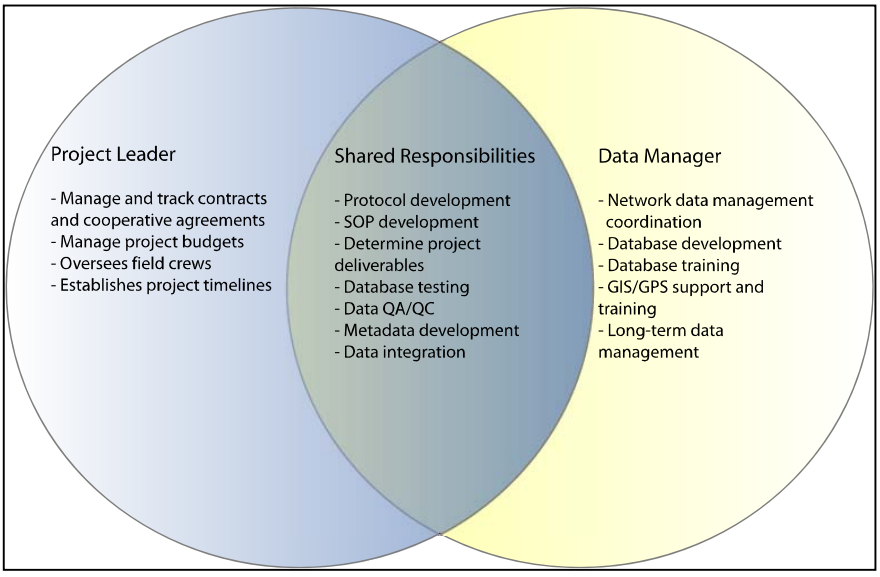


FIGURE ES-1: DIVISION OF PROJECT RELATED DATA MANAGEMENT TASKS
(adapted from the CAKN Data Management Plan)

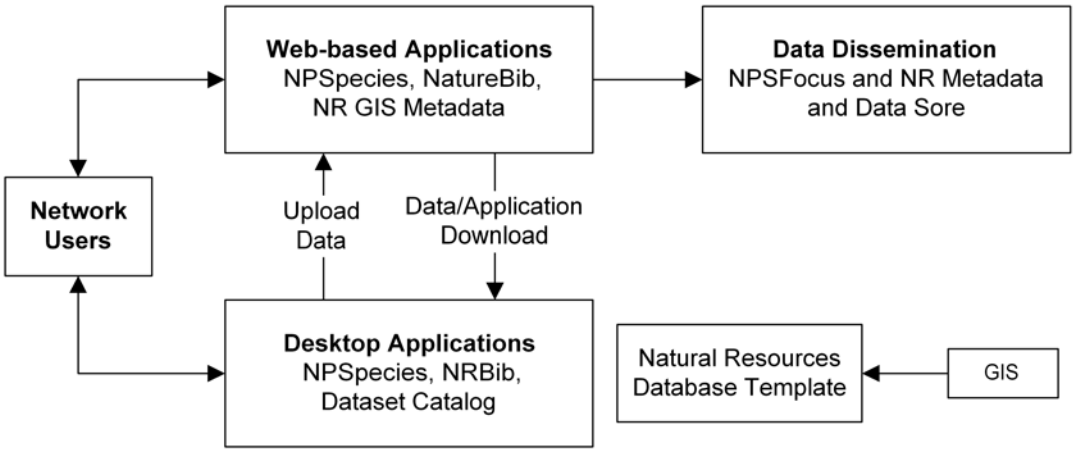


FIGURE ES-2: NATURAL RESOURCES DATABASE FRAMEWORK

An additional and integral component of the I&M data management infrastructure is the NPS STORET application for managing all data acquired during Network water quality monitoring. Water quality data collected as part of the network's monitoring program have distinct data management requirements. Data must be managed according to guidelines from the NPS Water Resources Division (WRD). This includes using the NPSTORET desktop database application at the parks to help manage data entry, documentation, and transfer to WRD. The NCRN will oversee the use of NPSTORET according to the network's integrated water quality monitoring protocol. Figure ES-3 illustrates the flow of information and data between the network, national program offices and EPA's main STORET repository.

Project Work Flow —All projects conducted by the NCRN have the same general workflow characteristics. Understanding the progressive stages of a project and the life cycle of the resulting data, we can more easily communicate the overall objectives and specific steps of the data management process. In addition, this awareness helps us to manage the staffing resources needed to produce, maintain, and deliver quality data and information. More details about data acquisition, quality assurance, documentation, dissemination and maintenance can be found in later chapters of this plan. Figure ES-4 illustrates the various project stages and highlights the data management tasks associated with each stage.

From the perspective of managing workflow, there are two main types of projects:

Short-term projects, which may include individual park research projects, inventories, or pilot work done in preparation for long-term monitoring.

Long-term projects, which will mainly be the implemented monitoring projects central to the I&M program, but which may also include multi-year research projects and monitoring performed by other agencies and cooperators. Long-term projects will often require a higher level of documentation, peer review and program support.

From a data management standpoint, a primary difference between short- and long-term projects is an increased need to adhere to standards for long-term projects to ensure internal compatibility over time. Both short-term and long-term projects share many workflow characteristics, and both generate data products that must be managed and made available.

DATA QUALITY ASSURANCE AND QUALITY CONTROL

Director's Order #11B states that all information (e.g., brochures, research and statistical reports, policy and regulatory information, and general reference information) distributed by the NPS (including information obtained from sources outside of the NPS) must be accurate, reliable and timely in nature. In order to disseminate accurate information, we must have confidence in the data we use. All data analyses, reports, and publications require data of

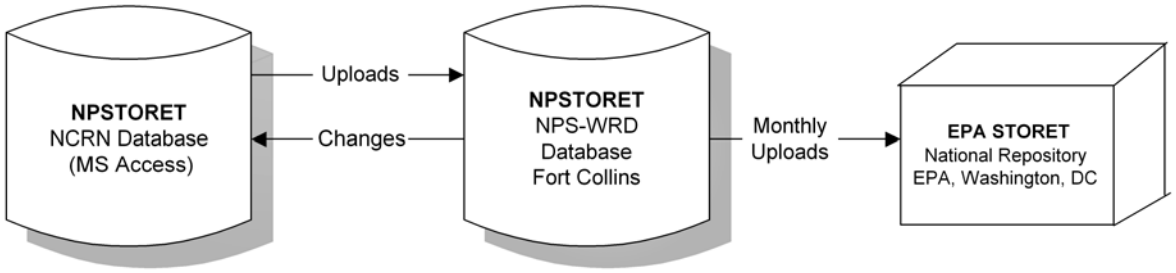


FIGURE ES-3: NATURAL RESOURCES DATABASE FRAMEWORK

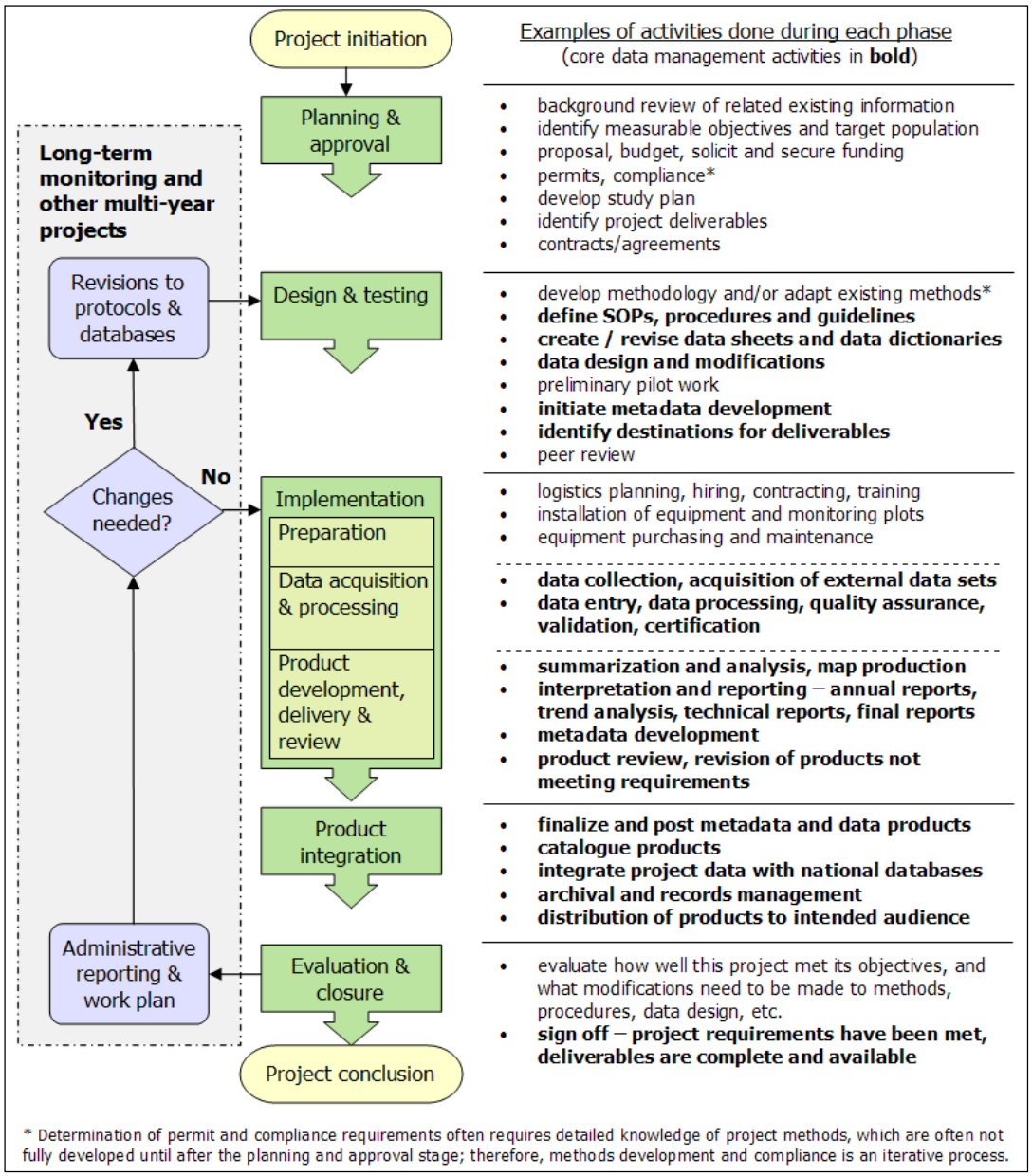


FIGURE ES-4: GENERAL STAGES OF PROJECT LIFE CYCLE

documented quality that minimize error and bias. Data of inconsistent or poor quality can result in incorrect interpretations and conclusions and improper management applications.

Palmer (2003) defines *Quality Assurance* (QA) as "an integrated system of management activities involving planning, implementation, documentation, assessment, reporting, and quality improvement to ensure that a process, item, or service is of the type and quality needed and expected by the consumer." He defines *Quality Control* (QC) as "a system of technical activities to measure the attributes and performance of a process, item, or service relative to defined standards." Quality Assurance procedures maintain quality throughout all stages of data development. Quality Control procedures monitor or evaluate the resulting data products.

QA/QC mechanisms are designed to prevent data contamination, which occurs when a process or event other than the one of interest affects the value of a variable and introduces two fundamental types of errors into a data set:

- *Errors of commission* — include those caused by data entry and transcription errors or malfunctioning equipment. They are common, fairly easy to identify, and can be effectively reduced upfront with appropriate QA mechanisms built into the data acquisition process, as well as QC procedures applied after the data have been acquired.

- *Errors of omission* — often include insufficient documentation of legitimate data values, which could affect the interpretation of those values. These errors may be harder to detect and correct, but many of these errors should be revealed by rigorous QC procedures.

We appraise data quality by applying verification and validation procedures as part of the quality control process. These procedures are more successful when preceded by effective quality assurance practices.

Data verification checks that the digitized data match the source data (i.e., does the database match the field data sheets?).

Data validation checks that the data make sense (e.g., is an ambient temperature reading of 120°F in the middle of winter really accurate?).

It is essential that we validate all data as truthful and do not misrepresent the circumstances and limitations of their collection. Failure to follow SOPs for data entry, validation, and verification will render a data set suspect. Although data entry and data verification can be handled by personnel who are less familiar with the data, validation requires in-depth knowledge about the data.

Communicating Data Quality — The Network will use data documentation and metadata to notify end users, project managers, and network management of data quality. A descriptive document for each data set/database will provide information on the specific QA/QC procedures applied and the results of the review. Descriptive documents or formal FGDC-compliant metadata will document quality for spatial and non-spatial data files posted on the Internet.

DATA DOCUMENTATION

Data documentation is a critical step toward ensuring that data sets are useable for their intended purposes well into the future. This involves the development of *metadata*. FGDC defines metadata as:

Metadata — information about the content, quality, condition and other characteristics of data.

Additionally, metadata provide the means to catalog datasets, within intranet and internet systems, thus making their respective datasets available to a broad range of potential data users.

Purpose of Metadata—Data sets sometimes take on lives of their own. Some seem to have the ability to reproduce and evolve on multiple hard drives, servers and other storage media. Others remain hidden in digital formats or in forgotten file drawers. In addition, once data are discovered, a potential data user is often left with little or no information regarding the quality, completeness, or manipulations performed on a particular "copy" of a dataset. Such ambiguity results in lost productivity as the user must invest time tracking information down, or, worst case scenario, renders

the dataset useless because answers to these and other critical questions cannot be found. As such, data documentation must include an upfront investment in planning and organization. Figure ES-5 illustrates the I&M Program's metadata system.

The importance for metadata is now universally accepted within the data management community. Recent legislative and policy efforts emphasize that data documentation must include and upfront investment in planning and organization.

Metadata Process/Workflow—The NCRN I&M Program deals with many different types of data sets from varying sources. The amount of documentation needed for different datasets may vary but the following is the basic steps to follow when documenting data sets.

Step 1. Identify Relevant Data Sets and Compile Pertinent Metadata

Step 2. Create Dataset Catalog Record

Step 3. Select Metadata Tool and Complete Record

Step 4. Make Information Available

DATA DISSEMINATION AND OWNERSHIP

The National Park Service defines conditions for the ownership and sharing of collections, data, and results based on research funded by the United States government. All cooperative and interagency agreements, as well as contracts, should include clear provisions for data ownership and sharing as defined by the National Park Service:

All data and materials collected or generated using National Park Service personnel and funds become the property of the National Park.

Any important findings from research and educational activities should be promptly submitted for publication. Authorship must accurately reflect the contributions of those involved.

Investigators must share collections, data, results, and supporting materials with other researchers whenever possible. In exceptional cases, where collections or data are sensitive or fragile, access may be limited.

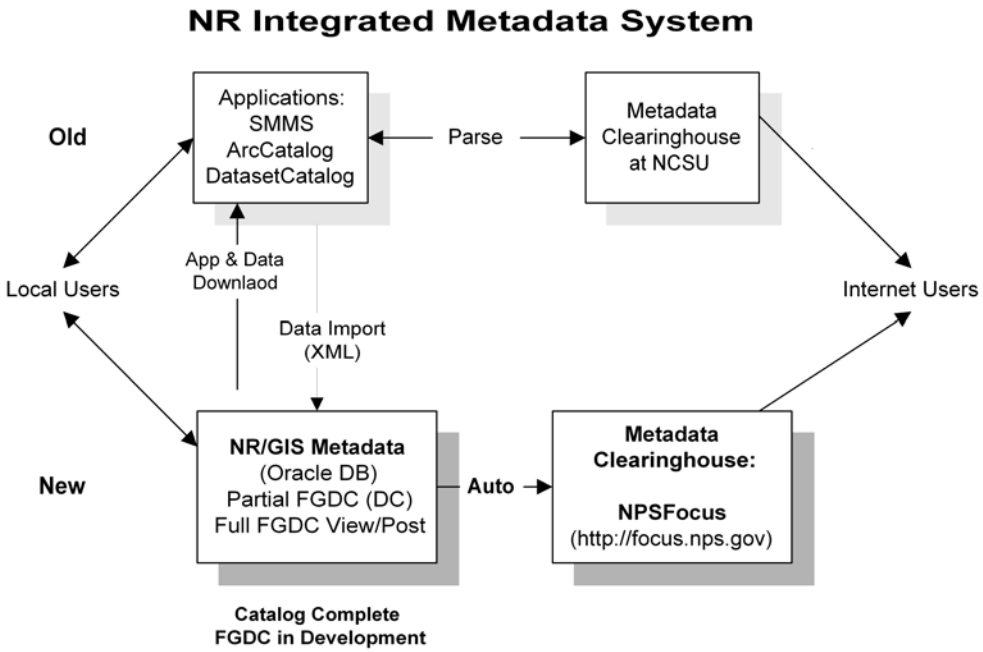


FIGURE ES-5: TAKEN AND MODIFIED FROM I&M DATA MANAGEMENT WORKSHOP, MARCH 2004

Data Classification: protected vs. public—All data and associated information from I&M activities must be assessed to determine their sensitivity. This includes, but is not limited to, reports, metadata, raw and manipulated spatial and non-spatial data, maps, etc. Network staff must carefully identify and manage any information that is considered sensitive. The Network must clearly identify and define those data needing access restrictions and those to make public.

The *Freedom of Information Act*, 5 U.S.C. § 552 (FOIA) stipulates that all US government agencies must provide access to data and information of interest, that are not protected from disclosure by exemptions, to the public. FOIA, as amended in 1996 to provide guidance for electronic information distribution, applies to records that are owned or controlled by a federal agency, regardless of whether or not the federal government created the records. Under the terms of FOIA, agencies must make non-protected records available for inspection and copying in public reading rooms and/or the Internet. Other records however, are provided in response to specific requests through a specified process. The Department of the Interior's revised FOIA regulations and the Department's Freedom of Information Act Handbook can be accessed at http://www.doi.gov/foia/ for further information. Network staff should consult their Regional FOIA Coordinator for more specific guidance whenever a particular instance of sharing information has the potential to involve protected information.

For example, information may be withheld regarding the nature and/or specific locations of the following resources recognized as 'sensitive' by the National Park Service. According to NPOMA, if the NPS determines that disclosure of information would be harmful, information may be withheld concerning the nature and specific location of:

- Endangered, threatened, rare or commercially valuable National Park System Resources (species and habitats)
- Mineral or paleontological objects
- Objects of cultural patrimony
- Significant caves

DATA MAINTENANCE

Data, documents, and anything that results from projects and activities that use Network data are all crucial pieces of information (figure ES-6). To ensure high-quality, long-term management and maintenance both digital and hard copy materials it is necessary to establish procedures that will permit a broad range of ers to easily obtain, share, and properly interpret both active and archived information while, at the same time, keeping the information secure.

Digital Data

In general, digital data maintained for lengthy periods of time will be one of two types: short-term data sets, for which data collection and modification have been completed (i.e., inventory projects); and long-term monitoring data sets, for which data acquisition and entry will continue indefinitely.

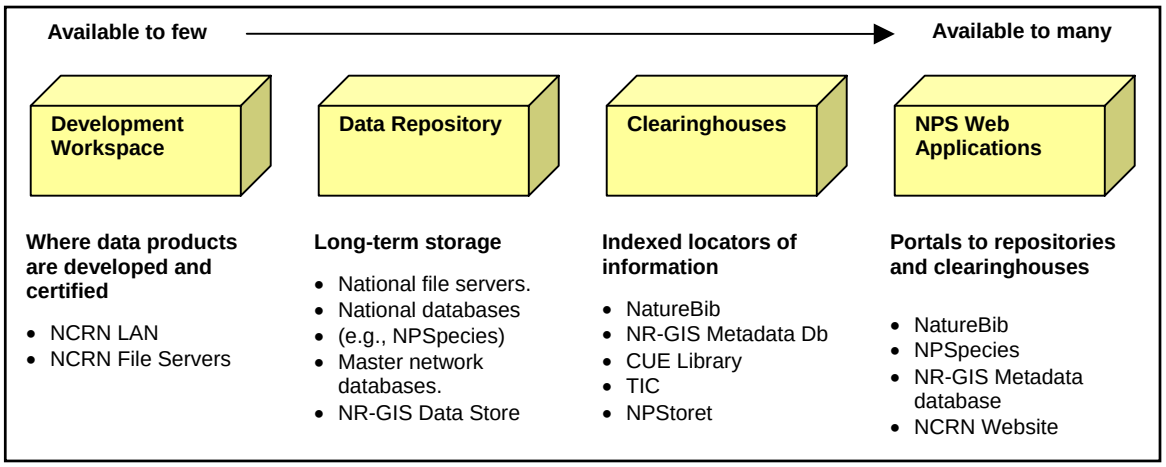


FIGURE ES-6: DATA DISTRIBUTION AND MAINTENANCE

Digital data will need to be stored in a repository that ensures both security and ready access to the data in perpetuity. All digital data sets are maintained on the NCRN file server under one of two directories depending on the depending on the status of the file.

Active Directory — used to house any file that is still in use or not yet considered complete. This includes datasets that have not yet undergone all of the necessary QA/QC procedures as well as draft documents.

Archive Directory — this directory contains all data sets that are complete (this mostly pertains to short-term projects such as inventories) along with completed reports, contracts and agreements. In the case of long-term monitoring data sets, because they will remain active indefinitely, the datasets will be archived periodically once the data have undergone QA/QC to preserve interim datasets. For data security purposes, this directory will be formatted as 'Read Only' for all personnel except the Data Manager and the Regional IT server support staff.

Digital Data Backup—The risk of data loss can come from a variety of sources, including catastrophic events (e.g., fire, flood), user error, hardware failure, software failure or corruption, and security breaches and vandalism. Performing regular backups of data and arranging for off-site storage of backup sets are the most important safeguards against data loss.

The NCRN I&M Program shares its file server with the National Capital Region Natural Resources and Science staff at the Center for Urban Ecology (CUE). Accordingly, backup procedures must be able to accommodate not only data produced by the I&M Program but all of the data and information stored by the regional staff as well. Due to the number of users utilizing the file server, the quantity of data included in the backup procedure quite large and constantly increasing.

Currently the NCRN backup plan calls for a complete daily backup of both the local directories and the shared directories (save two) that will enable a complete restoration of the file server for a period of thirty days. The two directories not included in the daily backup scheme are the GIS directory and the I&M Archive as the information in these directories changes infrequently. Data backups for these directories will take place once a week on a separate schedule.

Hard Copy Materials

The guidelines in this section apply to documents such as final reports prepared by staff or contractors, program administrative documents, contracts and agreements, memoranda of agreement, and other documents related to NCRN administration, activities and projects. These guidelines also apply to physical items such as natural history specimens, photographs, or audiotapes. In most instances, these documents and objects are essential companions to the digital data archives described earlier.

Documents—All paper documents managed or produced by the NCRN will be housed in one of two primary locations.

1. *NCRN central files*.
   The central files are maintained by the NCRN staff, at the Center for Urban Ecology. The NCRN will use acid-free paper and folders for all permanent records in the central files and store the documents in fire proof filing cabinets.

2. *Museum Resources Center (MARS)*
   MARS provides temperature and humidity-controlled facilities, a professional archival staff, and meets all museum standards set by NPS. This repository will be used for original documents and associated materials produced by the network (e.g., photographs, field notes, permits) that are a high priority to maintain under archival conditions.

For all materials submitted to MARS, NCRN will provide essential cataloging information such as the scope of content, project purpose, and range of years, to facilitate ANCS+ record creation and accession. NCRN will also ensure that materials are presented using archival-quality materials (acid-free paper and folders, polypropylene or polyethylene slide pages).

Specimens—It is recommended that all specimens collected under the auspices of the NCRN be archived at the regional Museum Resources Center (MARS); however, parks with the

proper storage facilities reserve the right to maintain their specimens in house. The NCRN will provide Park curators with associated data required for cataloging each specimen. These data will be in comma-delimited format (.CSV) format for automated uploading into ANCS+.

CONTENTS

CHAPTER 1: INTRODUCTION AND BACKGROUND

1. INTRODUCTION AND BACKGROUND 1-1

1.1 Purpose and Scope of the National Capital Region Network Data Management Plan 1-2

1.1.1 Data and Data Management: An Overview 1-2

1.2 The NCRN Inventory and Monitoring Program's Long-Term Data Management Goals 1-3

1.3 Types of Data Covered by the Information Management Plan 1-4

1.4 Information Management Goals and Objectives 1-4

1.5 Scope of the Data Management Plan 1-5

1.6 Revisions 1-5

1.7 Intended Audience and Layout for the NCRN Data Management Plan 1-6

CHAPTER 2: DATA MANAGEMENT ROLES AND RESPONSIBILITIES

2. DATA MANAGEMENT ROLES AND RESPONSIBILITIES 2-1

2.1 Roles and Responsibilities 2-2

2.2 Project Stewardship – Shared Responsibilities 2-2

2.2.1 Project Leader 2-2
2.2.2 Network Data Manager 2-4

2.3 Data Management Coordination 2-5

CHAPTER 3: DATA MANAGEMENT RESOURCES – INFRASTRUCTURE AND SYSTEMS ARCHITECTURE

3. DATA MANAGEMENT RESOURCES – INFRASTRUCTURE AND SYSTEMS ARCHITECTURE 2-1

3.1 Network Infrastructure 3-1

3.1.1 Maintaining Digital Files 3-2

3.2 National Information Management Systems 3-3

3.3 Network Systems Architecture 3-6

3.3.1 Project Database Standards 3-7
3.3.2 Project Tracking Application 3-7

CHAPTER 4: DATA MANAGEMENT PROCESS AND WORK FLOW

4. DATA MANAGEMENT PROCESS AND WORK FLOW 4-1

4.1 Project Work Flow 4-1

4.2 Data Life Cycle 4-4
4.3 Data Distribution and Integration 4-6

CHAPTER 5: DATA ACQUISITION AND PROCESSING

5. DATA ACQUISITION AND PROCESSING 5-1
5.1 Data Acquisition 5-1
5.1.1 Data Discovery / Data Mining 5-1
5.1.2 Programmatic Data Sources 5-2
5.1.3 Non-Programmatic Data Sources 5-4
5.2 Data Processing 5-5
5.2.1 General Data Processing Requirements 5-6

CHAPTER 6: QUALITY ASSURANCE / QUALITY CONTROL

6. QUALITY ASSURANCE / QUALITY CONTROL 6-1
6.1 Quality Assurance and Quality Control 6-1
6.1.1 National Park Service Mandate for Quality 6-1
6.2 Goals and Objectives 6-2
6.3 QA and QC Mechanisms 6-2
6.4 Data Collection 6-3
6.4.1 Methods for Reducing Data Collection Errors 6-5
6.5 Data Entry 6-6
6.5.1 Methods for Reducing Data Entry Errors 6-6
6.6 Verification and Validation Procedures 6-7
6.6.1 Methods for Data Verification 6-8
6.6.2 Methods for Data Validation 6-8
6.7 Data Quality Review and Communication 6-9

CHAPTER 7: DATA DOCUMENTATION

7. DATA DOCUMENTATION 7-1
7.1 Purpose of Metadata 7-1
7.2 Metadata Mandate 7-1
7.3 NPS Integrated Metadata System Plan and Tools 7-2
7.4 Metadata Process 7-3
7.4.1 Data Documentation 7-3

CHAPTER 8: REPORTING AND ANALYSIS

8. REPORTING AND ANALYSIS 8-1

8.1 Data Summary and Analysis 8-1

8.2 Reporting Requirements 8-1

CHAPTER 9: DATA DISSEMINATION

9. DATA DISSEMINATION 9-1

9.1 Data Ownership – The National Park Service Policy 9-1

9.1.1 Establishing Data Ownership Guidelines 9-1
9.1.2 Data Classification: protected vs. public 9-2
9.1.3 What to Protect 9-2
9.1.4 Access Restrictions on Sensitive Data 9-4

9.2 Data Distribution 9-5

9.2.1 Data Distribution Mechanisms 9-5
9.2.2 Public Access to Network Inventory and Monitoring Data 9-6

9.3 Data Feedback Mechanisms 9-7

9.3.1 Data Error Feedback Procedures 9-7

CHAPTER 10: DATA MAINTENANCE, STORAGE, AND ARCHIVING

10. DATA MAINTENANCE, STORAGE AND ARCHIVING 10-1

10.1 Digital Data Maintenance 10-1

10.1.1 Short-term Data Sets 10-1
10.1.2 Long-term Monitoring Data Sets 10-2
10.1.3 Quality Control of Converted Data 10-2
10.1.4 Version Control 10-2
10.1.5 Spatial Data 10-2

10.2 Storage and Archiving Procedures – Digital Data 10-3

10.2.1 Backup Procedures for Digital Data 10-3
10.2.2 Testing of Backup Files 10-3
10.2.3 Data and Network Security 10-3

10.3 Storage and Archiving Procedures–Documents and Objects 10-4

10.3.1 Documents 10-4
10.3.2 Specimens 10-5
10.3.3 Photographs 10-5

LITERATURE CITED

LITERATURE CITED R-1

GLOSSARY

GLOSSARY G-1

APPENDICES

Appendix A: Roles and Responsibilities for Personnel Involved in I&M Network Activity A-1
Appendix B: Digital Data Storage Standard Operating Procedure B-1
Appendix C: DRAFT General Guidance Document: Project Deliverables C-1
Appendix D: DRAFT General Guidance Document: Vouchers D-1
Appendix E: Twelve Basic Inventories Conducted by the Inventory and Monitoring Program E-1
Appendix F: Vital Signs Monitored by the National Capital Region Network F-1
Appendix G: DRAFT General Guidance Document: Data Loggers G-1
Appendix H: Non-Programmatic External Data that Support NCRN Vital Signs Monitoring H-1
Appendix I: NPSpecies Data Entry Standard Operating Procedure I-1
Appendix J: GPS Specifications Standard Operating Procedure J-1
Appendix K: Park Status Report Standard Operating Procedure K-1
Appendix L: Print Media Storage Standard Operating Procedure L-1

FIGURES

1-1: Map of the National Capital Region Network Parks 1-2
2-1: Distribution of Project Related Data Management Tasks 2-3
3-1: Model of the National-Level Application Architecture 3-4
4-1: Conceptual Model of Project Work Flow 4-2
4-2: Diagram of the Typical Project Data Life Cycle 4-5
4-3: Integrating Network Water Quality Data with National Databases 4-7
6-1: Data Evaluation Stages and QA/QC Workflow 6-4
6-2: Field Data Collection Options and the Degree to Which they Enforce Quality Control 6-5
7-1: Metadata Creation and Dissemination Tools 7-2
9-1: Storing and Disseminating Project Information 9-5

TABLES

1-1: Categories of Data Products and Project Deliverables 1-3
2-1: Categories of Data Stewardship Involving All NCRN Personnel 2-1

Chapter 1

Introduction

The Inventory and Monitoring (I&M) Program represents a long-term commitment by the National Park Service (NPS) to assess and document the status and trends of park ecological resources. In 1998, the *National Parks Omnibus Management Act* established the framework for the I&M Program to fully integrate natural resource monitoring and other scientific activities into the management processes of the National Park System.

The Act charges the Secretary of the Interior to "continually improve the ability of the National Park Service to provide state-of-the-art management, protection, and interpretation of and research on the resources of the National Park System," and to "… assure the full and proper utilization of the results of scientific studies for park management decisions." Section 5934 of the Act requires the Secretary of the Interior to develop a program of "inventory and monitoring of National Park System resources to establish baseline information and to provide information on the long-term trends in the condition of National Park System resources."

Mission Statement:

The mission of the National Capital Region Network Data Management Program is to support to the NCRN I&M Program by promoting data quality, integrity, security, interpretability, availability and longevity.

To carry out this mission, the NPS initiated a service-wide, natural resource I&M Program encompassing 270 parks that were identified as having significant natural resources. Ecologically similar parks among the 270 were grouped into 32 networks. Each I&M network has been tasked with documenting existing park vertebrates and vascular plants; developing a management-based ecological monitoring program with a written plan and protocols; and developing an Information Management Plan that encompasses all aspects of the network program.

The National Capital Region Network (NCRN) includes 11 parks:

- Antietam National Battlefield Park (ANTI)
- Catoctin Mountain Park (CATO)
- Chesapeake and Ohio National Historical Park (CHOH)
- George Washington Memorial Parkway (GWMP)
- Harpers Ferry National Historical Park (HAFE)
- Manassas National Battlefield (MANA)
- Monocacy National Battlefield (MONO)
- National Capital Parks/East (NACE)
- Rock Creek Park (ROCR)
- Prince William Forest Park (PRWI)
- Wolf Trap National Park for the Performing Arts (WOTR)

The NCRN I&M Program also supports the National Mall and Memorial Parks (NAMA). A map of the NCRN can be found in Figure 1-1.

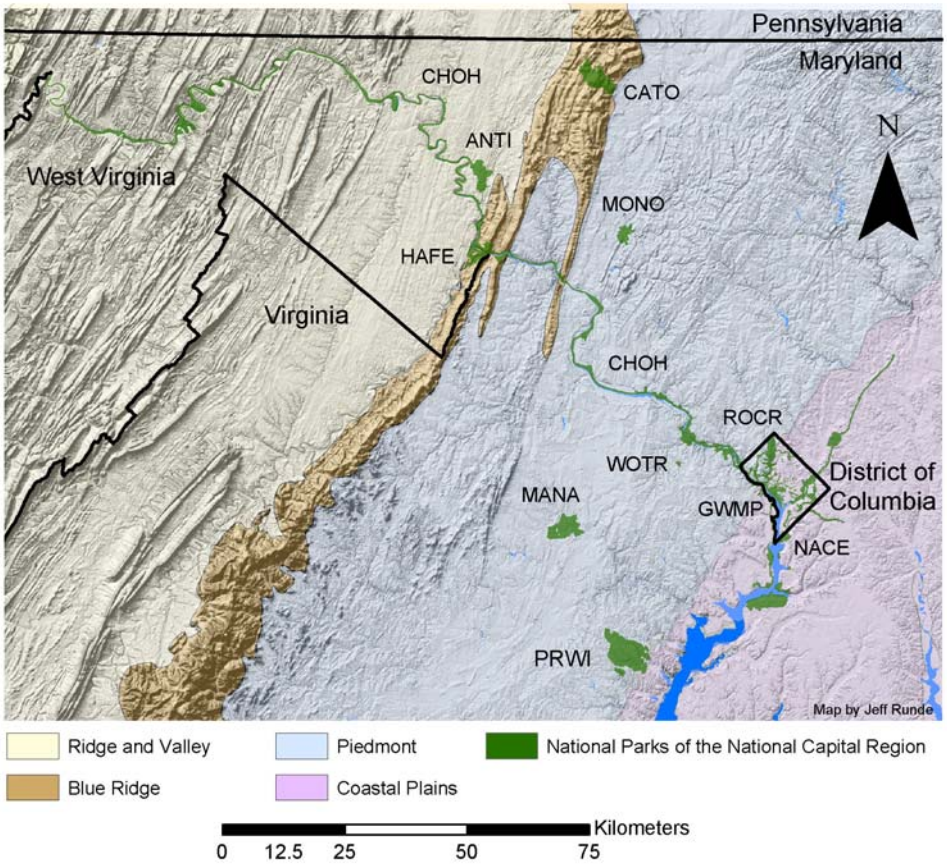


FIGURE 1-1: MAP OF THE NATIONAL CAPITAL REGION NETWORK PARKS

1.1 PURPOSE AND SCOPE OF THE NATIONAL CAPITAL REGION NETWORK DATA MANAGEMENT PLAN

Collecting natural resource data is our first step toward understanding the ecosystems within our national parks. These ecosystems are evolving, and so is our knowledge of them and how they work. We use "raw" data to analyze, synthesize, and model aspects of ecosystems. In turn, we use our results and interpretations to make decisions about the park's vital natural resources. Thus, *data* collected by researchers and maintained by the NCRN will become *information* through analyses, syntheses, and modeling.

One of the primary goals of the NPS is to "improve park management through greater reliance on scientific knowledge." Thus, the NCRN I&M Program will identify, catalog, organize, structure, archive, and provide high-quality natural resource data, which institutions and individuals alike can share. Clearly, researchers, managers, and administrators all rely on the NCRN to secure and maintain the quality of these data for a long time. The procedures and standards outlined in this plan will be used to manage any data the NCRN deems relevant to its mission.

1.1.1 Data and Data Management: An Overview

Any good set of data—whether collected last week or 20 years ago—must tell us enough about itself so that we can reliably preserve and use it. Anyone using these data will need to know as much as possible about how and why these data were collected. Thus, data sets should be accompanied by sufficient documentation about how and why it was collected.

Therefore, NCRN data management standards and guidelines cannot simply attend to the tables, fields, and values that make up a data set. The network must also provide a process for developing, preserving, and integrating the context that makes data interpretable and valuable. Although this means more time documenting, it leads us to clearer preservation and presentation of data.

We sometimes use the term "data" in a broad way that encompasses other products that are generated alongside primary tabular and spatial data. These products fall into five general categories: raw data, derived data, documentation, reports, and administrative records (table 1-1).

To meet I&M Program goals, and to ensure adequate context for primary data products, these categories of projects all require some level of management to ensure their quality and availability. We intend to integrate the manner in which the NCRN creates, manages, and provides the products of our research and analysis. Thus, we will use a more "holistic view" about how natural resource data are generated, processed, finalized, and provided. All phases of data and information processing are integrated, and information about each phase and its processes must be shared through good documentation.

1.2 THE NCRN INVENTORY AND MONITORING PROGRAM'S LONG-TERM DATA MANAGEMENT GOALS

The I&M Program's success at identifying, cataloging, organizing, structuring, archiving, and providing relevant natural resource information will largely determine the program's efficacy and image among critics, peers, and advocates.

The long-term goals of the I&M Program are to:

- establish natural resource inventory and monitoring standards throughout the national park system that transcend traditional program, activity, and funding boundaries
- inventory the natural resources and park ecosystems under NPS stewardship
- monitor park ecosystems to provide reference points for comparisons with other, altered environments
- integrate natural resource inventory and monitoring information into NPSplanning, management, and decision making
- share NPS accomplishments and information with other natural resource organizations and form partnerships for attaining common goals and objectives

TABLE 1-1: CATEGORIES OF DATA PRODUCTS AND PROJECT DELIVERABLES

Category	Examples
Raw data	GPS rover files, raw field forms and notebooks, photographs and sound/video recordings, telemetry or remote-sensed data files, biological voucher specimens.
Compiled/derived data	Relational databases, tabular data files, GIS layers, maps, species checklists.
Documentation	Data collection protocols, data processing/analysis protocols, record of protocol changes, data dictionary, FGDC/NBII metadata, data design documentation, quality assurance report, catalog of specimens/photographs.
Reports	Annual progress report, final report (technical or general audience), periodic trend analysis report, publication.
Administrative records	Contracts and agreements, study plan, research permit/application, other critical administrative correspondence.

To achieve the last two of these goals, we must develop a modern information management infrastructure (e.g., staffing, hardware, software). This infrastructure will include procedures to ensure that relevant natural resource data collected by NPS staff, cooperators, researchers, and others will be entered, quality-checked, analyzed, reported, archived, documented, cataloged, and made available to others for management decision-making, research, and education.

The NPS is a highly decentralized agency with complex data requirements. The primary audience for many of the products from the I&M Program is at the park level–to provide park managers with the information they need to make better-informed decisions and to work more effectively with other agencies and individuals for the benefit of park resources. However, certain data are also needed at the regional or national level for a variety of purposes. The National Park System Advisory Board (2001) has stated that the findings "must be communicated to the public, for it is the broader public that will decide the fate of these resources."

1.3 TYPES OF DATA COVERED BY THE INFORMATION MANAGEMENT PLAN

This plan covers three major categories of data that are coordinated or managed by NCRN. These are:

1. *Data developed or acquired directly by the network as a result of inventory, monitoring, or other projects.* This category includes project-related protocols, protocols, reports, spatial data, and associated materials such as field notes and photographs provided to NCRN by contractors or developed by NCRN staff. Projects can be short-term (one to two years duration) or long-term (ongoing monitoring). Also included in this category are collaborative efforts between the NCRN and other NPS or non-NPS entities.

2. *Data developed by other NPS entities*. The NCRN utilizes numerous datasets developed by parks, regional programs (not limited to the National Capital Region) or data produced at the national level.

3. *Data acquired and/or maintained by non-NPS sources*. Datasets developed by other government agencies and/or non-government organizations are a great asset to the NCRN. Included in this category are datasets that support certain NCRN vital sign programs.

These above categories can contain one or more of the following data formats:

- hard-copy documents (e.g., reports, field notes, survey forms, maps, references, administrative documents)
- objects (e.g., specimens, samples, photographs, slides)
- electronic files (e.g., Word files, email, websites, digital images)
- electronic tabular data (e.g., databases, spreadsheets, tables, delimited files)
- spatial data (e.g., shapefiles, coverages, remote-sensing data)

Each of these data formats has specific requirements for ongoing management and maintenance, which we address in this plan.

1.4 INFORMATION MANAGEMENT GOALS AND OBJECTIVES

As the basic and most important products of scientific research, data and information represent a valuable, and often, irreplaceable resource. Because field experiments and associated data collection are often time and budget consuming,

management of data and information products plays an important role in any scientific program. For long-term ecological monitoring programs, such as the NCRN I&M Program, retention and documentation of high quality data are the foundation upon which the success of the overall I&M Program rests.

The NCRN approach to data management is to develop a Data Management Plan that is "user friendly" to a varied audience from park natural resource managers to data managers. This means providing guidance on data management practices at a number of different levels. To facilitate this guidance, we developed the NCRN plan along three basic principles:

1. keep the plan simple, flexible, and evolving
2. make it useful to all: park GIS and data management staff, regional technical staff, resource management staff, and cooperating scientists
3. include the data users in the decision making process whenever possible

1.5 SCOPE OF THE DATA MANAGEMENT PLAN

The NCRN *Data Management Plan* outlines how we intend to implement and maintain a system that will serve the data and information management needs of our I&M Program. This plan reflects our commitment to the establishment, maintenance, description, accessibility, and long-term availability of high-quality data and information. This document is intended to provide an overarching data management strategy that establishes guidance and specific policy. It is beyond the scope of this plan to detail every data management procedure. Much more detailed data management guidance is provided through guidance documents and standard operating procedures (SOPs), many of which can be found in the appendices of this plan, as well as specific monitoring protocols.

The NCRN Data Management Plan is intended to provide an overarching data management strategy.

The NCRN Data Management Plan describes how the network will:

- support I&M Program objectives
- define data management roles and responsibilities
- acquire and process data
- assure data quality
- document, analyze, summarize and disseminate data and information
- maintain nationally developed data management systems
- maintain, store and archive data

1.6 REVISIONS

The NCRN data management plan covers I&M Program needs based on the most current information systems technology relevant through 2005. As technology, program mandates, and network needs change and evolve over time, data management guidance and strategy must evolve as well. Much of the more dynamic information been placed in specific guidance documents

and SOPs while the more this plan contains information that is more static. Even so, the network data management plan will inevitably require revisions and updates. At the very least the plan will be reviewed every five years in conjunction with the network monitoring plan review.

1.7 INTENDED AUDIENCE AND LAYOUT FOR THE NCRN DATA MANAGEMENT PLAN

This plan provides the long-term documentation of the data management process employed by the NCRN I&M Program. Current and future network staff represents the primary audience. Various chapters and SOPs will apply to any individuals with specific tasks that fall under I&M projects. These individuals may include principal investigators and their staff, resource managers and other park staff, or regional GIS staff.

Credits

Developed by:	Chapter was adapted from materials developed by Doug Wilder (CAKN)
Other contributions included:	Data Management Plan Working Group; Marcus Koenen (NCRN)

Chapter 2

Data Management Roles and Responsibilities

Data management is not an end in itself, but instead is a means of maximizing the quality and utility of our natural resource information. It is a complex process characterized as much by attitudes and habits as it is by infrastructure, standards, and procedures. Although primary responsibility resides with the NCRN Data Manager, good data management could not possibly be accomplished by the data manager alone. Data management is truly a collaborative endeavor that involves many people with a broad range of tasks and responsibilities. As such, a valid data management system must be developed and continually modified to meet the needs of everyone with a role in coordinating, generating, maintaining, and using natural resource information in its many forms. This is a diverse group made up of superintendents, natural resource managers, park scientists, GIS staff, information technology specialists, project managers and technicians, and interpretive staff. A successful data management program is maintained by reinforcing communication, awareness, and acceptance among everyone with responsibilities related to the origin, quality, disposition, and use of the data. The NPS stresses the importance of proper data management and I&M Program guidance states that at least 30% of program resources should be allocated towards data management, analysis, and reporting.

30% Rule

I&M Program guidance states that at least 30% of program resources should be allocated towards data management, analysis, and reporting.

For the NCRN I&M Program to work effectively, everyone within the network, as well as those associated with the network through contracts and agreements will have stewardship responsibilities in the production, analysis, management, and/or end use of the data. Table 2-1 summarizes the data stewardship roles of various network personnel. Each of these broad categories has principal, or 'must-do', responsibilities as well as many potential ancillary tasks. As coordinator of these tasks, the fundamental role of the NCRN data manager will be to understand and determine program and project requirements, to create and maintain data management infrastructure and standards, and to communicate and work with all responsible individuals. A more comprehensive list of personnel roles and responsibilities can be found in appendix A.

TABLE 2-1: CATEGORIES OF DATA STEWARDSHIP INVOLVING ALL NCRN PERSONNEL

Stewardship Category[a]	Related Activities	Principal jobs or positions
Data Production	Creating data or information from any original or derived source. This includes recording locations, images, measurements, and observations in the field, digitizing source maps, keying in data from a hardcopy source, converting existing data sources, image processing, and preparing and delivering informative products, such as summary tables, maps, charts, and reports.	Project Crew Member Project Crew Leader Data/GIS Specialist or Technician
Data Analysis	Using data to predict, qualify, and quantify ecosystem elements, structure, and function as part of the effort to understand these components, address monitoring objectives, and inform park and ecosystem management.	Quantitative Ecologist Resource Specialist
Data Management	Preparing and executing policies, procedures, and activities that keep data and information resources organized, available, useful, compliant, and safe.	Network Data Manager Project Leader GIS Manager Information Technology Specialist
End Use	Obtaining and applying available information to develop knowledge that contributes to understanding and managing park resources.	Network Coordinator Park managers and superintendents Public and Others

Note:

a. Each position is listed in only one category according to overriding responsibilities. However, most positions contribute in each category.

The remainder of this chapter discusses comprehensive data management roles and responsibilities that generally apply to all network activities. Each vital sign monitoring protocol and inventory study plan contains specific instructions for assignments and tasks that nest within this overall framework. Individuals who carry out monitoring protocols and inventory study plans are responsible for reading and understanding these instructional guidelines.

The fundamental role of the network data manager will be to understand and determine program and project requirements, to create and maintain data management infrastructure and standards, and to communicate and work with all responsible individuals.

2.1 ROLES AND RESPONSIBILITIES

A *role* is a function or position (for example, Network Data Manager).

A *responsibility* is a duty or obligation associated with ones position or *role* (for example perform data quality assurance / quality check [QA/QC]).

The current and expected capacity, diversity, and rate of change in information technology makes managing the large amount of information a greater task than any individual can expect to do alone. Knowledgeable individuals from scientific, administrative, and technological disciplines must work together to ensure that data are collected using appropriate methods, and that the resulting data sets, as well as refined products developed from those datasets, are well managed. It is not uncommon for a typical I&M data set to be handled, viewed, or stewarded by most network personnel. It is therefore vital that all those involved with the NCRN I&M Program be well versed in the network's data management standards and practices.

2.2 PROJECT STEWARDSHIP – SHARED RESPONSIBILITIES

Successful data stewardship requires that people involved in all NCRN activities learn and understand network expectations for continuous data management. Project leaders and the data manager comprise the central data management team for inventory and monitoring projects. Each is responsible for certain aspects of project data and both share responsibility for some overlapping tasks (figure 2-1). Because of the collaborative nature of project data management, good communication among project leaders and data managers is essential to meeting program goals.

2.2.1 Project Leader

Since the data management aspects of every inventory or monitoring project normally require the expertise and involvement of several people over a period of months or years, a project leader will usually be designated and charged with keeping track of project objectives, budget, requirements, and progress. This role is usually assigned to a network ecologist or someone with expertise or experience in the particular subject of the project. The project leader is responsible for data quality during all phases of the project, including collecting, entering, handling, reviewing, summarizing, and reporting data. Developing project documentation and metadata are crucial elements of the project leader's role, and the project leader will work closely with the data manager to ensure that good data stewardship practices are followed.

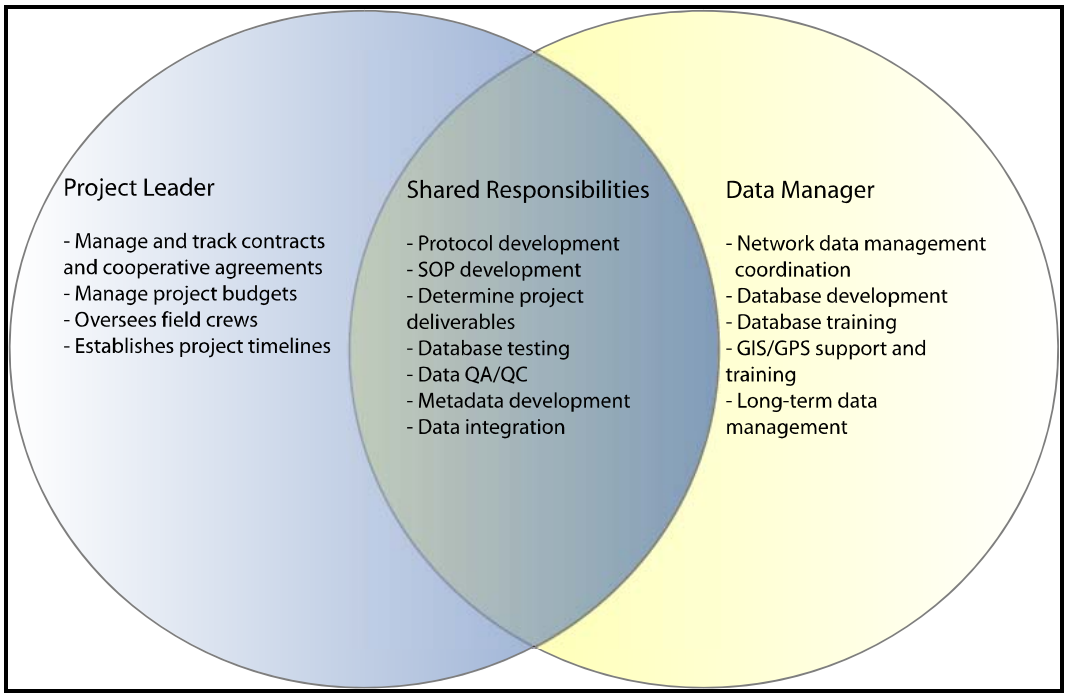


FIGURE 2-1: DISTRIBUTION OF PROJECT RELATED DATA MANAGEMENT TASKS

The project leader's data management responsibilities include:

- Developing project metadata.
- Documenting and implementing standard operating procedures for data collection and data handling.
- Implementing and applying quality control quality assurance measures. This includes ensuring that all field personnel are properly collecting, entering, verifying and validating the data.
- Scheduling regular project milestones such as data collection periods, data processing target dates and reporting schedules.
- Preparing and disseminating regular summary reports and trend analyses.
- Ensuring that all those involved in the project receive the proper field and data management training.

The project leader will also work closely with the data manager to:

- Develop quality assurance and quality control procedures.
- Identify data management training needs.
- Coordinate the development and changes to project databases and field forms.
- Identify sensitive information.
- Define the progression of data from raw, unrefined data to meaningful information.
- Increase the interpretability and accessibility of existing natural resources information.
- Manage the archival process to ensure that project documentation, original field forms, databases, and reports are archived regularly.

2.2.2 Network Data Manager

The network data manager is responsible for ensuring the compatibility of project data with program standards, for designing the infrastructure for the project data, and for ensuring long-term data integrity, security, and availability.

The data manager's responsibilities include:

- Developing overarching data management guidelines that apply to all network projects.
- Developing and maintaining the proper data management infrastructure.
- Developing metadata standards and tools to ensure that all project data is accompanied by proper documentation.
- Creating and maintaining project databases in accordance with program standards.
- Providing data management training to all Network personnel.
- Establishing and implementing methods to protect sensitive information.
- Establishing procedures for data dissemination.

The data manager will work closely with the project leader to:

- Define the scope of the project.
- Become familiar with how project data are collected, handled and used.
- Review QA/QC aspects of project protocols.
- Identify elements that can be incorporated into project databases to enhance quality control.
- Create a streamlined data entry interface.
- Ensure that project metadata is complete.

The network data manager is also responsible for maintaining all of the network's spatial data and the associated metadata. A primary function of this role is to distribute and disseminate spatial data. This involves providing spatial data to field crews and cooperators, and posting data to local and national repositories. Much of the base cartographic data that is a standard requirement for most projects (e.g., park boundary files, hydrology, roads/trails, etc.) are maintained by either regional staff such as regional GIS specialists, or in the case of park boundaries the regional Lands Office, or maintained by the parks. It is imperative that the data manager coordinate with the various offices to ensure that the base cartographic data is current and accurate. Some additional functions of this role are listed below.

GIS DATA MANAGEMENT RESPONSIBILITIES INCLUDE:

- Maintain and disseminate accurate GIS data sets.
- Prepare and maintain FGDC compliant metadata for all GIS data sets.
- Develop procedures for field collection of spatial data including the use of GPS and other spatial data collection techniques.
- Display, analyze, and create maps from spatial data to meet project objectives.
- Collaborate with regional GIS specialists to ensure that the Network maintains the most current park/regional GIS data. Determine the GIS data and analysis needs for the project.
- Create relationships between GIS and non-spatial data.
- Create database and GIS applications to facilitate the integration and analysis of both spatial and non-spatial data.

2.3 DATA MANAGEMENT COORDINATION

The Natural Resource Challenge states that collaboration among the NPS, other public agencies, universities, and non-governmental organizations is necessary to effectively acquire, apply, and promulgate the scientific knowledge gained in national parks. The I&M Program encourages coordination among participants at all levels to help ensure that data collected by NPS staff, cooperators, researchers, and others are entered, quality-checked, documented, analyzed, reported, archived, cataloged, and made available for management decision-making, research, and education.

Key contacts for the network data manager (some of which were already mentioned above) include park, network, and regional GIS and data managers and the project leaders for each monitoring or inventory project. Consistent and productive communication among these personnel leads to common understanding and better synchronization of network and park data management activities. Park and network staff work together to manage resource information using a variety of methods. These include personal visits, phone calls, email, joint meetings and training sessions, as well as the meetings and work of the Natural Resources Advisory Team (NAT), Science Advisory Committee (SAC), Technical Advisory Committee (TAC), and the Board of Directors (BOD).

Data Managers can also facilitate communication by participating in regional workgroups such as the NCR GIS Workgroup. The GIS Workgroup meets quarterly and is chaired by one of the regional GIS specialists and provides an important opportunity to relate information on computer/software updates, initiate collaborative efforts, and discuss regional standards. Standard agenda items for these meetings are documented in meeting minutes coordinated by the chair of each group.

It is important for the data manager to contribute and interact on a national level. Data managers throughout the I&M Program regularly coordinate with each other and national program staff via annual meetings, conference calls, workgroups, a listserv, web sites, and informal communication. This model of cooperation and communication is effective, and it can be applied to resource information management issues and used by administrators.

In addition to supporting network data management needs, the data manager will work closely with park and regional staff to ensure that information is maintained in a manner that promotes data sharing. Accordingly, the NCRN data manager will provide park managers with data management support by assisting with database development (based on the Natural Resource Database Template), coordinating the entry of park data into national databases (for example, NPSpecies and NatureBib), and providing training for best data management practices.

Credits

Developed by:	This chapter was adapted from materials developed by Rob Daley (GRYN), Teresa Leibfreid (CUPN), Doug Wilder (CAKN) and Patrick Flaherty (APHN)
Other contributions included:	Data Management Plan Working Group

Chapter 3

Data Management Resources – Infrastructure and Systems Architecture

Sound data management requires a solid technological framework (both infrastructure and architecture) to ensure that data are secure and easily obtainable. The computer resource *infrastructure* is composed of computers and servers that are functionally or directly linked through computer networking services. This infrastructure represents the foundation upon which our network information system is built and includes not only local hardware systems but also systems located in regional and national offices. Systems *architecture* signifies the applications, database systems, repositories, and software tools that make up the framework of our data management enterprise.

The NCRN I&M Program relies heavily on park, regional, and national information technology (IT) personnel and resources to maintain the computer resource infrastructure. This includes, but is not limited to, hardware replacement, software installation and support, security updates, virus-protection, telecommunications networking, and backups of servers. Therefore communication with park and regional IT specialists is essential to ensure adequate resources and service continuity for our system's architecture. This chapter describes the infrastructure in general terms and focuses specifically on the systems architecture that is central to data management. It is important to keep in mind that technology relating to computer hardware resources is constantly advancing and as such the network's infrastructure is constantly evolving.

3.1 NETWORK INFRASTRUCTURE

An important element of a data management program is a reliable, secure network of computers and servers. The NCRN digital infrastructure has three main components: park and network-based local area networks (LAN), regional wide area network (WAN) and servers maintained at the national level. Park, regional, and national IT specialists maintain this infrastructure and administer all aspects of system security and backups.

The NCRN I&M Program, housed at the Center for Urban Ecology (CUE), is supported by regional IT staff and uses a 100 Base-T LAN to connect to the NCR WAN and domain. The LAN uses standard TCP/IP protocol transmission and routable IP for all applications.

The LAN follows standard NPS security procedures, all of which are overseen by the NCR IT Security Manager. Specifically, all users employ Windows XP Professional network security features, including user-specific, password-protected logins. In addition, the file server administration is restricted only to NCR domain administrators and the server itself is located in a locked, climate-controlled room to which only the server administrator, data manager, and I&M regional coordinator have access.

The NCRN uses a Dell 2600 file server, using a RAID 5 (redundant array of independent disks) hot-swappable drive configuration to minimize data loss in the event of drive failure, providing more than 600 GB of data storage. The server is shared with the regional Natural Resources and Science staff also housed at CUE. The server also houses a Quantum SuperDLT tape drive and uses Veritas Backup Exec 9.1 to backup all computer drives. For added security, all backup tapes not currently in use are stored off site by the NCRN I&M data manager. A more detailed description of the NCRN server backup procedures can be found in chapter 10 as well as in appendix B.

Additional data storage will be provided via external hard drive arrays, which will primarily be used to house the digital imagery obtained in conjunction with the Network's Land Use Change monitoring protocols. This information will not be included in the daily backups but will be backed up onto DVDs.

In addition to information maintained on the NCRN data server, the network has access to information and applications housed on regional and national servers. The following is list of the network's information system resources and the applications and data hosted by the various components.

National Servers

- Master applications – integrated client-server versions of NatureBib, NPSpecies, NR-GIS Metadata Database
- Centralized repositories – NR-GIS Metadata and Data Store, Protocol Clearinghouse
- Public access sites – portals to NatureBib, NPSpecies, NPSFocus, and websites for monitoring networks

Regional Services

- Regional GIS data – base cartographic and imagery datasets such as park boundary files, hydrology, etc.
- GIS and image processing applications – access to certain specialized applications such as ERDAS Imagine Mr. SID compression software

Network Data Server

- Master project databases – compiled data sets for monitoring projects and other multi-year efforts that have been certified for data quality and archived to ensure data security
- Common lookup tables – e.g., parks, projects, personnel, species
- Project tracking application – used to track project status, contact information, product due dates
- Network digital archive – network repository for finished versions of project deliverables for network projects (e.g., reports, methods documentation, data files, metadata, etc.)
- Local applications – desktop versions of national applications such as NPSpecies and Dataset Catalog
- Working files – working databases, draft geospatial themes, drafts of reports, administrative records
- GIS files – base spatial data, imagery, and project-specific themes

3.1.1 Maintaining Digital Files

All employees housed at CUE are provided with a personal "local" directory on the CUE file server where access is limited to that individual. A shared directory contains the NCRN program directory as well as directories for each regional program housed at CUE. The NCRN I&M Program has developed a hierarchical directory structure for storing and organizing digital files. Only those directly involved with the NCRN I&M Program are provided with "write" access to this directory. Additional restrictions are applied

to subdirectories to ensure data security. For example, only the NCRN data manager and regional IT specialist have 'write' privileges to the *Archive* directory. The following is an abbreviated description of the file structure.

I&M Directory Structure

- Admin – Documents related to program administration such as program reports.
- Photos – Program digital photograph library.
- Projects – Contains documents relating to both short-term and long-term projects.
  - Active directories – Repository for all project files that are active or subject to change on a regular basis. This mainly relates to project databases.
  - Archived directories – Repository for project materials that generally remain unchanged.
  - INBOX – A directory where network staff can place materials that should be stored in the I&M Archives directory. The data manager is responsible for transferring these materials into the proper archive folders since other staff do not have 'write' privileges to the archive directory.

GIS data is shared by all regional employees located at CUE, so this directory is not nested under the I&M directory structure. Data located within this directory are arranged in a park-specific fashion. This directory is designed so that only the regional GIS staff and the NCRN data manager have write privileges. All of the base cartographic and imagery data are stored here. Additional data can be added as long as they are accompanied by FGDC compliant metadata. Any GIS file undergoing edits is stored elsewhere until it is considered complete. Refer to appendix B for more information on the NCRN digital data storage procedures and directory structures.

3.2 NATIONAL INFORMATION MANAGEMENT SYSTEMS

The need for effective natural resource information management extends across NPS divisional boundaries and management strategies must be defined at the highest level possible. In this context, integrated inventory and monitoring of natural resources is multidisciplinary and requires national-level, programmatic data and information management strategies for success.

The basic strategy of natural resource and therefore inventory and monitoring information management is to provide integrated natural resource databases and information systems that enhance NPS manager and staff access and use of timely and valid data and information for management decisions, resource protection, and interpretation. Inventory and monitoring information needs are broadly separated into two categories:

- *Detailed data and information needed for onsite resource management and protection*. The information used to guide natural resource management decisions must be specific to inform and be useful to management staff at parks and central offices.
- *Summary information needed to describe the resources and their condition*. This kind of information usually needs to be aggregated across the NPS for use by NPS and Department of the Interior managers and central office personnel to answer requests from Congress and for budget, program, and project planning.

The NPS Natural Resource Program Center (NRPC) and the I&M Program actively develop and implement a national-level, program-wide information management framework. NRPC and I&M staff integrate desktop database applications with internet-

based databases to serve both local and national-level data and information requirements. NRPC staff members work with regional and support office staff to develop extensible desktop GIS systems that integrate closely with the database systems. Centralized data archiving and distribution capabilities at the NRPC provide for long-term data security and storage. NRPC sponsors training courses on data management, I&M techniques, and remote sensing to assist I&M data managers with developing and effectively utilizing natural resource information.

National-level Application Architecture

To achieve an integrated information management system, three of the national-level data management applications (NatureBib, NPSpecies, and NR-GIS Metadata Database) utilize a distributed application architecture with both desktop and internet-accessible (master) components (figure 3-1).

NatureBib—NatureBib is the master database for bibliographic references that merges a number of previously separate databases such as Whitetail Deer Management Bibliography (DeerBib), Geologic Resource Bibliography (GRBib), and others. It also contains citation data from independent databases like NPSpecies and the Dataset Catalog and NR-GIS Metadata Database. It currently focuses on natural resource references, but may eventually be linked to references on cultural resources and other park operations. As with NPSpecies and NR-GIS Metadata Database, it is possible to download data from the master web version into the MS Access desktop version that can be used locally on computers with limited internet connectivity (http://www.nature.nps.gov/nrbib).

NPSpecies—NPSpecies is the master species database for the NPS. The database lists the species that occur in or near each park, and the physical or written evidence for the occurrence of the species (e.g., references, vouchers, and observations). Taxonomy and nomenclature are based the interagency Integrated Taxonomic Information System (ITIS). The master version of NPSpecies for each park or network can be downloaded from the master website into an MS Access version of NPSpecies. The internet-based version is the master database, which can be accessed via password-protected logins administered by park, network, and regional data stewards assigned for each park and network. The master database requires that species lists are certified by networks before any data will be made available to the public. NPSpecies is linked to NatureBib for bibliographic references that provide written evidence of a species' occurrence in a park and will be linked to NR-GIS Metadata Database to document biological inventory products. The MS Access application and additional details can be found at the NPSpecies website (http://science.nature.nps.gov/im/apps/npspp/index.htm).

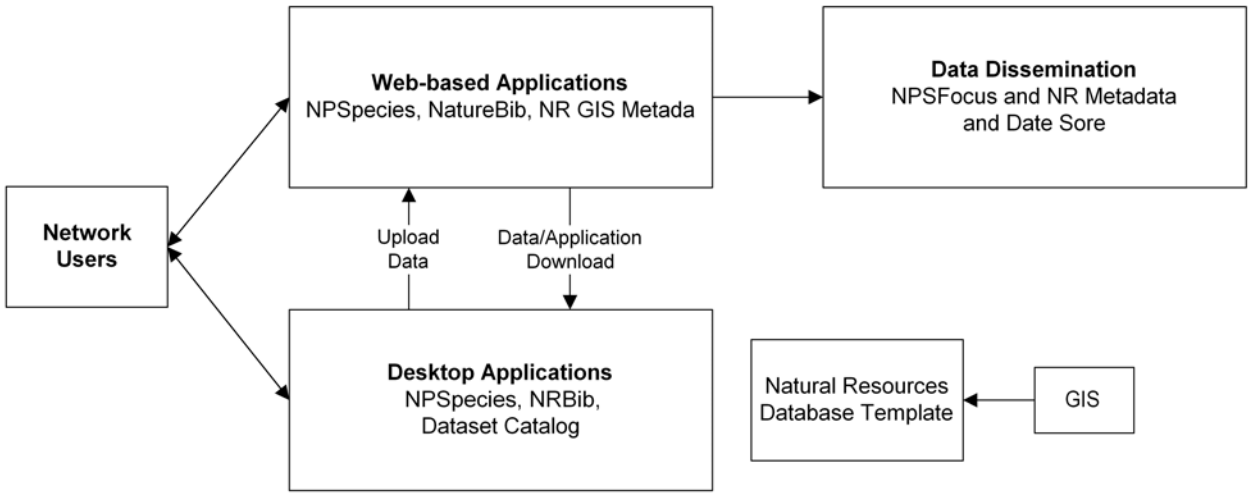


FIGURE 3-1: MODEL OF THE NATIONAL-LEVEL APPLICATION ARCHITECTURE

Dataset Catalog and NR-GIS Metadata Database—Dataset Catalog is a desktop metadata database application developed by the I&M Program to provide a tool that parks, networks, and cooperators can use to inventory and manage data set holdings. Although not designed as a comprehensive metadata tool, the Dataset Catalog is used for cataloging abbreviated metadata about a variety of digital and non-digital natural resource data sets. The Dataset Catalog helps parks and networks begin to meet Executive Order 12906 mandating federal agencies to document all data collected after January 1995. It provides brief metadata and a comprehensive list about all resource data sets for use in data management, project planning, and more stringent metadata activities. As with other service-wide applications, the master metadata database (NR-GIS Metadata Database) is available through a website and will be linked to NPSpecies (the NPS species database) and NatureBib (the bibliographic database). It will be possible to download a version in MS Access format from the master website (Dataset Catalog: http://science.nature.nps.gov/im/apps/datacat/index.htm
and NR-GIS Metadata Database: http://science.nature.nps.gov/nrdata).

NPS Metadata Tools and Editor—The NPS Metadata Tools and Editor is a metadata management and editing application that implements two separate interfaces: either (1) as an extension within ArcCatalog (v8.3 or v9.0); or (2) as a stand-alone desktop interface. The application incorporates the functions previously available in the NPS Metadata Tools Extension for ArcGIS developed by the NPS Midwest Region GIS Support Office. The application is intended to be the primary editor for metadata that will be uploaded to the NR-GIS Data Store. The NPS Metadata Tools and Editor application integrates with the NR-GIS Data Store information system by producing XML metadata based on the NPS Metadata Profile. This XML metadata file can then be uploaded to the NR-GIS Metadata and Data Store application. Metadata editing is done with editing style sheets that transform a standard XML metadata file into an editable record. The initial editing functions include those elements required for upload to the NR-GIS Data Store (the NPS_Basic_Edit stylesheet) and Biological Profile metadata elements (the NPS_BioProfile_Edit stylesheet).

Other National-Level Inventory and Monitoring Information Management and GIS Applications

NPSTORET—STORET is an interagency water quality database developed and supported by the Environmental Protection Agency's (EPA) to house local, state, and federal water quality data collected in support of managing the nation's water resources under the Clean Water Act. STORET is used by NPS as a repository of physical, chemical, biological, and other monitoring data collected in and around national park units by park staff, contractors, and cooperators. The NPS operates its own service-wide copy of STORET and makes periodic uploads to the EPA STORET National Data Warehouse so that data collected by and for parks will be accessible to the public. NPS Director's Order 77 indicates that the NPS should archive water quality data in STORET, and the NPS Water Resources Division (WRD) requires that any data collected as part of a funded WRD project get archived in STORET. NPSTORET (also known as Water Quality Database Templates) the NPS master database designed to facilitate park-level standardized reporting for STORET. The database is still in development, but metadata, protocols, data dictionaries, and reporting capabilities are available through a front-end form. Upon implementation, network staff and cooperators will be able to use the MS Access version of NPSTORET either as a direct database for data entry and management, or as a means of submitting data for upload to STORET by WRD staff. The MS Access application and additional details can be found at: http://www.nature.nps.gov/water/infodata.htm. Additional information on STORET can be found at: http://www.epa.gov/storet.

Natural Resource Database Template—The Natural Resource Database Template (NRDT) is a flexible, relational database in MS Access for storing inventory and monitoring data (including raw data collected during field studies). This relational database can be used as a standalone database or in conjunction with the GIS software (e.g., ArcView or ArcGIS) to enter, store, retrieve, and otherwise manage natural resource information. The template has a core database structure that can be modified and extended by different parks and networks depending on the components of their inventory and monitoring

program and the specific sampling protocols they use. A description of the Database Template application, a data dictionary, and example implementations are located on the NR Database Template website (http://science.nature.nps.gov/im/apps/template/index.htm).

Natural Resource Monitoring Protocols Clearinghouse—The Natural Resource Monitoring Protocol Clearinghouse (i.e., Protocol Database) is a web-based clearinghouse of sampling protocols used in national parks to monitor the condition of selected natural resources. The database provides a summary of, and in many cases allows the user to download a digital copy of, sampling protocols that have been developed by the prototype monitoring parks or other well-established protocols used in national parks. The Protocol Database also makes it possible to download database components (e.g., tables, queries, data entry forms) in MS Access that are consistent with the Natural Resource Database Template that have been developed for a particular protocol. See the Protocol Database website for available protocols (http://science.nature.nps.gov/im/monitor/protocoldb.cfm).

NR-GIS Data Store—The NR-GIS Data Store is a key component of the data dissemination strategy employed by the I&M Program. The NR-GIS Data Store is a graphical search interface that links dataset metadata to a searchable data server on which datasets are organized by NPS units, offices and programs. The interface allows customized public or protected searches of natural resource datasets, inventory products and GIS data produced by the I&M and Natural Resource GIS Programs. Each park or network is able to post and curate its data on the server. The NR-GIS Data Store will be integrated with the master NR-GIS Metadata Database application to streamline programmatic data documentation and dissemination processes. The simple browse function of this server can be accessed at: http://nrdata.nps.gov/. See the NR-GIS Data Store website for further information (http://science.nature.nps.gov/nrdata).

NPS Technology Information Center—The Technology Information Center (TIC) provides a digital repository for all types of documents, data, GIS maps, and architectural blueprints for NPS. TIC also provides a tracking system and can generate copies of requested documents upon request. (https://amoebawww.den.nps.gov/amoeba/TIC/TIC.NSF)

3.3 NETWORK SYSTEMS ARCHITECTURE

Rather than developing a single, integrated database system, the NCRN data design relies upon modular, standalone project databases that share design standards and links to centralized data tables. Individual project databases are developed, maintained, and archived separately. There are numerous advantages to this strategy:

- Data sets are modular, allowing greater flexibility in accommodating the needs of each project area. Individual project databases and protocols can be developed at different rates without a significant cost to data integration. In addition, one project database can be modified without affecting the functionality of other project databases.

- By working up from modular data sets, we avoid a large initial investment in a centralized database and the concomitant difficulties of integrating among project areas with very different—and often unforeseen—structural requirements. Furthermore, the payoff for this initial investment may not be realized down the road by greater efficiency for interdisciplinary use.

This system of stand-alone project databases is the most appropriate approach given the current network infrastructure and data management priorities. The network recognizes the benefit of moving away from standalone applications to a more comprehensive client server environment and when resources and infrastructure permit, is likely something the NCRN I&M Program will move towards.

3.3.1 Project Database Standards

Project database standards are necessary for ensuring compatibility among data sets, which is vital given the often unpredictable ways in which data sets will be integrated, summarized and analyzed. When well thought out, standards also help to encourage sound database design and facilitate interpretability of data sets. Databases that are developed for park and network projects will all contain the following main components:

- *Common lookup tables* — Links to entire tables that reside in a centralized database, rather than storing redundant information in each database. It is a good strategy to centralize this information so that users have access to the most updated versions in a single, known place. Centralizing also avoids redundancy and versioning issues among multiple copies. These tables typically contain information that is not project-specific (e.g., lists of parks, personnel, and species).

- *Core tables and fields based on network and national templates* — These tables and fields are used to manage the information describing the "who, where, and when" of project data. Core tables are distinguished from common lookup tables in that they reside in each individual project database and are populated locally. These core tables contain critical data fields that are standardized with regard to data types, field names, and domain ranges.

- *Project-specific fields and tables* — The remainder of database objects can be considered project-specific, although there will typically be a large amount of overlap among projects. This is true even among projects that may not seem logically—for example, a temperature field will require similar data types and domain values. As much as is possible, efforts will be made to develop these project-specific objects to be compatible with those maintained by other networks and cooperators managing similar data sets—especially if integration with other data sets is important for meeting project objectives.

3.3.2 Project Tracking Application

To support program coordination and annual reporting, and to improve accountability for the products of our natural resource inventory and monitoring efforts, our network will implement a project tracking system. The NCRN has chosen to adopt a MS Access based project tracking database developed by the Southeast and Southwest Alaskan Networks. Network personnel are currently in the process of modifying and adapting the database for use in the NCRN. The database will be available to all network staff and each of these staff will be able to make certain changes to update information about project status, deliverable details, etc. Project managers will be responsible for updating and maintaining the database for projects they manage. The primary purpose with the database will be to:

Compatibility with national standards

As much as possible, NCRN standards for fields, tables and other database objects will mirror those conveyed through the Natural Resource Database Template. Where there are differences between local and national standards, documentation of the rationale for these differences will be developed. In addition, documentation and database tools (e.g., queries that rename or reformat data) will be developed to ensure that data exports for integration are in a format compatible with current national standards.

- Maintaining the list of projects — By maintaining a single list of natural resource data projects, it is much easier to quickly find project-related information (e.g., status, funding sources and amounts, objectives, contact information) and summarize that information for administrative reports.

- Tracking project budgets — The funding for each project, including the account(s) from which the funding was allotted, will be entered. All project invoices and expenditures will be entered and tracked.

- Tracking deliverables — For each project a comprehensive list is maintained of what deliverables are expected and when. Once they are delivered and posted or archived, this function shifts to being a finding aid for available products. Deliverables are first specified at project initiation and information is updated at various project milestones (e.g., contracting, product delivery, archival).

- Product metadata — it is also the intent of the network to incorporate metadata links to the Dataset Catalog so that network personnel can enter product metadata while updating the project database.

- Reporting — The database is designed to automate reports such as the Annual Administrative Report and Work Plan (AARWP) based on project information for each fiscal year.

Credits

Developed by	Patrick Flaherty (APHN) and John Boetsch (NCCN)
Other contributions included:	Section 3.2 - National Inventory and Monitoring Data Management Framework was adapted from material written by Lisa Nelson (WASO).

Chapter 4

Data Management Process and Work Flow

This chapter considers the general work flow characteristics of projects that produce natural resource data, and then gives an overview of how natural resource data are generated, processed, finalized and made available. Data management activities that relate to the various stages of a project are highlighted. By describing the progressive stages of a project and the life cycle of the resulting data, we can more easily communicate the overall objectives and specific steps of the data management process. In addition, this awareness helps us manage the staffing resources needed to produce, maintain, and deliver quality data and information.

4.1 PROJECT WORK FLOW

From the perspective of managing workflow, there are two main types of projects:

- *Short-term projects*, which may include individual park research projects, inventories, or pilot work done in preparation for long-term monitoring.

- *Long-term projects*, which will mainly be the monitoring projects implemented by the I&M Program, but which may also include multi-year research projects and monitoring performed by other agencies and cooperators. Long-term projects will often require a higher level of documentation, peer review and program support.

Both short-term and long-term projects share many work flow characteristics, and both generate data products that must be managed and made available. From a data management standpoint, a primary difference between short- and long-term projects is an increased need to adhere to standards for long-term projects to ensure internal compatibility over time.

Projects can be divided into five primary stages: *planning and approval*; *design and testing*; *implementation*; *product integration*; *evaluation and closure* (figure 4-1). Each stage is characterized by a set of activities carried out by staff involved in the project. Primary responsibility for these activities rests with different individuals according to the different phases of a project. Additional discussion of the different roles and responsibilities of park and network staff can be found in chapter 2 of this plan. For short-term projects this is a finite process ending in closure usually after one or two years. This is a cyclical process for long-term projects, however, the frequency depends on the individual sampling protocol.

Planning and Approval

At the planning and approval stage, many preliminary decisions are made about project scope and objectives. Funding sources, permits, and compliance are all addressed in this phase. Primary responsibility rests with project leaders and program administrators. Although this phase lacks specific data management activities, it is important that data managers remain informed of projects in this phase. This is especially true as timelines for deliverables are finalized. All contracts, agreements, and permits should include standard language that describes the formats, specifications, and timelines for project deliverables. A copy of the NCRN Guidelines for Project Deliverables, which contains numerous SOPs dealing with specific products including voucher specimens, will be supplied to every principal investigator after a contract or agreement has been finalized. A copy of the guidelines documents that outline requirements for project deliverables and vouchers can be found in appendix C and D.

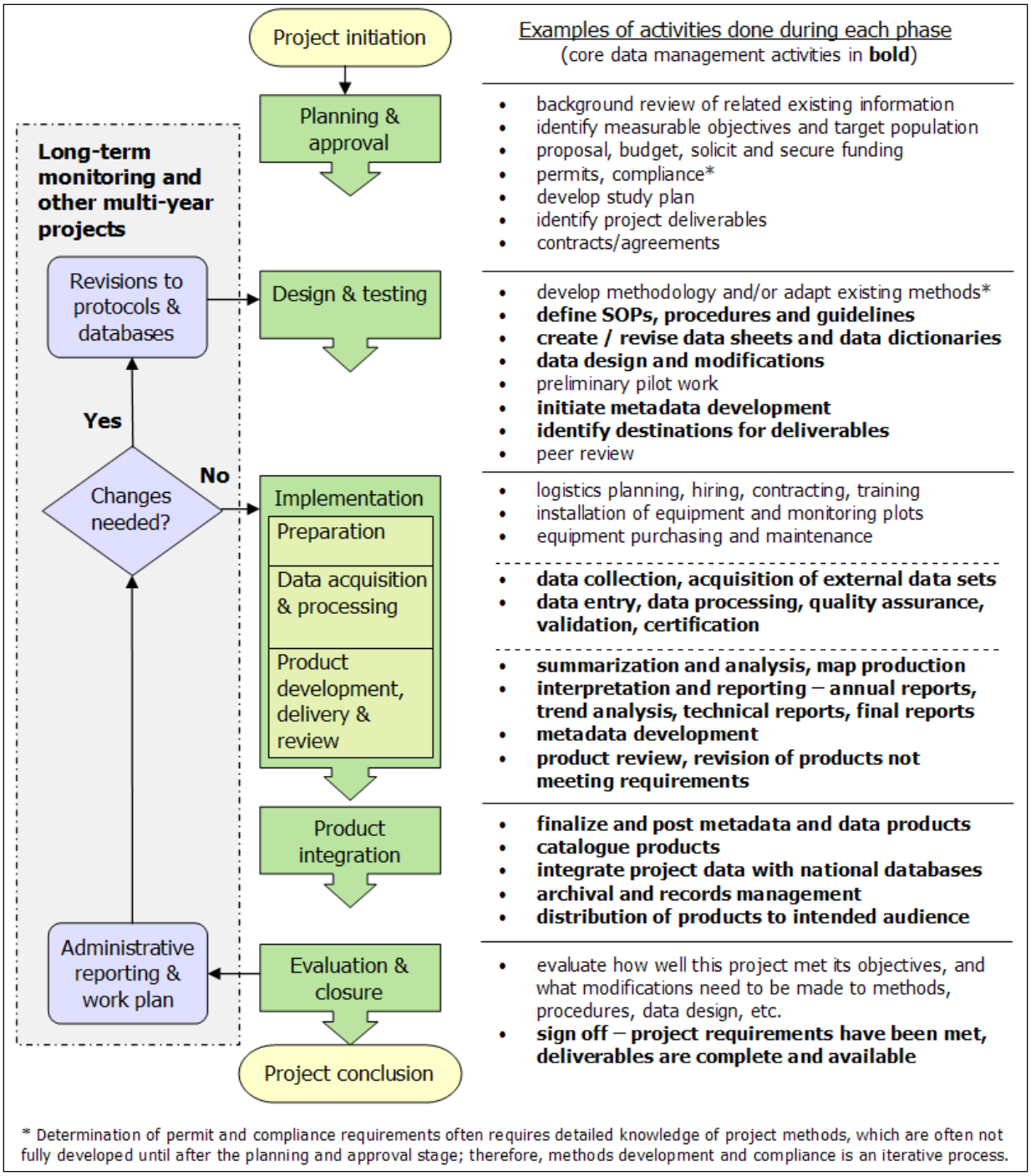


FIGURE 4-1: CONCEPTUAL MODEL OF PROJECT WORK FLOW

During this phase, all of the details are worked out regarding how data will be acquired, processed, analyzed, reported, and made available to others. The project leader is responsible for developing and testing project methodology, or for modifying existing methods to meet project objectives. It is critical that the project leader and the data manager work together throughout this phase. The collaboration and dialog between these two helps to build and reinforce good data management throughout the project, especially during the crucial stages of data acquisition, processing, and retrieval. An important part of this collaboration is developing the data design and data dictionary, where the specifics of database implementation and parameters that will be

collected are defined in detail. Devoting adequate attention to this aspect of project is *possibly the single most important part* of assuring the quality, integrity, and usability of the resulting data. Once the project methods, data design, and data dictionary have been developed and documented, a database can be constructed to meet project requirements.

Implementation

During the implementation phase, data are acquired, processed, error-checked, and documented. The project leader oversees all aspects of implementation – from logistics planning, contracting, training, and equipment procurement to data acquisition, report preparation and final delivery. Throughout this phase, data management personnel function primarily as facilitators, providing training and support for database applications, GIS, GPS and other data processing applications; facilitating data summarization, validation, and analysis; and assisting with the technical aspects of documentation and product development. The specific roles of data management staff during this phase will depend primarily on the technical capabilities of the project staff. As much as is possible, these roles should be worked out in advance of implementation.

The implementation phase is also when reports, maps, GIS themes, and other products are developed and delivered. Toward the end of this phase, project staff members work to develop and finalize the deliverables that were identified in the project planning documents (i.e., protocol, study plan, contract, agreement, or permit). In general, all raw and derived data products, metadata, reports, and other documentation should be delivered to the project manager assigned to the project. Administrative records should be delivered to appropriate park and network staff as specified. All project deliverables should be developed and delivered according to product specifications.

Product Integration

During the product integration phase, data products and other deliverables are integrated into national and network databases, metadata records are finalized and posted in clearinghouses, and products are distributed or otherwise made available to their intended audience. Another aspect of integration is merging data from a working database to a master database maintained on the network server. This occurs only after the annual working data set has been certified for quality by the project leader. Certain projects may also have additional integration needs, such as when working jointly with other agencies for a common database.

Product integration includes creating records for reports and other project documents in NatureBib, posting imaged documents to the appropriate repository, posting metadata records that have been completed and submitted by project leaders, and updating NPSpecies to reflect any new species occurrence information derived from the project. This will allow the information from the project to be searchable and available to others via service-wide search engines.

Evaluation and Closure

Upon project closure, records are updated to reflect the status of the project and its associated deliverables in a network project tracking application. For long-term monitoring and other cyclic projects, this phase occurs at the end of each field season and leads to an annual review of the project. For non-cyclic projects, this phase represents project completion. After products are catalogued and made available, program administrators, project leaders, and data managers should work together to assess how well the project met its objectives, and to determine what might be done to improve various aspects of the methodology, implementation, and formats of the resulting information. For monitoring protocols, careful documentation of all changes is required. Changes to methods, SOPs, and other procedures are maintained in a tracking table associated with each document. Major revisions may require additional peer review.

4.2 DATA LIFE CYCLE

During various phases of a project, data take on different forms and are maintained in different places as they are acquired, processed, documented, and archived. This data life cycle is characterized by a sequence of events that we can model to facilitate communication. These events involve interactions with the following objects:

- *Raw data* — Analog data recorded by hand on hard-copy forms and digital files from handheld computers, GPS receivers, telemetry data loggers, etc.

- *Working database* — A project-specific database for entering and processing data for the current season (or other logical period of time). This might be the only database for short-term projects where there is no need to distinguish working data for the current season from the full set of validated project data.

- *Certified data and metadata* — Completed data and documentation for short-term projects, or one season of completed data for long-term monitoring projects. Certification is a confirmation by the project leader that the data have passed all quality assurance requirements and are complete and ready for distribution. Metadata records include the detailed information about project data needed for its proper use and interpretation (see chapter 7).

- *Master database* — Project-specific database for storing the full project data set, used for viewing, summarizing, and analysis. In the case of on-going monitoring projects, only data that have passed all quality assurance steps are incorporated here.

- *Reports and data products* — Information that is derived from certified project data.

- *Edit log* — A means of tracking changes to certified data (often incorporated into project databases).

- *National databases and repositories* — Applications and repositories maintained at the national level, primarily for the purpose of integration among NPS units and for sharing information with cooperators and the public.

- *Local archives and digital library* — Local storage of copies of data, metadata and other products generated by projects. Archives are for hard-copy items and off-line storage media, whereas the digital library is maintained live on a server.

Although the data life cycle may vary depending on specific project needs and objectives, the typical life cycle for NCRN data proceeds as follows (figure 4-2):

1. *Acquire data* — For data recorded by hand in the field, data forms should be reviewed regularly (at least daily) for completeness and validity in order to capture errors as close to their origin as possible.

2. *Data entry / impor*t — Analog data are entered manually and digital data files are uploaded to the working database.

3. *Archive raw data* — Copies of all raw data files are archived intact. Digital files are copied to the project archive directory. Hard copy forms are either scanned and placed in project archive directory, or copied and placed in the archives.

4. *Verification, processing and validation* — The accuracy of raw data transcription is verified; data is processed to remove missing values and other data flaws; and data is validated using database queries to capture missing data, out-of-range values, and logical errors.

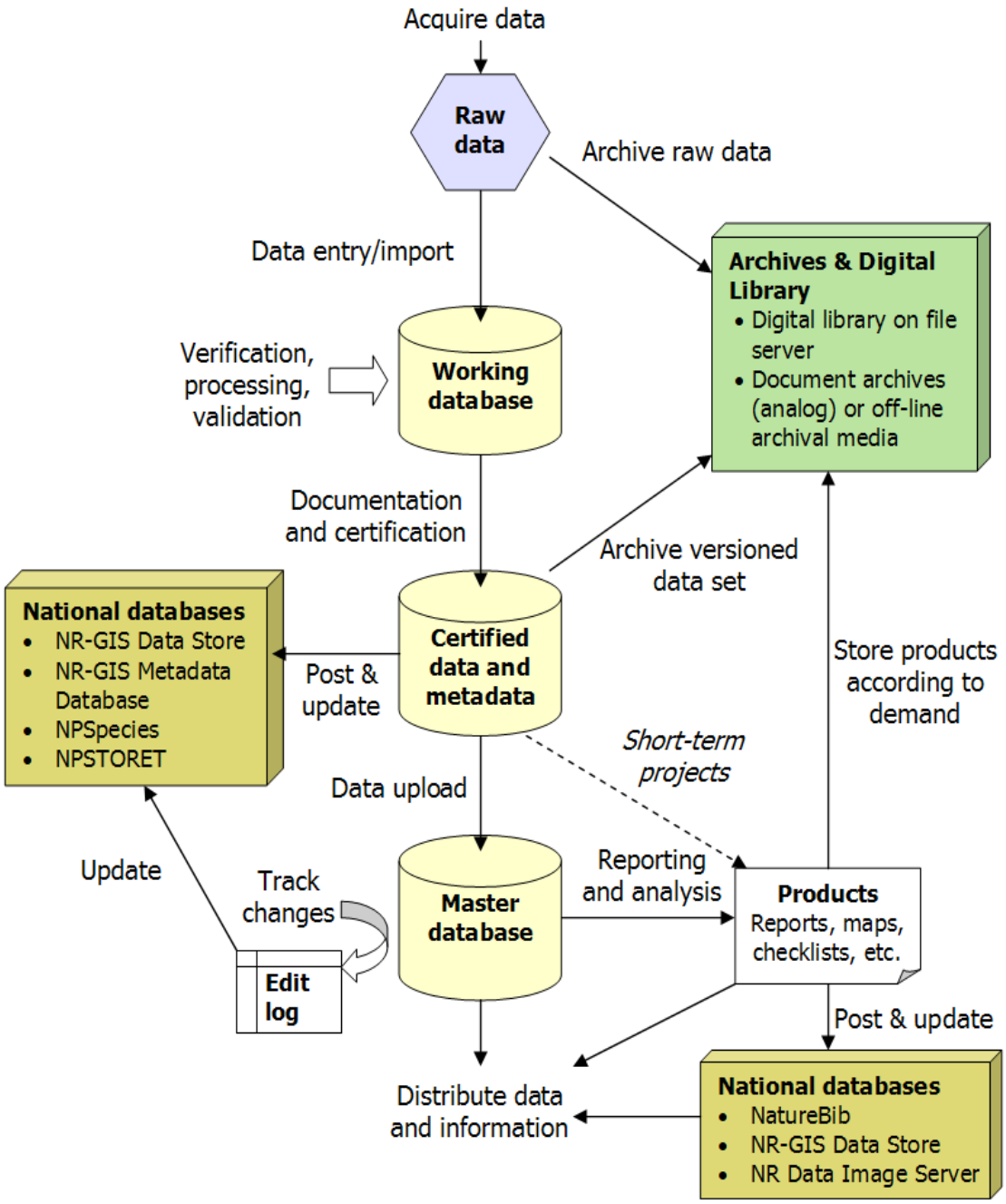


FIGURE 4-2: DIAGRAM OF THE TYPICAL PROJECT DATA LIFE CYCLE

5. *Documentation and certification* — Project metadata is developed or updated, and the data set is certified. Both data and metadata are checked to assure they meet proper standards and are ready to be posted and delivered.

6. *Archive versioned data set* — Copies of the certified data and metadata are placed in the digital library. This can be accomplished by storing a compressed copy of the working database, or by exporting data to a more software-independent format (e.g., ASCII text; see chapter 10).

7. *Upload data* — Certified data are uploaded from the working database to the master project database. This step might be skipped for short-term projects where there is no need to distinguish working data for the current season from the full set of certified project data.

8. *Reporting and analysis* — Certified data are used to generate data products, analysis, and reports, including semi-automated annual summary reports for monitoring projects.

9. *Store products* — Reports and other data products are stored according to format and likely demand – either in the digital library, on off-line media, or in the document archives.

10. *Post products and data to national databases* — To make data available to others, reports and other products are posted to national repositories such as NR-GIS Data Store or the NR Data Image Server and national databases such as NPSpecies, NPSTORET, and NR-GIS Metadata Database. In addition, products are catalogued in NatureBib and posted to the Biodiversity Data Store. Data products may not be posted if they contain protected information about the nature or location of rare, threatened or endangered species, or other natural resources of management concern (see chapter 9).

11. *Track changes* — All subsequent changes to certified data are documented in an edit log, which accompanies project data and metadata upon distribution. Significant edits will trigger reposting of the data and products to national databases and repositories.

This sequence of events occurs in an iterative fashion for long-term monitoring projects, whereas this sequence is followed only once for short-term projects. For projects spanning multiple years, decision points include whether or not a separate working database is desirable, and the extent to which product development and delivery is repeated year after year.

4.3 DATA DISTRIBUTION AND INTEGRATION

As projects are concluded or, in the case of long-term projects, reach reporting milestones; products are finalized by the project lead; and products are sent to the appropriate person for integration, posting, and distribution. In most cases the data manager will review the product for conformance with format standards then store the product in the appropriate repository. After storing the products, they are indexed via the metadata records that are part of the NR-GIS Metadata Database. The metadata records provide pointers to data and data products. Distribution then follows as data discovery allows potential users to find and either request or download the data sets from their repositories.

In addition to storing and distributing data products, product integration also involves updates to national databases such as NPSpecies and NPSTORET. Both of these databases have local desktop databases that can be updated with data collected during the course of a project. Desktop databases are then uploaded and synchronized with the national databases on a regular basis.

For NPSTORET, all water quality data collected by the network will be managed according to guidelines established by the NPS Water Resources Division. The NCRN Water Quality Specialist will maintain a desktop copy of NPSTORET where all data collected in connection with NCRN water monitoring protocols will be stored. All data in the local database will be transferred, at least annually, to NPS Water Resource Division for upload to the STORET database (figure 4-3).

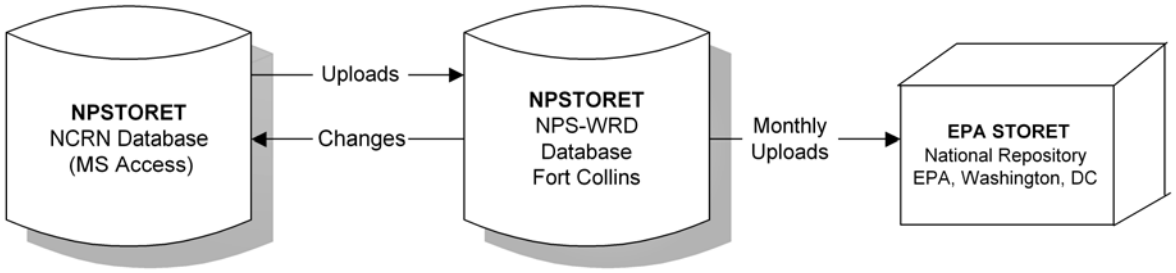


FIGURE 4-3: INTEGRATING NETWORK WATER QUALITY DATA WITH NATIONAL DATABASES

Credits

Developed by	John Boetsch (NCCN)
Other contributions included:	Velma Potash (CACO); Sara Stevens (NCBN); Data Management Plan Working Group Special thanks to Dorothy Mortenson (SWAN) and Doug Wilder (CAKN) for significant contributions and ideas

Chapter 5

Data Acquisition and Processing

The I&M Program, in support of NPS Natural Resources Challenge, is responsible for acquiring information required by park managers to properly managing and maintaining the natural resources of their park. To accomplish this task, the NCRN I&M Program collects information from multiple sources and processes it so that it meets national and network standards. This chapter describes the steps involved in acquiring data as well as the initial stages of data processing. Collecting new data and acquiring existing data from outside sources are essential to strengthening the scientific foundation of the I&M program and providing the park managers with high quality data necessary to properly manage park resources.

There are two general classifications of data handled by the I&M Program:

1. Programmatic Data — any data produced from projects that are initiated (funded) by the I&M Program or projects that in some way involve the I&M Program.

2. Non-programmatic Data — includes data collected by entities other than the NCRN I&M Program. Non-programmatic data is further classified as follows:

    - Non-programmatic NPS Data — any data produced by the NPS that did not involved the inventory and monitoring program.

    - Non-programmatic External Data — any data produced by agencies or institutions other than the National Parks Service.

These definitions do not rank the importance of the different data types to the I&M Program. The definitions provide a way of distinguishing between data gathered from different sources. The importance or value placed on a dataset will be based on the quality or usefulness of data itself and the impact it has on the NCRN I&M Program and parks.

5.1 DATA ACQUISITION

As mentioned above, the NCRN collects data and information from a variety of sources. Some is original data acquired as part of I&M field studies but a great deal of information acquired by the network originates from sources outside the I&M Program. This section will detail various sources of data.

5.1.1 Data Discovery / Data Mining

The NCRN often acquires a great deal of data relating to network projects prior to conducting any field work through the data discovery process. Data discovery or data mining is the process of searching for existing data/information, from both programmatic and non-programmatic sources, that may be useful I&M Program mission. The data mining process usually takes place during the project planning or scoping phase to help gather information that assists in the development of project methodologies and designs. Many of the data sources listed below are accessible via the internet but some require visiting local research or academic institutions, museums or local parks to conduct the searches. For more information refer to the following link: http://www1.nrintra.nps.gov/im/datamgmt/DataMining/datamining.htm.

Biological/Natural Resources Data

- NPSpecies/NatureBib
- Park and regional offices/archives
- Local document library
- Library catalogs (e.g., academic or research institutions)
- Voucher collections (museums, parks, universities)
- Online literature databases (e.g., Biosis and FirstSearch)

Geographic Data

- Regional GIS Archive
- Park GIS Archive
- NPS Metadata and Data Store
- Federal, state and local GIS clearinghouses

Data discovery is an integral part of project development but data discovery efforts should not be limited solely to project development needs. This should be an ongoing process requiring regular data searches, communication, and visits to NCRN parks in order to ensure that the I&M Program maintains as much relevant material pertaining to the parks as possible. It should be noted that such sources need not be directly connected to network parks but may instead pertain to methodologies or protocols that could assist network personnel with the development of a more productive program. Encouraging data sharing with network parks will assist in this process and may alleviate the need for regular searches of park records. The goal of the NCRN I&M Program is to conduct general data discovery efforts on an annual basis. *(Note: the SOP is in development.)*

Information about parks in the NCRN has been collected for many years, beginning long before the inception of the I&M Program. Much of the information collected during the data mining process is likely to predate the I&M Program and this data is referred to as *legacy data*.

Legacy Data—Data collected prior to the inception of the I&M Program.

5.1.2 Programmatic Data Sources

Data collection initiated by the I&M Program usually takes place as part of either natural resources inventories or vital signs monitoring projects, but can also include collaborative efforts between the I&M Program and other regional programs or institutions. Data for these projects are typically collected by I&M personnel, park staff, or by cooperators/contractors.

Natural Resources Inventories

- Resource inventories were designed to identify the primary resources of each park. Appendix E lists the 12 basic inventories conducted by the I&M Program.

Vital Signs Monitoring Projects

- The I&M Program and the network parks identified 21 priority vital signs for monitoring and they are identified in appendix F. Vital signs are measurable physical, chemical, and biological elements and processes of park ecosystems that are selected to represent the overall health or condition of a park's natural resources. Early detection of potential problems allows park managers to take steps to restore ecological health of park resources before serious damage can happen.

Field Studies

Biological inventories and long-term monitoring projects are the most common examples of field studies conducted by the I&M Program. In addition to general SOPs that define network-wide requirements, protocol-specific SOPs are developed detailing procedures and/or methodologies specific to each protocol. The data manager will work closely with the principal investigator and network staff to develop these guidelines and methodologies connected with data collection, processing, storage, and maintenance of project data. This may range from detailing the proper usage of data entry forms or databases to outlining calibration procedures for automated data loggers. Refer to the individual protocols for protocol specific SOPs. The data manager and project leader are responsible for ensuring that all SOPs are followed. There are numerous methods for collecting data in the field ranging from paper field forms to automated computers and sensors that collect information independent of user input. The collection technique used will depend on the project and the specific data collection requirements. Refer to appendix G for guidance on the various data collection methods available and recommended procedures that accompany the use of each method.

Data Entry, Formats and Standards

The data entry requirements for various projects will depend on the data collection techniques used in the field. If data are collected on paper field forms then the data entry requirements (and QA/QC requirements) will be greater than if data are collected via hand-held computers or data loggers.

Database Standard

The NCRN I&M Program has adopted the Natural Resources Database Template (NRDT) as the foundation for its databases and will provide assistance in database development for any programmatic project.

All field studies conducted by the I&M Program have a MS Access database associated with them. The NCRN I&M Program has adopted the NRDT (refer to the NRDT website) as the foundation for its databases and will provide assistance in database development for any programmatic project. The database template is highly flexible and can be modified and customized for each project to meet the needs and requirements of the researcher. The database promotes quality control and data entry efficiency by incorporating mechanisms such as pick lists and auto-populated fields. A more detailed description of the NRDT and the NCRN database requirements can be found in appendix C.

Changes to Data Collection Procedures/Protocols

Changes to established data collection procedures are discouraged unless there are valid reasons for altering the methodologies. Ideally, all problems should be identified during the design and testing stages of the project and changes implemented prior to the collection of any field data. Protocols should attempt to identify any foreseeable issues that might occur as well as contingencies to address them. Inevitably, unforeseen problems may occur which require procedures/protocol revision after data collection has begun. Significant changes to the protocols must be approved by the project lead, principal investigator, Contract Technical Representative (COR) if applicable, and the data manager. The project leader and/or COR must evaluate the proposed changes and determine if additional peer review is required before accepting them.

Altering data collection procedures or protocols may also occur as a result of the comprehensive review that all monitoring protocols undergo every five years. During the review, data is analyzed to determine if the current protocol accomplished its goal. If it is concluded that the protocol in its present form has not achieved the desired results changes could be recommended. Once again, all changes must be approved by the project lead, principal investigator, COR, and data manager.

5.1.3 Non-Programmatic Data Sources

Two distinct kinds of non-programmatic data were defined in the introduction of this chapter: (1) *NPS data* and (2) *non-NPS data*. The following section will further describe the data that fall into these classes and the manner in which the NCRN manages such data.

Non-Programmatic Data – NPS Data

A large percentage of data collected in network parks are collected by park personnel working on projects initiated at the park level or by other NPS regional or national programs. The data collected and products produced by such efforts provide a great deal of information about park natural resources and are therefore very relevant to the mission of the I&M Program. It is important that park, regional, and network staff work closely together to ensure that information is maintained in a manner that promotes data sharing.

Park Data—NCRN parks often use base funding or receive funding through NRPP (Natural Resources Protection and Preservation) or Regional (R-NRPP) programs.

- Park-based biological inventories — Data from park initiated biological inventories can be used to supplement network-level inventories conducted by the I&M Program.
- Park-based monitoring projects — Parks also generate valuable park-level monitoring projects (such as vegetation and water quality).
- Park and multi-park based or regional projects — Other studies or projects conducted at the park or regional level that do not fall into one of the previous two categories (e.g., restoration projects).

Regional and National Programs—NPS regional and national programs support all of the parks within the NCR and also provide a good resource for natural resources information. Many of the following programs collect information that is valuable to the I&M Program. For example, the NCRN Exotic Plant Management Team personnel often advise the I&M Program of new occurrences of exotic plant species in network parks.

- Air — The I&M program will rely on the data, trends analyses, and reports from national monitoring networks to obtain trends for its air vital signs.
- EPMT — Exotic Plant Management Teams (EPMTs) maintains APCAM (Alien Plant Control and Monitoring Database) in order to track exotic species and their treatments (EPMT Web Page).
- Fire Program — Data concerning the occurrence of fires within the NCRN are maintained both at the individual park level and regionally. National databases such as Fire-Pro, SACS and the soon-to-be implemented Fire Program Analysis (FPA) package (http://fpa.nifc.gov/) have been and will be used to maintain information regarding resources devoted to fires as well as fire occurrences. The NPS is also involved in efforts such as the Joint Fire Science Program (http://jfsp.nifc.gov/) that provides scientific information and support for fuel and fire management programs.
- GIS — The I&M Program is supported by regional GIS specialists to help ensure that spatial data are available and accurate. Much of this data is also available through the Spatial Data Clearinghouse (http://www.nps.gov/gis/data_info/).
- Water — The regional Water Resources program synthesizes, analyzes, and interprets water resources data collected by parks. All of the field data collected by the parks and regional water resources program is entered into STORET.

- Wildlife Management — The regional wildlife biologist provides management support to the parks in the NCR. The Wildlife Management Program coordinates long-term monitoring and assessment of wildlife populations.

The NCRN I&M Program benefits by being collocated with many of the above regional programs. This allows for greater communication and data sharing between regional and network personnel.

Non-Programmatic Data – External Data—As was the case in the previous section, relevant information important to the mission of the I&M Program can often be obtained from entities outside of the NPS. In fact, certain vital signs (Air and Land Use/Land Cover Change) rely almost entirely on data collected by other agencies or institutions. The agencies or organizations that collect these data have the expertise to conduct the proper quality control procedures and the capability to function as a repository and clearinghouse for the validated data.

When the data are not maintained in-house, two modes of acquisition are required: a low-level monitoring of trends (i.e., is there a problem?), and a higher-intensity, site-specific acquisition of data so that the extent of a problem can be determined (i.e., how much of a problem and where do I get the hourly, site-specific data?) Appendix H indicates which agencies collect the data, which agencies or organizations summarize the data, and how the I&M Program data manager will access both types of data.

5.2 DATA PROCESSING

All data used and maintained by the NCRN, regardless of source or origin, must meet certain standards. To meet these standards, all data and/or data products are subjected to certain processing steps. All of these steps will be discussed in more detail in the following chapters of this plan so they will only be touched on briefly here. These steps include *data quality assurance and control*, *data documentation,* and *data integration*.

It is often easier to perform these processing steps on programmatic data because such datasets were created by network personnel and have remained under the control of the network since their creation. This is not the case for non-programmatic datasets that may have originated at the park level or that were created by other agencies. Quality control, documentation, and integration of non-programmatic data are often more of a challenge.

Programmatic Raw Data—Field crews and project staff are responsible for entering the data into the project database in a timely manner (i.e., within a week of collecting it). Once the data are entered into the project database, project personnel must perform QA/QC checks on the data (see chapter 6 of this plan for more information on QA/QC). If the data were entered directly into the database in the field, the crew should still review the data to ensure that it makes sense and that something was not entered incorrectly. This should be done prior to integrating the new data with the master project database. All monitoring protocols will include an SOP detailing the project database including data entry and review instructions.

Field crews are required to periodically forward the data manager the project data files. Refer to individual protocols for requirements regarding forwarding data. The data manager maintains the master copy of the database and updates it with data files received from the field crews. All raw data are archived along with any field data sheets. All datasets developed by the NCRN will be documented with metadata. The metadata will be updated throughout the life of the dataset and accompany any data that is made available for distribution. Refer to chapter 7 of this plan for more information on metadata documentation.

Vouchers—Voucher specimens, including photographs, are the best way to document the occurrence of a species in the network parks. Accordingly, any voucher specimen or photograph collected in a network during an I&M-related project must be processed according to the guidelines and procedures described in appendix D. It is the intent of the NCRN I&M Program to

work with the network parks and regional curatorial staff to try and develop a standard voucher processing procedure that can be applied to all regional, network, and park projects.

Non-Programmatic Data—Of the two distinct types of non-programmatic data, NPS data products (national, regional and park) will often require less processing as the I&M Program shares many of the file standards with network parks and regional programs. The NCRN, through integration efforts with network parks, will encourage parks to adopt and use I&M standard applications (such as the NRDT, NPSpecies and Dataset Catalog) in the hopes of developing a more streamlined, seamless data transfer system.

Goal

The NCRN plans to integrate with network parks in an effort to promote common standards and seamless data sharing.

Many of the non-NPS agencies or organizations that collect and house information used by the NCRN are required to meet their own quality control and documentation standards. Network personnel accessing data from other agencies will familiarize themselves with the standards used by those agencies and assess the data quality prior to using any of the data. If necessary, additional documentation will be added to the metadata. The lack of network involvement in data collection and initial processing steps underlines the importance of obtaining comprehensive data set documentation from those who manage the data.

When accessing data to determine basic trends analysis for vital signs such as air quality contributing agencies prepare basic summary statistics and trend analyses which reduces the processing burden of the I&M Program. Trend analyses performed by these national programs (appendix H) will be the default parameters monitored by the I&M Program. However, if more intensive analyses are intended, such as conducting in-depth analyses at a specific location rather than monitoring regional trends, raw data may be required which will require more processing steps. In such cases, the specific protocols provide the necessary data processing requirements.

5.2.1 General Data Processing Requirements

The following paragraphs deal with processing requirements that are common to both programmatic and non-programmatic data. The actual steps taken to process the data may vary but *all* data sets must meet the same requirements

Data Mining—All information collected during data discovery process is maintained at the CUE either electronically or in hard copy format (depending on how it was collected). The type of data determines how it is processed and maintained. Any data collected during data discovery should be accompanied by as much documentation (metadata) as possible. This is especially true for legacy data because as the datasets age, less and less is known about them.

Digital Library

The NCRN intends to scan all hard copy materials and create a digital library.

If legacy data are collected in a digital format the information should be converted to current file formats compatible with the current software standards. This will help integrate older datasets with current standards. The process of converting from one format or standard to another requires that data is checked to confirm that information has not been lost or altered in the conversion process. See chapter 10 for more information.

In the past, hard copy materials were maintained and cataloged for reference purposes. In the future, the NCRN intends to scan hardcopy references and materials, saving them as .pdf files in order to create a digital library.

NPSpecies, NatureBib and Biodiversity Data—All information regarding species occurrence and biodiversity, whether the data originated form an I&M project or not, should be entered into the NPSpecies database. Each record entered into the database must be supported with some form of evidence whether it be an observation, reference, or voucher specimen.

Incomplete information often leads to questionable records. More specifics on entering data into NPSpecies can be found in appendix I. A description of the NPSpecies and NatureBib databases can be found in chapter 3.

Geospatial Data—All geospatial data (GIS and GPS) must be stored in the proper format (e.g., shapefile or geodatabase feature class), have the correct spatial reference information and be checked for topological correctness. Each geospatial data file must also be accompanied by FGDC compliant metadata. Geographic datasets should also meet National Mapping Accuracy Standards (NMAS). For more specifics regarding GIS and GPS data refer to the GIS SOP in appendix C and the GPS SOP in appendix J.

Adhering to these standards is important for every project's geospatial data but especially relevant for imagery and datasets acquired in connection with the Land Use/Land Cover Change vital signs. Remote sensing datasets such as satellite imagery or aerial photography will require varying levels of processing depending on how they are received. These may include geospatial processing and/or spectral processing. Ideally, all spatial datasets will be received in a geo-referenced format and may only require geographic transformations to meet NCRN standards. Varying degrees of spatial and spectral processing may be necessary to adequately answer the proposed questions. For specifics regarding the processing steps of imagery acquired for the Land Use/Land Cover Change vital signs, refer to the protocols.

Credits

Developed by	Geoffrey Sanders (NCRN)
Other contributions included:	Marcus Koenen (NCNR); Debbie Angell (SODN); Data Management Plan Working Group

Chapter 6

Quality Assurance / Quality Control

6.1 QUALITY ASSURANCE (QA) AND QUALITY CONTROL (QC)

The I&M data management program must ensure that our projects produce and maintain data of the highest possible quality. Director's Order #11B, "Ensuring Quality of Information Disseminated by the National Park Service," states that all information (e.g., brochures, research and statistical reports, policy and regulatory information, and general reference information) distributed by the NPS (including information obtained from sources outside of the NPS) must be accurate, reliable, and timely in nature. In order to disseminate accurate information, we must have confidence in the data we use. All data analyses, reports, and publications require data of documented quality that minimize error and bias. Data of inconsistent or poor quality can result in incorrect interpretations and conclusions and improper management applications.

The NCRN will establish and document protocols for identifying and reducing error at all stages in the data lifecycle. These stages include project planning, data collection, data entry, verification and validation, processing, and archiving. Although a data set containing no errors would be ideal, the cost of attaining 95%–100% accuracy may outweigh the benefit. Therefore, we consider at least two factors when setting data quality expectations:

- frequency of incorrect data fields or records
- significance of error within a data field

We are more likely to detect an error when we work with clearly documented data sets and understand what a "significant" error is within *that* data set. The significance of an error can vary with data sets and depends on where the error occurs. For example, a two-digit number off by one decimal place is a significant error. A six-digit number, with the sixth digit off by one decimal place, is not a significant error. But one incorrect digit in a six-digit species number could indicate a different species. That is a significant error.

6.1.1 National Park Service Mandate for Quality

Current computer hardware and software technology are constantly changing and evolving, sometimes in unpredictable ways. If we expect our current data to be useful to future users, data must survive changes in technology. We can promote data longevity through high-quality documentation and maintenance during all phases of data management: during data collection, entry, verification, and validation. Well-documented data sets are especially important when sharing data.

NPS Director's Order #11B, issued in 2002, promotes information and data quality. It defines *'quality'* as incorporating three key components–*objectivity*, *utility*, and *integrity*.

Objectivity consists of:

- *presentation*, which focuses on whether disseminated information is being presented in an accurate, clear, complete, and unbiased manner within a proper context, and
- *substance,* which focuses on ensuring accurate, usable, and reliable information.

Utility refers to the usefulness of the information to its intended users, from the perspectives of both the network and the general public.

Integrity refers to the security of information, for example, protection from unauthorized access or revision to ensure that the information is not compromised through corruption or falsification.

Director's Order #11B also specifies that information must be based on reliable data sources, which are accurate, timely, and representative of the most current information available. These standards apply to NPS-generated information, and information provided to the NPS by other parties if the NPS disseminates or relies upon this information.

6.2 GOALS AND OBJECTIVES

We must ensure that a project produces data of the right type, quality, and quantity to meet project objectives and user needs. Quality criteria should be set at a level proportionate to project-specific objectives, and these criteria should indicate the level of quality acceptable for resulting data products. The EPA (2003) defines data quality objectives as qualitative and quantitative statements that:

- clarify the intended use of the data
- define the type of data needed to support the decision
- identify the conditions under which the data are to be collected
- specify tolerable limits on the probability of making a decision error due to uncertainty in the data

The most effective mechanism for ensuring that a project produces data of the right type, quality, and quantity, is to provide procedures and guidelines to assist the researcher in accurate data collection, entry, and validation. NCRN will initiate a comprehensive set of SOPs and data-collecting protocols for quality control, field methodologies, field forms, and data entry applications with some built-in validation.

Although specific QA/QC procedures will depend upon the individual vital signs being monitored and must be specified in the protocols for each monitoring vital sign, some general concepts apply to all network projects. These general guidelines will ensure that all data collected are checked for integrity before being integrated into the monitoring program databases.

6.3 QA AND QC MECHANISMS

Palmer (2003) defines *Quality Assurance* (QA) / *Quality Control* (QC) as the following:

- *Quality Assurance* (QA): "an integrated system of management activities involving planning, implementation, documentation, assessment, reporting, and quality improvement to ensure that a process, item, or service is of the type and quality needed and expected by the consumer."
- *Quality Control* (QC): "a system of technical activities to measure the attributes and performance of a process, item, or service relative to defined standards." QA procedures maintain quality throughout all stages of data development. QC procedures monitor or evaluate the resulting data products.

QA/QC mechanisms are designed to prevent data contamination, which occurs when a process or event other than the one of interest affects the value of a variable and introduces two fundamental types of errors into a data set:

- Errors of commission: include those caused by data entry and transcription errors or malfunctioning equipment. They are common, fairly easy to identify, and can be effectively reduced upfront with appropriate QA mechanisms built into the data acquisition process, as well as QC procedures applied after the data have been acquired.

- Errors of omission: often include insufficient documentation of legitimate data values, which could affect the interpretation of those values. These errors may be harder to detect and correct, but many of these errors should be revealed by rigorous QC procedures.

QA/QC procedures applied to ecological data include four procedural areas (or activities), ranging from simple to sophisticated, and inexpensive to costly:

1. defining and enforcing standards for electronic formats, locally defined codes, measurement units, and metadata
2. checking for unusual or unreasonable patterns in data
3. checking for comparability of values between data sets
4. assessing overall data quality

Much QA/QC work involves the first activity (defining and enforcing…), which begins with data design and continues through acquisition, entry, metadata development, and archiving. The progression from raw data to verified data to validated data implies increasing confidence in the quality of the data through time (figure 6-1).

6.4 DATA COLLECTION

Careful, accurate recording of field observations in the data collection phase of a project will help reduce the incidence of invalid data in the resulting data set. Unlike a typographical error that occurs when a recorded observation is incorrectly transferred from a paper field form to a digital database, an incorrect entry in the field cannot be easily corrected. Therefore, attention to detail during data collection is crucial to overall data quality.

The different data collection options discussed in chapter 5 and appendix F have pros and cons when it comes to data quality control and assurance. Not all of these collection methods will be realistic or feasible options for all projects. Even though paper field forms may not incorporate the QC mechanisms that some of the other devices do, they may still be the most appropriate option for some projects. These mechanisms for data collection are further discussed in the NCRN Data Collection Guidelines document (appendix F) but figure 6-2 illustrates the degree to which the different data collection mechanisms promote data quality and integrity.

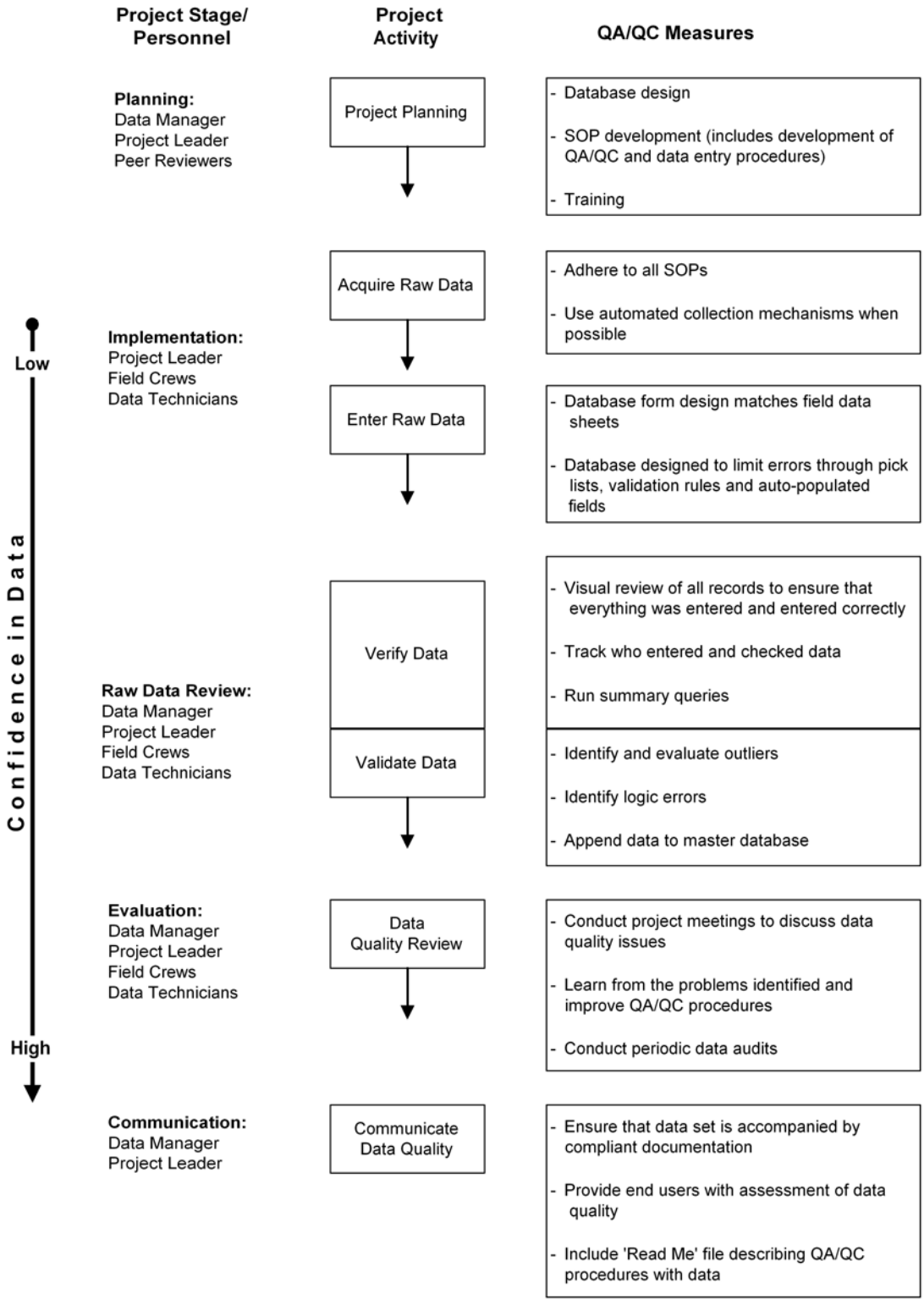


FIGURE 6-1: DATA EVALUATION STAGES AND QA/QC WORKFLOW

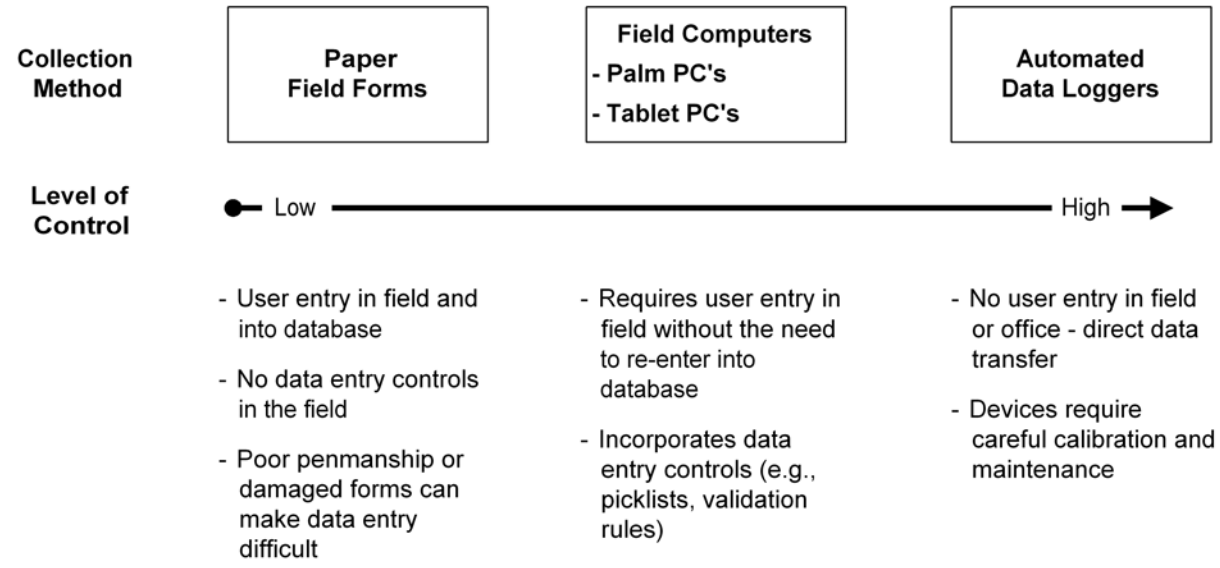


FIGURE 6-2: FIELD DATA COLLECTION OPTIONS AND THE DEGREE TO WHICH THEY ENFORCE QUALITY CONTROL

6.4.1 Methods for Reducing Data Collection Errors

A number of alternatives exist for field data collection; all have specific QA/QC procedures associated with them. Some general QA/QC guidelines are listed below but refer to appendix F for more specific guidance relating to the various data collection methods.

- Use a formatted, project-specific data sheet as opposed to a field notebook.
- Use a handheld computer (PDAs or tablet PCs) for data collection whenever possible.
- Use automated data loggers where appropriate.
- Use a handheld tape recorder.

Consider calibration, maintenance, and minimum timing requirements of field equipment. Accurate field measurements are possible only if field equipment is regularly calibrated and maintained. Where appropriate, consult reference manuals for recommended calibration and maintenance procedures. Once in the field, allow sufficient time for field equipment to adjust to its environment so it will record accurate measurements (for example, when using water quality probes and GPS units). Researchers should consider maintaining records of equipment calibration and failures that accompany their field data whenever possible.

Be organized and keep a log. Organization is the key to good data collection methods. Maintaining a log of important decisions and events will help clarify information and contribute to an accurate report.

Ensure that field crews receive proper training. Although protocols and SOPs are in place, they cannot guarantee that high quality data will be collected. Prior to routine data collection for a project, conduct training sessions to ensure that field personnel have a clear understanding of data collection procedures described in the SOPs. A training program may also include a process to certify that field staff understand and can perform the specified data collection procedures. The development of a training

manual may be helpful for long-term monitoring data collection efforts and those that will involve a large number of field staff. Palmer and Landis (2002) provide an outline for a training manual and suggestions for planning training sessions.

Perform quantitative assessments of data quality. Repeating measurements is the primary tool for performing quantitative assessments of data. Project managers (Project COTR) should periodically review the work of field technicians to ensure that their work does not drift from standards during the course of the field season. Quantitative assessments may be considered if staff and funding are available, and Palmer and Landis (2002) describe several approaches that can be used.

6.5 DATA ENTRY

'Data entry' is the initial set of operations where we transfer raw data from paper field forms into a computerized form linked to database tables. Spreadsheets should not be used for data entry (data can be exported to a spreadsheet for manipulations post entry). When data are gathered or stored digitally in the field (e.g., on a data logger), data entry takes place in the field and data is transferred (downloaded) to a file in an office computer where they can be further manipulated or uploaded directly into the project database.

Transferring data from field projects into the computer seems a fairly simple task. But the value of the data depends upon their accuracy, and we must feel confident about the overall quality of the data. Without proper preparation and some established guidelines, the quality and integrity of the data can be questionable. Ideally, data entry occurs as soon as possible—immediately after data collection is completed, or as an on-going process during long projects—by a person who is familiar with the data. The primary goal of data entry is *to transcribe the data from paper records into the computer with 100% accuracy*. Yet, we know that a few transcription errors are unavoidable during data entry. Thus, all data should be checked and corrected during a data verification process.

The primary goal of data entry is *to transcribe the data from paper records into the computer with 100% accuracy*.

The data manager, along with the project manager, should provide training in the use of the database to all data entry technicians and other users. The project manager makes certain that data entry technicians understand how to enter data and follow the protocols. Data entry technicians are responsible for becoming familiar with the field data forms and differences in handwriting. They must also become familiar with the database software, database structure, and any standard codes for data entry used by the network. At a minimum, data entry technicians should know how to open a data entry form, create a new record, and exit the database properly. They must learn how to commit both a "field" entry and a "complete record" entry and to correct mistakes made while typing.

6.5.1 Methods for Reducing Data Entry Errors

Enter or download data in a timely manner. All data should be entered or downloaded into the project database as soon as possible, preferably at least once a week. Try to avoid delaying data entry until all the project data have been collected. Downloaded "raw" data files should be maintained separately from the main dataset. If the field crew is operating remotely, data files should also be forwarded to the Network data manager for storage and incorporation into the master dataset.

Design efficient data entry forms and methods. A data entry form that mimics the field data forms can effectively reduce manual data entry errors due to the 1:1 correspondence of the attributes. A strategy to distinguish between validated data and newly entered data should be adopted. Data can be entered into an empty, fresh database table to avoid contaminating existing data and the new data appended to the master data only after formal verification, validation, and documentation. Alternatively, we can include validation attributes that indicate which data have been checked and validated by the project manager in the database. Regardless of strategy, we must clearly document the process for validation in the protocol data management SOP.

Build automated error checking features into the database. The most robust QA/QC measures for data entry should be built into the database design to perform automatic validation checks of data. Data entry forms reduce transcription errors through auto-filled fields, range limits, pick lists, and spelling checks. They provide controlled access to the database (i.e., forms are set for data entry only, which prevents accidental deletion or alteration of existing data). They control the sequence of data entry (i.e., certain fields require an entry before more information can be entered). They warn the operator when errors are made and provide an opportunity for correction before the data are committed to a file.

- Auto-filled fields. Whenever possible, the data in a field should be auto-filled by the computer.

- Range limits. Where the appropriate values for a particular field span a finite range, the database should check the entered value against the specified minimum and maximum values for that parameter. When a value is outside the accepted range, a warning message appears and asks the user to reenter a valid value. For some fields, values outside a specified "normal" range may be acceptable. In this case, the warning message asks the user to verify the entry before continuing.

- Pick lists. The data entry application may also use pop-up pick lists for standardized text items where spelling errors can occur. A pick list may also be used when only certain entries are acceptable.

- Unique constraints. Duplicate and incorrect data entry can often be caught with the application of unique constraints on data entry fields. These constraints are particularly useful when importing data from other applications.

Provide a clean, organized work environment. Desktop space near the computer should be free of clutter and distractions that could the technician to lose her place. There should be enough space for two stacks of paper documents, one from which data are being entered and one from which data have been entered. A pad or notebook and some fine colored markers should also be available for making notes. (The need for a clean workspace also applies to the verification and validation phases.)

If possible, use two data entry technicians for data entry. When one technician reads the data from the field data forms and another enters them into the computer, the work is often faster and results in a lower error rate. If only one person is available, that person should work at a slower pace to avoid errors. Like many monotonous tasks, data entry can be done in a personal rhythm that reduces the tediousness of the work.

6.6 VERIFICATION AND VALIDATION PROCEDURES

We appraise data quality by applying verification and validation procedures as part of the QC process. These procedures are more successful when preceded by effective quality assurance practices. *Data verification* checks that the digitized data match the source data, while *data validation* checks that the data make sense. It is essential that we validate all data as truthful and do not misrepresent the circumstances and limitations of their collection. Failure to follow SOPs for data entry, validation, and verification will render a data set suspect. Although data entry and data verification can be handled by personnel who are less familiar with the data, validation requires in-depth knowledge about the data.

The data manager establishes SOPs for verification and validation and provides them to the project manager. The project manager will ensure that the SOPs are followed. While there are certain standard verification and validation steps that all projects should follow, due to variation among projects, specific verification and validation procedures will be developed for each project or protocol to ensure that data quality is properly evaluated. The project manager or designee will validate the data after verification is complete. The project manager is also responsible for reviewing all data products and reports before they are released outside the Network. The data and project managers will evaluate the results of verification and validation and

determine any procedural or data form revisions that may be indicated by the results. Technicians will follow the SOPs for verification of data, make required changes, and document those changes.

6.6.1 Methods for Data Verification

Each of the following methods has a direct correlation between effectiveness and effort. The methods that eliminate the most errors can be very time consuming while the simplest and cheapest methods will not be as efficient at detecting errors.

1. Visual review at data entry. The data entry technician verifies each record after input, comparing the values recorded in the database with the original values from the hard copy and immediately corrects any errors. This method is the least complicated since it requires no additional personnel or software. Its reliability depends entirely upon the person keying data and thus, is probably the least reliable data verification method.

2. Visual review after data entry. All records are printed upon the completion of data entry. The values on the printout are compared with the original values from the hard copy. Errors are marked and corrected in a timely manner. When one technician performs this review, the method's reliability increases if someone other than the person keying data performs the review. Alternatively, two technicians can perform this review. One technician reads the original data sheets (the reader), and the second reads the same data on the printout (the checker).

3. Duplicate data entry. The data entry technician completes all data entry, as normal. Random records are selected (every *n*th record) and entered into an empty replica of the permanent database, preferably by someone other than the person keying the permanent data. Then we use a query to automatically compare the duplicate records from the two data sets and report any mismatches of data. Then we manually review any disparities and correct if necessary. This method adds the overhead of retyping the selected records, as well as the creation of a comparison query, but it becomes increasingly successful as the value of *n* decreases. Professional data entry services frequently use this method.

We can also use the entered data to calculate simple summary statistics with statistical software. These summary statistics can help us catch a duplicate or omitted entry. For example, we can view the number of known constant elements, such as the number of sampling sites, plots per site, or dates per sample. We can pose the same question in different ways; differences in the answer provide clues to errors. The more checks we devise to test the completeness of the data, the greater our confidence that we have completely verified the data.

To minimize transcription errors, our policy is to verify 100% of records to their original source by permanent staff. If the errors are found during review, then the entire data set is verified again.

6.6.2 Methods for Data Validation

Although the dataset may have been correctly transcribed from original field notes or forms, they still might be inaccurate or illogical. For example, entries of stream pH of 25.0 or a temperature of 95°C in data files raise doubt about their accuracy; and such entries almost certainly are incorrect, whether or not they were properly transcribed from field forms. This process of reviewing computerized data for range and logic errors we call *validation*, and it can accompany data verification *only* if the operator has comprehensive knowledge about the data. More often, validation is a separate operation carried out *after* verification by a project specialist who can identify generic and specific errors in particular data types. Validation efforts should also include a check for the completeness of a data set since field sheets or other sources of data could easily be overlooked.

General step-by-step instructions are not possible for data validation because each data set has unique measurement ranges, sampling precision, and accuracy. Nevertheless, validation is a critically important step in the certification of the data. Invalid data commonly consist of slightly misspelled species names or site codes, the wrong date, or out-of-range errors in parameters with well defined limits (e.g., elevation). But more interesting and often puzzling errors are detected as unreasonable metrics (e.g.,

stream temperature of 70°C) or impossible associations (e.g., a tree 2 feet in diameter and only 3 feet high). We call these types of erroneous data *logic errors* because using them produces illogical (and incorrect) results. The discovery of logic errors has direct, positive consequences for data quality and provides important feedback to the methods and data forms used in the field.

The following general methods can be used to validate data. Specific procedures for data validation depend upon the vital sign being monitored and will be included in the monitoring protocols.

Data entry application programming. Certain components of data validation are built into data entry forms. The simplest validation during data entry is range checking, such as ensuring that a user attempting to enter a pH of 20.0 gets a warning and the opportunity to enter a correct value between 1.0 and 14.0 (or better yet, within a narrow range appropriate to the study area). Not all fields, however, have appropriate ranges that are known in advance, so knowledge of what are reasonable data and a separate, interactive validation stage are important.

A caveat should be interjected regarding the operative word "illegal". Even though a value above or below a given threshold has never before been observed and the possibility that it could occur seems impossible, such an observation is not always an illegal data point. Edwards (2000) points out that one of the most famous data QA/QC blunders to date occurred when NASA's computer programs deleted satellite observations of ozone concentrations that were below a specified level, seriously delaying the discovery of the ozone hole over the South Pole.

Outlier Detection. According to Edwards (2000), "the term outlier is not (and should not be) formally defined. An outlier is simply an unusually extreme value for a variable, given the statistical model in use." Any data set will undoubtedly contain some extreme values, so the meaning of 'unusually extreme' is subjective. The challenge in detecting outliers is in deciding how unusual a value must be before it can (with confidence) be considered 'unusually' unusual.

Data quality assurance procedures should not try to eliminate outliers. Extreme values naturally occur in many ecological phenomena; eliminating these values simply because they are extreme is equivalent to pretending the phenomenon is 'well-behaved' when it is not. Eliminating data contamination is a better way to explain this quality assurance goal. If contamination is not detected during data collection, it is usually only be detected later if an outlying data value results. When we detect an outlier, we should try to determine if some contamination is responsible.

Other exploratory data analyses. Histograms, line plots, and basic statistics can reveal possible logic and range errors. Palmer and Landis (2002) suggest that in some cases, calculations for assessments of precision, bias, representativeness, completeness, and comparability may be applicable and that for certain types of measurements, evaluation of a detection limit may also be warranted (the authors provide examples of procedures that may be applicable). Normal probability plots, Grubb's test, and simple and multiple linear regression techniques may also be used (Edwards 2000; the author provides SAS and Splus code for constructing normal probability plots and examples of output showing normal and non-normal distributions).

6.7 DATA QUALITY REVIEW AND COMMUNICATION

Given that only data and information that adhere to NPS quality standards can be released, strict quality review must be followed. Documentation of the QA/QC standards used in producing the information and that substantiate the quality of the information must be formally documented. Information disseminated to the public must be approved by the appropriate reviewing officials and programs. Also, mechanisms must be in place for receiving and addressing comments/complaints pertaining to the quality of data.

Data Quality Review Methods—The I&M Program will establish guidelines and protocols to ensure compliance with Director's Order #11B. These protocols will document both internal and external review procedures for data and information disseminated outside the network, as well as a process for processing complaints about data quality.

Edwards (2000) suggests the initiation of quality circles, regular meetings of project managers, the data manager, and data management personnel for discussing data quality problems and issues. These meetings promote teamwork attitudes while focusing brainpower on data quality issues. Participants become more aware of quality issues and learn to anticipate problems. Moreover, all participants develop a greater appreciation of the importance of their role in data quality and the entire monitoring effort.

Value of Feedback from QA/QC Procedures—Quality assurance procedures may need revision to improve quality levels if random checks reveal an unacceptable level of data quality. Quality checks should not be performed with the sole objective of eliminating errors, as the results may also prove useful in improving the overall process. For example, if the month and day are repeatedly reversed in a date field, the data entry technicians may require retraining about the month/day entry order. If retraining is unsuccessful in reducing the error's occurrence, the computer program may need to be rewritten so that month and day are entered separately, field length limits are enforced, or a pick list is created. In this manner, the validation process will serve as a means of improving quality as well as controlling the lack of quality.

We can modify field data forms to avoid common mistakes when necessary. With knowledge of validation errors and exploratory data results in hand, the field data forms as the source of the logic errors can be more easily re-evaluated. Often minor changes, small annotations, or adding check boxes to a field form can remove ambiguity about what to enter on the form. Perhaps surprisingly, when we find the same type of validation error occurring repeatedly in different data sets, the field form–not the field crew–is usually at fault. Repeated errors found during validation can also mean that protocols or field training are at fault, which can then be recognized and corrected.

Monitoring Conformance to Plans and Standards—Data managers can use periodic data audits and quality control inspections to maintain and improve their data quality program. They must verify that staff is operating in conformance with the data quality procedures specified in this plan and the protocol specific data management plans. They should track and facilitate the correction of any deficiencies. These quality checks promote a cyclic process of continuous feedback and improvement of the both the data and quality planning process.

Periodic checks by the data manager to see if network staff are adhering to the data quality procedures established in the Data Management Plan and protocol SOPs may include verification of the following:

- Data collection and reporting requirements are being met (this may require auditing field crews or contractors)
- Data collection and reporting procedures are being followed (this may require auditing field crews or contractors)
- Verification and validation procedures are being followed
- Data file structures and maintenance are clear, accurate and according to plan
- Revision control of program documents and field sheets are adequate
- Calibration and maintenance procedures are being followed (this may require auditing field crews or contractors)
- Seasonal and temporary staff have been trained in data management practice

- Metadata collection and construction for the program proceeds in a timely manner
- Data are being archived and catalogued appropriately for long term storage

The results of quality assessments should be documented and reported to the research staff and the network coordinator. The project manager and coordinator are responsible for ensuring that nonconformities in data management practices are corrected.

Communicating Data Quality—The network will use data documentation and metadata to notify end users, project managers, and network management of data quality. A descriptive document for each data set/database will provide information on the specific QA/QC procedures applied and the results of the review. Descriptive documents or formal FGDC-compliant metadata will document quality for spatial and non-spatial data files posted on the internet.

Credits

Developed by:	This chapter was adapted from materials developed by Debbie Angell (SODN).
Other contributions included:	Data Management Plan Working Group

Chapter 7

Data Documentation

Data documentation is a critical step toward ensuring that data sets are useable for their intended purposes well into the future. This involves the development of *metadata,* or information about the data. The Federal Geographic Data Committee (FGDC) defines metadata as:

Metadata – information about the content, quality, condition and other characteristics of data.

Metadata provides a means of documenting the qualitative attributes of datasets, especially those that are distributed to audiences via both the intranet and internet. Knowing that a dataset is accompanied by complete and compliant metadata documentation adds a level of assurance to both the data steward and the potential users because such documentation helps to ensure that the data will be used properly.

7.1 PURPOSE OF METADATA

Data sets sometimes take on lives of their own. Some seem to have the ability to reproduce and evolve on multiple hard drives, servers, and other storage media. Others remain hidden in digital formats or in forgotten file drawers. Once data are discovered, a potential data user is often left with little or no information regarding the quality, completeness, or manipulations performed on a particular "copy" of a dataset. Such ambiguity results in lost productivity as the user must invest time tracking information down, or, in a worst-case scenario, renders the dataset useless because answers to these and other critical questions cannot be found. As such, data documentation must include an up-front investment in planning and organization.

The information provided in metadata notifies potential users about the quality of datasets and helps to ensure that datasets are used properly.

The information needed to understand and ensure the proper use of data is located in a variety of places such as project planning documents and protocols, as well as project databases and management systems. Additional product documentation, such as the degree to which the product us undergone QA/QC, is important so that the user understands the limitations of the data.

7.2 METADATA MANDATE

The importance for metadata is now universally accepted within the data management community. Recent legislative and policy efforts emphasize that data documentation must include an up-front investment in planning and organization.

Executive Order 12906, signed by President William Jefferson Clinton in 1994, mandates federal agencies to "...document all new geospatial data it collects or produces, either directly or indirectly..." using the FGDC Content Standard for Digital Geospatial Metadata (CSDGM, *www.fgdc.gov/contstan.html*). In addition, EO 12906 directs agencies to plan for legacy data documentation and provide metadata and data to the public.

The FGDC Biological Data Profile contains all the elements of the CSDGM and includes additional elements for describing biological data sets. Metadata created in compliance with the Biological Data Profile can be added to the National Biological

Information Infrastructure (NBII, *www.nbii.gov/datainfo/metadata*) Clearinghouse. Although not a requirement, completion of the Biological Data Profile for appropriate data sets is recommended.

All GIS data layers must be documented with applicable FGDC and National Park Service (NPS) metadata standards. The NPS GIS Committee requires all GIS data layers be described with FGDC standards and the NPS Metadata Profile (*nrdata.nps.gov/profiles/NPS_Profile.xml*). While there are numerous tools available for developing metadata, the NPS Integrated Metadata System Plan recommends three desktop applications: Dataset Catalog, ArcCatalog, and the NPS Metadata Tools and Editor. Figure 7-1 illustrates the general strategy for creating and distributing metadata.

This approach uses existing desktop metadata creation applications, as well as an online integrated metadata database (NR-GIS Metadata) and a web based data server (NR-GIS Data Server). NR-GIS Metadata and Data Store will comprise a web-based system to integrate both data dissemination and metadata maintenance. It will be possible to update dataset catalog records in the NR-GIS Metadata database or in the source desktop application (e.g., ArcCatalog, Dataset Catalog, NPS Metadata Tools and Editor). Non-sensitive NR-GIS Metadata records are automatically posted to NPS Focus.

7.3 NPS INTEGRATED METADATA SYSTEM PLAN AND TOOLS

The NPS Integrated Metadata System Plan is limited to three recommended desktop applications including Dataset Catalog (developed by the Inventory and Monitoring [I&M] Program) and two commercial off the shelf metadata tools: ArcCatalog and SMMS. A fourth tool, the Metadata Parser (mp) and its utility in metadata creation is also briefly discussed.

- Dataset Catalog: Dataset Catalog is a tool for cataloging abbreviated metadata on geospatial and biological data sets pertaining to park(s) and/or a network. It provides parks and/or networks a means whereby they can inventory, organize, and maintain information about data set holdings locally. While Dataset Catalog is not intended to be an exhaustive metadata listing, it does assist parks and networks in beginning to meet the mandates of EO 12906. The I&M Program recommends that all relevant datasets at I&M parks and networks be cataloged in at least simple Dataset Catalog format.

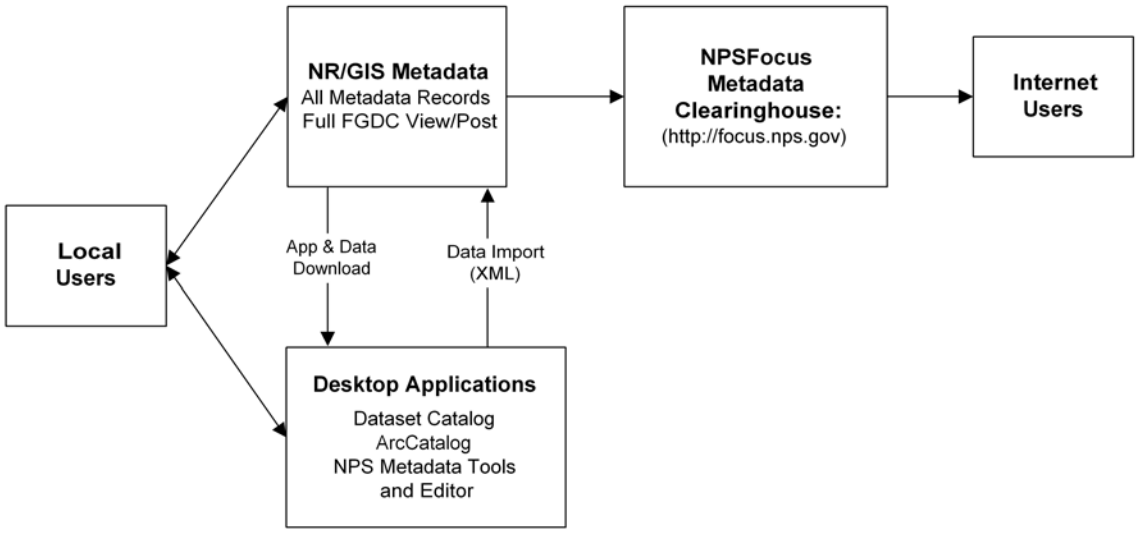


FIGURE 7-1: METADATA CREATION AND DISSEMINATION TOOLS

- ArcCatalog: ArcCatalog is a management tool for GIS files contained within the ArcGIS Desktop suite of applications. With ArcCatalog, users can browse, manage, create, and organize tabular and GIS data. In addition, ArcCatalog comes with support for several popular metadata standards that allow one to create, edit, and view information about the data. There are editors to enter metadata, a storage schema, and property sheets to view the data. With ArcCatalog, users can view GIS data holdings, preview geographic information, view and edit metadata, work with tables, and define the schema structure for GIS data layers. Metadata within ArcCatalog is stored exclusively as Extensible Markup Language (XML) files. The NPS Integrated Metadata System Plan recommends ArcCatalog for gathering GIS-integrated geospatial metadata. An optional but highly recommended extension for ArcCatalog is the NPS Metadata ArcCatalog Extension developed by NPS Midwest Region GIS Technical Support Center.

- NPS Metadata Tools and Editor: NPS Metadata Tools and Editor is a metadata management and editing application that implements two separate interfaces: either 1) as an extension within ArcCatalog (v8.3 or v9.0), or 2) as a standalone desktop interface. The application is intended to be the primary editor for metadata that will be uploaded to the NR-GIS Data Store. The NPS Metadata Tools and Editor application integrates with the NR-GIS Data Store information system by producing XML metadata based on the NPS Metadata Profile. This XML metadata file can then be uploaded to the NR-GIS Metadata and Data Store application.

7.4 METADATA PROCESS

The goal is to create formal, comprehensive metadata documentation that meets both federal (FGDC) and NPS standards. The following section describes how datasets from varying sources are treated with regard to metadata creation. A detailed outline of the steps involved in the creation of compliant dataset and product documentation can be found in the NCRN Metadata SOP (appendix K).

Metadata will initially be created using one of the applications described above. The initial documentation will be limited but as the project/life of the product progresses, the metadata document will become more complete and robust. Dataset documentation is complete when the final product is certified by the project leader. Currently, XML is the current standard for NPS metadata. Accordingly, all spatial and non-spatial data files will be accompanied by an XML metadata file.

7.4.1 Data Documentation

Data utilized by the National Capital Region Network (NCRN) can be grouped into three broad categories based on origin. These categories include programmatic datasets (data collected by, and/or under the direct purview of, the NCRN), non-programmatic datasets (ongoing data collection efforts conducted by entities outside the direct purview of the NCRN), and legacy datasets (primarily data collected prior to the inception of the NCRN).

Programmatic Data — For new projects, metadata development will begin up front, with an interview with the principal investigator to explain what will be needed to properly document the data. In most instances, this will include completion of a basic metadata survey for inclusion in the data manager's project file, as well as submission of supporting documentation (proposal, SOPs, etc.). In addition, a database structure will be developed by, or in close consultation with, the data manager, to ensure compliance with the principles and procedures contained within this plan. Updates and revisions to the metadata will take place as the project progresses and the data evolves.

Non-programmatic Data — Networks and/or prototypes are not the only entities gathering relevant inventory and/or monitoring data pertinent to park management. As mentioned in chapter 5, non-programmatic data falls into two categories: (1) NPS data or data collected by other NPS entities, and (2) non-NPS data or data collected by groups or agencies not associated with the NPS. Data acquired through data mining efforts does not enjoy the same degree of stewardship afforded project data. Nevertheless, non-programmatic datasets require the same degree of documentation as datasets produced by the NCRN. The NCRN will make every effort to capture and assimilate all relevant data through data mining.

Non-programmatic datasets require the same level of documentation as those datasets originally produced by the NCRN.

Ideally, complete and compliant metadata will accompany the dataset or data product but oftentimes this will not be the case as other agencies or organizations may not have as strict metadata requirements. If the available dataset documentation is limited, the outside entity should be contacted and a request made for available metadata and/or a metadata interview will be conducted. A processing and revision log will be maintained with the dataset for capture of additional metadata. Any datasets created based on, or derived from, external datasets should be accompanied by their own metadata record including information on the original data as well as documentation on the process used to create the new dataset.

Legacy Data — In many cases, legacy data are initially identified as part of data mining efforts and, as mentioned previously, data acquired in this manner often lacks quality documentation due to a lower level of stewardship. If pertinent information is missing and the originators of the data or the data steward are not known or available, an "adequate" level of documentation may not be possible. However, the data and all supporting documentation related to it should be assembled and reviewed. Many legacy data sets will need to be converted to a standard database format for incorporation and future analyses. Data entry, validation, and verification procedures will follow those contained within this document. A processing and revision log will be maintained with the dataset for capture of pertinent metadata. If programmatic data products are derived from legacy datasets, the metadata for the derived data product should indicate what datasets were used and contain links to the metadata documents for the relevant legacy datasets.

Chapter 8

Reporting and Analysis

One of the critical goals of the Inventory and Monitoring (I&M) Program is to provide meaningful information that can be used to drive natural resource management in National Parks. In order to provide information, data must first be processed, checked, summarized, analyzed and finally reported to the appropriate information users. The National Capital Region Network (NCRN) I&M Program produces a variety of reports to synthesize new findings and summarize results. To accomplish the tasks of data analysis and reporting the data manager must work closely with other network staff, especially the project leaders and quantitative ecologist, to determine the most appropriate approach to data summary and analysis. This chapter presents the general requirements and processes for summarizing, analyzing, and reporting data.

8.1 DATA SUMMARY AND ANALYSIS

The first step in developing in summarizing and analyzing data is in identifying project objectives. The objectives will dictate what data will be collected and how they must be analyzed. Ideally, databases will be designed with these objectives in mind. In addition, database consideration must keep export functions in mind in order to format data appropriately and efficiently for analysis.

> Goals
> Data must be properly analyzed and reported to managers and decision makers in a timely manner. This can be accomplished only if project goals are clearly identified.

Each monitoring protocol details the required analyses. MS Access can be used to generate basic summary statistics (e.g., mean, sum, maximum, minimum, standard deviation). Data should be exported to more robust statistics packages (e.g., MiniTab or Systat) when more complex analyses are required. Distance 4.0, for example, will be used to estimate animal population abundance from distance sampling data (see the NCRN Deer Monitoring Protocol).

The project leader, quantitative ecologist and data manager must work together to identify the proper way to summarize or export the data from project databases for analysis and reporting. In some cases, the Data Manager will need to develop the necessary queries, macros, modules and reports to extract the desired data as needed.

8.2 REPORTING REQUIREMENTS

The national I&M Program requires that all networks submit reports to inform program administrators, regional and national NPS staff, as well as the public, of the status of the projects and the Network program in general. Reporting requirements may vary depending on the project or vital sign but all requirements will be clearly defined in the protocol (as well as in the agreement or contract if applicable). For instance, if the protocol for Land Use Change specifies that analyses are conducted every five years, the reporting schedule should be synchronized with the analyses.

Standard reports are detailed in chapter 7 of the NCRN Monitoring Plan (Koenen et al. 2005). The standard operating procedures (SOPs) are listed below:

- SOP Analysis and Trend Reports – designed to report the results of trend analyses and to report correlations between resources. These reports will not be prepared annually as the data required to generate such reports will most likely take multiple years to collect. The exact amount of time/data required to conduct the analyses and generate the report will vary among vital signs (SOP – to be developed).

- SOP Park Status Reports – designed to provide a park specific summary of data and information collected by the NCRN I&M Program. These reports should be directed towards the specific management needs of a park and address how data produced by the I&M Program relates to management needs and performance goals such as GPRA. An SOP for developing park status reports can be found in appendix K.

- SOP Annual and Final Reports – prepared by the project leader or cooperator. Annual and final reports contain a summary of the entire project including major findings. Conclusions include public interest highlights, management recommendations, and research recommendations. SOPs pertaining to these reports are included in the Project Deliverables Guidelines document in appendix C.

Credits

Developed by:	Chapter was adapted from materials developed by Gareth Rowell (Prairie Cluster) and Michael Williams (HTLN), John Boetsch (NCCN) and Doug Wilder (CAKN)
Other contributions included:	Marcus Koenen (NCRN); John Paul Schmit (NCRN)

Chapter 9

Data Dissemination

9.1 DATA OWNERSHIP – THE NATIONAL PARK SERVICE POLICY

The National Park Service (NPS) defines conditions for the ownership and sharing of collections, data, and results based on research funded by the United States government. All cooperative and interagency agreements, as well as contracts, should include clear provisions for data ownership and sharing as defined by the NPS:

- All data and materials collected or generated using NPS personnel and funds become the property of the national park.
- Any important findings from research and educational activities should be promptly submitted for publication. Authorship must accurately reflect the contributions of those involved.
- Investigators must share collections, data, results, and supporting materials with other researchers whenever possible. In exceptional cases, where collections or data are sensitive or fragile, access may be limited.

The Office of Management and Budget (OMB) ensures that grants and cooperative agreements are managed properly. The OMB circulars establish some degree of standardization government-wide to achieve consistency and uniformity in the development and administration of grants and cooperative agreements. Specifically, OMB Circular A-110 establishes property standards within cooperative agreements with higher institutions and non-profit organizations. Section 36 of Circular A-110, "Intangible Property" describes administrative requirements pertinent to data and ownership. The following is a brief synopsis of how products and data are treated with regard to ownership:

Products

- Institutions receiving NPS funding can copyright materials they produce but the NPS can reserves the right to use, publish or reproduce this work for federal purposes at no expense.

Data

- The NPS also has the right to use the original data produced by the recipients of an award.
- The NPS can also authorize others to use this data for federal purposes.

9.1.1 Establishing Data Ownership Guidelines

The National Capital Region Network (NCRN) has established guidelines for the ownership of data and other research information. To ensure that proper ownership, format, and development of network products is maintained, all cooperative or interagency work must be conducted as part of a signed collaborative agreement. Every cooperative or interagency agreement or contract involving the NCRN must either include or incorporate by reference OMB Circular A-110 cited under the *Reports and Deliverables* Section of all agreements and contracts. The following is an example of appropriate language to use when citing Circular A-110:

"As the performing organization of this agreement, *institution or organization name* shall follow the procedures and policies set forth in OMB Circular A-110."

Every cooperative or interagency agreement or contract must include a list of deliverables and products clearly defined within each agreement or contract. Details on formatting and media types for all required deliverables will be provided to the cooperator or contractor in the form of SOP's. Agreements and contracts must list all products expected to result from the project. These include, but are not limited to, field notebooks, photographs (hardcopy and digital), specimens, raw data, and reports.

The following statement must be included in the Reports and Deliverables section of all NCRN agreements and contracts:

"All reports and deliverables must follow the most recent version of the National Capital Region Product Specifications."

Researchers should also provide a schedule of deliverables (as specified in the SOP) that includes sufficient time for NPS review of draft deliverables before scheduled final submissions. This schedule can be included in the network's project management/tracking software so network staff are aware of project timelines.

9.1.2 Data Classification: protected vs. public

All data and associated information from I&M activities must be assessed to determine their sensitivity. This includes, but is not limited to, reports, metadata, raw and manipulated spatial and non-spatial data, maps, etc. Network staff must carefully identify and manage any information that is considered sensitive. The network must clearly identify and define those data needing access restrictions and those to make public.

The Freedom of Information Act, 5 U.S.C. § 552 (FOIA) stipulates that all United States Government agencies must provide access to data and information of interest, that are not protected from disclosure by exemptions, to the public. FOIA, as amended in 1996 to provide guidance for electronic information distribution, applies to records that are owned or controlled by a federal agency, regardless of whether or not the federal government created the records. Under the terms of FOIA, agencies must make non-protected records available for inspection and copying in public reading rooms and/or the Internet. Other records however, are provided in response to specific requests through a specified process. The Department of the Interior's revised FOIA regulations and the Department's Freedom of Information Act Handbook can be accessed at http://www.doi.gov/foia/ for further information. Network staff should consult their Regional FOIA Coordinator for more specific guidance whenever a particular instance of sharing information has the potential to involve protected information.

9.1.3 What to Protect

NPS is directed to protect information about the nature and location of sensitive park resources. Public access to data can be restricted under the following legislation:

- NPS Director's Order #66 (draft)
- National Parks Omnibus Management Act (16 U.S.C. 5937)
- National Historic Preservation Act (16 U.S.C. 470w-3)
- Federal Cave Resources Protection Act (16 U.S.C. 4304)
- Archaeological Resources Protection Act (16 U.S.C. 470hh)

Through these regulations, information that could result in harm to natural resources can be classified as 'protected' or 'sensitive' and withheld from public release (National Parks Omnibus Management Act [NPOMA]).

The following guidance for determining whether information should be protected is suggested in the draft Director's Order #66 (the final guidance may be contained in the Reference Manual 66):

- Has harm, theft, or destruction occurred to a similar resource on federal, state, or private lands?
- Has harm, theft, or destruction occurred to other types of resources of similar commercial value, cultural importance, rarity, or threatened or endangered status on federal, state, or private lands?
- Is information about locations of the park resource specific enough so that the park resource is likely to be found at these locations at predictable times now or in the future?
- Would information about the nature of the park resource that is otherwise not of concern permit determining locations of the resource if the information were available in conjunction with other specific types or classes of information?
- Even where relatively out-dated, is there information that would reveal locations or characteristics of the park resource such that the information could be used to find the park resource as it exists now or is likely to exist in the future?
- Does NPS have the capacity to protect the park resource if the public knows its specific location?

Natural Resource information that is sensitive or protected requires the following:

- Identification of potentially sensitive resources
- Compilation of all records relating to those resources
- Determination of what data must not be released to the public
- Management and archival of those records to avoid their unintentional release

Classification of sensitive I&M data will be the responsibility of the NCRN staff, the park superintendents and resource managers, and investigators working on individual projects. Network staff will classify sensitive data on a case-by-case, project-by-project, basis. They will work closely with investigators for each project to ensure that potentially sensitive park resources are identified, and that information about these resources is tracked throughout the project.

Network staff are also responsible for identifying all potentially sensitive resources to principal investigator(s) working on each project. The investigators, whether network staff or partners, will develop procedures to flag all potentially sensitive resources in any products that come from the project, including documents, maps, databases, and metadata. When submitting any products or results, investigators should specifically identify all records and other references to potentially sensitive resources. Note that partners should not release any information in a public forum before consulting with network staff to ensure that the information is not classified as sensitive or protected. The NCRN will provide guidance to investigators to help inform them of potentially sensitive resources.

For example, information may be withheld regarding the nature and/or specific locations of the following resources recognized as 'sensitive' by the NPS. According to NPOMA, if the NPS determines that disclosure of information would be harmful, information may be withheld concerning the nature and specific location of:

- Endangered, threatened, rare or commercially valuable National Park System Resources (species and habitats)
- Mineral or paleontological objects
- Objects of cultural patrimony
- Significant caves

Note that information already in the public domain can, in general, be released to the public domain. For example, the media has reported in detail the return of condors to the Grand Canyon. If an individual requests site-specific information about where the condors have been seen, this information, in general, can be released. However, the locations of specific nest sites cannot be released.

9.1.4 Access Restrictions on Sensitive Data

NCRN personnel are responsible for managing access to sensitive data handled by the program. All potentially sensitive park resources will be identified and investigators working on network projects will be informed that:

- All data and associated information must be made available for review by network staff prior to release in any format
- Any information classified as protected should not be released in any format except as approved in advance by the NPS

The network coordinator, NPS project manager, or data manager identifies all potentially sensitive park resources to the principal investigator for each project. Reciprocally, the principal investigators for each project must identify any known references to potentially sensitive park resources.

For each project, the network staff provides a complete list of all references to potentially sensitive park resources in each park to the park superintendent for review. Each superintendent then determines which information should be protected.

When preparing or uploading information into any network database, the network staff ensures that all protected information is properly identified and marked. The network staff will work together to ensure that all references to protected information are removed or obscured in any reports, publications, maps, or other public forum.

The main distribution applications and repositories developed by the I&M Program, (see section 9.2.1) are maintained on both secure and public servers, and all records that are marked ‘sensitive’ during uploading will only become available on the secure servers. Network staff should use the criteria mentioned in section 9.2.2 prior to assigning a sensitivity level to specific data records prior to uploading information to the NPSpecies and NatureBib databases.

Thus, access to data on sensitive park resources can be limited to network staff or research partners. However, limits to how these data are subsequently released must also be clearly defined. It is crucial that the network staff institute quality control and quality assurance measures to ensure that the person doing the uploading of records into the online applications is familiar with the procedures for identifying and entering protected information.

9.2 DATA DISTRIBUTION

One of the most important goals of the I&M Program is to integrate natural resource inventory and monitoring information into NPS planning, management, and decision making.

To accomplish this goal, procedures must be developed to ensure that relevant natural resource data collected by NPS staff, cooperators and researchers are entered, quality-checked, analyzed, documented, cataloged, archived, and made available for management decision-making, research, and education. Such procedures have been documented throughout this plan as well as in the associated SOP's. Providing well-documented data in a timely manner to park managers is especially important to the success of the program. The NCRN Program will make certain that:

- Data are easily discoverable and obtainable
- Data that have not yet been subjected to full quality control will not be released by the network, unless necessary in response to a FOIA request
- Distributed data are accompanied by complete metadata that clearly establishes the data as a product of the NPS I&M Program
- Sensitive data are identified and protected from unauthorized access and inappropriate use

9.2.1 Data Distribution Mechanisms

Prior to finalizing project data and data products information and data are stored locally and only available to network staff. Once products have been finalized and archived (see chapter 10 for details on long term data maintenance) they need to be made available to others. Accomplishing this requires that a range of information systems such as product repositories, clearinghouses, and web applications are used. Each of these systems has a different purpose and function, as shown in figure 9.1.

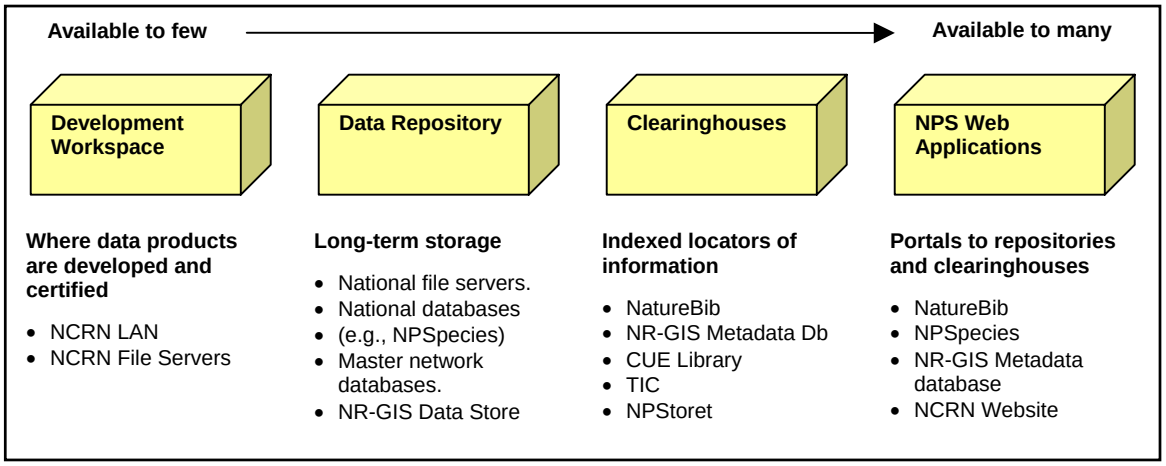


FIGURE 9-1: STORING AND DISSEMINATING PROJECT INFORMATION

The NCRN's main mechanism for distribution of the network's inventory and monitoring data will be the internet. Use of the internet will allow the dissemination of data and information to reach a broad community of users. As part of the NPS I&M Program, web-based applications and repositories have been developed to store a variety of park natural resource information. The network will use the following applications and repositories to distribute data developed by the program:

National Repositories

- NatureBib
- NPSpecies
- Biodiversity Data Store
- NPS Technology Information Center (TIC
- Natural Resource and GIS Metadata and Data Store

Network Repository

- NCRN Website–provides detailed information about the network and it's I&M Program. Data and products will either be available through the site, or users will be directed to where the data are stored (National Capital Region Network website).

For specific information relating to the resources listed above, refer to Chapter 3 (specifically: 3.2 National Information Management Systems) of this plan. Any information not readily available through any of the above-mentioned repositories can be acquired by contacting the NCRN data manager directly whose contact information will be available on the network website. The NCRN also plans to establish an e-mail account dedicated to handling both data requests as well as comments regarding data already disseminated. All network personnel will have access to this account and be able to process any comments or feedback received. The NCRN data manager will be responsible for processing data requests. Comments can also be submitted to the data manager directly.

Because network data will reside in the repositories listed above, this data will automatically be searchable via the integrated metadata and image management system and search gateway called NPS Focus. NPS Focus has been released as an Intranet version only (http://focus.nps.gov/). A public version is projected in the near future.

There are others methods of disseminating information in addition to data clearing houses and web sites. The NCRN will take advantage of its collocation with regional programs especially the regional Resource Learning Center. The NCRN is presently working with the Regional Learning Center Coordinator to develop new ways of disseminating data collected as part of I&M activities. These include Inventory summaries and highlights and park resource highlights and fact sheets. Through coordination with the Resource Learning Center, information can be made available to a wider audience and in the appropriate formats.

9.2.2 Public Access to Network Inventory and Monitoring Data

The network will regularly provide updated information about inventories and monitoring projects, including annual reports and detailed project reports through the network web site. Information on species in the national parks, including all records generated through the I&M Program, will be maintained and assessable through the NPSpecies database. Bibliographic references that refer to National Park System natural resources will be accessible through the NatureBib database. Documents, maps, and data sets containing resource information from all sources, and their associated

metadata, will be assessable through the Biodiversity Data Store and/or NR-GIS Metadata and Data Store. Each of these databases/repositories will be available via both a secure server and a public server, and the public can access all information in these databases except those records marked as 'sensitive.'

9.3 DATA FEEDBACK MECHANISMS

The NCRN website will provide an opportunity for NPS staff, cooperators and the public to provide feedback on data and information gathered as part of the network's I&M Program. The metadata attached to each dataset or report available to the public will include information on where to send comments or questions regarding the data or document. In addition there the network website will also have information on who to contact regarding a specific dataset report. Annual reporting of progress will be presented to the Board of Directors and to the Technical Committee on a yearly basis, and feedback will be expected during and following these presentations.

9.3.1 Data Error Feedback Procedures

The following feedback procedures describe the process that the NCRN will use to receive and verify data errors identified by public and private data users:

- Data users send a notification about an alleged error to a "comments and requests" e-mail address posted on the NCRN website. This e-mail account will be checked daily by various network personnel.
- The NCRN then determines if the data questioned by the notifier are correct or incorrect. If the data are incorrect, network staff makes the appropriate corrections and notifies the original data collectors (cooperator, other agency, park staff, etc.). Either way, network staff informs the notifier that their feedback has been processed and whether or not corrections were necessary.
- The NCRN staff then input the information into a data error log table incorporated in either each of the network monitoring databases or a specific error tracking database developed for the network.
- Once the data are corrected the data repository will be refreshed with the corrected information.

Credits

Developed by:	Adapted from materials developed by Sara Stevens (NCBN)
Other contributions included:	Data Management Plan Working Group

Chapter 10

Data Maintenance, Storage, and Archiving

Effective long-term data maintenance depends on thoughtful and appropriate data documentation. An essential part of any archive is its accompanying explanatory materials (Olson and McCord 1998). This chapter will refer to, and in some cases elaborate on, metadata standards and dataset documentation procedures that are more fully explained in chapter 7 (Data Documentation) of the this plan.

Data, documents, and anything that results from projects and activities that use network data are all crucial pieces of information. To ensure high-quality long-term management and maintenance of this information, the National Capital Region Network (NCRN) will implement procedures to protect information over time. These procedures will permit a broad range of users to easily obtain, share, and properly interpret both active and archived information.

10.1 DIGITAL DATA MAINTENANCE

In general, digital data maintained for lengthy periods of time will be one of two types: short-term data sets, for which data collection and modification have been completed (i.e., inventory projects); and long-term monitoring data sets, for which data acquisition and entry will continue indefinitely.

Technological obsolescence is a significant cause of information loss, and data can quickly become inaccessible to users if they are stored in out-of-date software programs or on outmoded media. Effective maintenance of digital files depends on the proper management of a continuously changing infrastructure of hardware, software, file formats, and storage media. Major changes in hardware can be expected every 1–2 years and in software every 1–5 years (Vogt-O'Connor 2000). As software and hardware evolve, data sets must be consistently migrated to new platforms, or they must be saved in formats that are independent of specific platforms or software (e.g., ASCII delimited files).

10.1.1 Short-term Data Sets

Upon project finalization, project managers are responsible for creating a set of ASCII comma-delimited text files for all short-term data sets created or managed by the NCRN. Each database table will have a corresponding ASCII file. These files will be accompanied by a README.TXT file that explains the contents of each file, file relationships, and field definitions. These ASCII files are in addition to the native version of the dataset (typically in database or spreadsheet format). Creating these text files will help ensure that the data are usable in a wide range of applications or platforms. All finalized files will be stored on the NCRN server's archive section in the appropriate project folder.

In addition to creating ASCII files, NCRN will also update completed and archived data sets that may have been created in older versions of MS-Access, with the goal of having no dataset more than two versions behind the current version used by NCRN. There is a risk of reducing performance in the process of conversion; for example, complex data entry forms or reports may not function properly in an upgraded version. To the extent possible, we will maintain proper functionality of data entry forms and reports. However, the priority will be to ensure basic table and relationship integrity. All previous versions of the data set will be saved.

10.1.2 Long-term Monitoring Data Sets

Long-term monitoring data sets require regular updates and conversion to current database formats. All active or long-term databases will conform to the current NPS and I&M software version standards.

Monitoring projects will also have variable long-term data archiving requirements. Raw data sets that are later manipulated or synthesized may need to be stored in perpetuity. Modifications to protocols will typically require complete data sets to be archived before modifications are implemented. Depending on the monitoring project, it may be necessary to preserve interim data sets (data "milestones") over the long term. The NCRN will save archived data sets or subsets destined for long-term archiving, whenever possible, in their native formats in addition to ASCII text files.

10.1.3 Quality Control of Converted Data

All ASCII files created from databases will undergo quality control (QC) to ensure that the number of records and fields correspond to the source data set and that conversion has not created errors or data loss. A second reviewer (preferably a program scientist) will evaluate the ASCII files and documentation to verify that tables, fields, and relations are fully explained and presented in a way that is useful to secondary users.

Databases that are converted from one version of MS-Access to an upgraded version will require additional QC. This is especially important if the databases are being actively used for data entry or analysis. Forms, queries, reports, and data entry all will be thoroughly tested.

10.1.4 Version Control

All documents and applications produced by the NCRN will have a “revision log” included. Documents such as protocols, guideline documents and standard operating procedures (SOPs) will all include a revision table to document “who” made the change(s) and “when” they were made.

All draft documents will be tracked by the date they were last edited. When the document becomes finalized, the finished version will be distinguished by a version number (e.g., version 1). Network databases will include a table to document changes made to the database.

Previous versions of databases will be saved in their native format and archived in addition to the current version. Documentation of version updates and associated details will be part of the archive metadata document, and revision information and history will also be included in tables within the database files themselves. File names of the archived revisions should clearly indicate the revision number or date.

10.1.5 Spatial Data

Spatial data sets that are essential to the NCRN will be maintained in a format that remains fully-accessible by the current ArcGIS version. ArcGIS has maintained compatibility with previous data formats, and while shape-files have retained all functionality in ArcGIS, Arc Info coverages may require conversion to ArcGIS shapefiles or Geodatabase Feature Classes in the future. At this time there is no practical way to save GIS data in a software or platform-independent format.

Both uncorrected and corrected GPS data (e.g., .SSF and .COR files) will be archived in their native format in addition to the corresponding GIS files that are created.

10.2 STORAGE AND ARCHIVING PROCEDURES – DIGITAL DATA

Digital data will need to be stored in a repository that ensures both security and ready access to the data in perpetuity. Appendix L explains and illustrates the directory structure employed by the NCRN wherein there are two distinct storage directories, *Active* and *Archive* are used to house data and information depending on their degree of completeness or finality.

Active Directory – used to house any file that is still in use or not yet considered complete. This includes datasets that have not yet undergone all of the necessary QA/QC procedures as well as draft documents.

Archive Directory – this directory contains all data sets that are complete (this mostly pertains to short-term projects such as inventories) along with completed reports, contracts and agreements. In the case of long-term monitoring data sets, because they will remain active indefinitely, the datasets will be archived periodically (most likely annually) once the data have undergone QA/QC to preserve interim datasets. For data security purposes, this directory will be formatted as 'Read Only' for all personnel except the Data Manager and the Regional IT server support staff.

10.2.1 Backup Procedures for Digital Data

The risk of data loss can come from a variety of sources, including catastrophic events (e.g., fire, flood), user error, hardware failure, software failure or corruption, and security breaches and vandalism. Performing regular backups of data and arranging for off-site storage of backup sets are the most important safeguards against data loss.

The NCRN I&M Program shares its file server with the National Capital Region Natural Resources and Science staff at the Center for Urban Ecology (CUE). Accordingly, backup procedures must be able to accommodate not only data produced by the I&M Program but all of the data and information stored by the regional staff as well. Due to the number of users utilizing the file server, the quantity of data included in the backup procedure is quite large and constantly increasing. Thus, the NCRN will reevaluate the backup procedures often, at least annually (but most likely semi-annually) to ensure that the present procedures are adequate to accommodate not only the Network data but regional data as well. It is the intent of the NCRN to conduct a full tape backup on a daily basis and maintain those tapes for at least a month's time. This will allow for complete restoration of the server files for up to a month. Refer to the SOP for digital data storage in appendix L where the current NCRN server backup procedure is described.

10.2.2 Testing of Backup Files

All backups should be run with "verify" toggled. "Verify" compares source files against the backup files and detects any discrepancies in file size or other errors. However, to protect against undetected hardware or tape failures, a selective restore should be conducted monthly on each active tape set. Tapes will be replaced every two years to prevent data loss due to deterioration.

10.2.3 Data and Network Security

Because NCRN is located within National Park Services offices, local and wide area networks currently conform to Department of Interior security guidelines.

Only I&M staff and system administrators have permission to edit files stored under the I&M directory, and restrictions have been established on archived data files. As mentioned above, all 'Archive' directories are designated as 'Read Only' for all staff with the exception of the data manager and IT server administrator. Therefore, all changes to files located in the Archive directories

must be routed through the data manager, who is responsible for ensuring that documentation associated with the data set are also updated.

10.3 STORAGE AND ARCHIVING PROCEDURES–DOCUMENTS AND OBJECTS

The guidelines in this section apply to documents such as final reports prepared by staff or contractors, program administrative documents, contracts and agreements, memoranda of agreement, and other documents related to NCRN administration, activities and projects. These guidelines also apply to physical items such as natural history specimens, photographs, or audio tapes. In most instances, these documents and objects are essential companions to the digital data archives described earlier.

Direction for managing these materials (as well as digital materials) is provided in NPS Director's Order 19: Records Management (2001) and its appendix, NPS Records Disposition Schedule (NPS-19 Appendix B, revised 5-2003). NPS-19 states that all records of natural and cultural resources and their management are considered mission-critical records. NPS-19 further states:

> "Mission critical records are permanent records that will eventually become archival records. They should receive the highest priority in records management activities and resources and should receive archival care as soon as practical in the life of the record."

Section N of appendix B, which provides guidelines on natural resource-related records (including, specifically, the results of Inventory and Monitoring Programs), indicates that all natural resource records are considered "permanent," that is, are to be transferred to the National Archives when 30 years old. It also indicates that non-archival copies of natural resource-related materials are" potentially important for the ongoing management of NPS resources" and should not, in any instance, be destroyed.

10.3.1 Documents

All paper documents managed or produced by the NCRN will be housed in one of three locations. Refer to appendix L for specifics on how the NCRN handles print materials.

1. *NCRN central files.*

   These files contain project files, administrative documents and non-record copies of documents that are archived at an off-site facility (see item 2, below). Examples include: meeting minutes, correspondence, memoranda of understanding, contracts and agreements, research permits, interim, original data sheets and selected final reports produced by the program or under its auspices. The NCRN will use acid-free paper and folders for all permanent records in the central files. In addition to maintaining these paper records, NCRN will maintain electronic versions, when possible, on the NCRN server. The central files are maintained by the NCRN staff, at the Center for Urban Ecology in fire proof filing cabinets.

2. *Museum Resources Center (MARS)*

   MARS provides temperature and humidity-controlled facilities, a professional archival staff, and meets all museum standards set by NPS. This repository will be used for original documents and associated materials produced by the network that are a high priority to maintain under archival conditions. Examples include: original inventory reports and accompanying slides and maps; original vegetation mapping reports; NCRN Phase 1, 2 and 3 reports.

Copies of these reports will be maintained in the NCRN central files, and all will have an electronic equivalent (e.g., pdf) for distribution or reproduction.

For all materials submitted to MARS, NCRN will provide essential cataloging information such as the scope of content, project purpose, and range of years, to facilitate ANCS+ record creation and accession. NCRN will also ensure that materials are presented using archival-quality materials (acid-free paper and folders, polypropylene or polyethylene slide pages).

Many NCRN reports and documents encompass data from multiple parks, which has made it difficult to accession archival copies into a specific network park museum. The NCRN and regional staff are currently working together to develop a SOP for accessioning "multi-park" materials at the Regional Museum Resources Center. In these instances MARS will prepare associated ANCS+ records that reference all parks included in a report or document, and will prepare finding aids to help potential users locate the materials.

3. *Network Parks' Central Files and/or Museums*

The NCRN will provide high-quality copies of park-related documents resulting from Network projects, along with electronic versions, to park resource management staff. Parks may choose to accession these materials into their museums, incorporate them into their central files, or house them in their resource management library, as they deem appropriate. The NCRN will not manage documents at the Park level.

10.3.2 Specimens

The Network will provide specimens collected under the auspices of the NCRN to the Network Park in which they were collected for inclusion, or to a repository approved by a Park (where the specimens are considered on loan). The NCRN recommends that all voucher specimens be stored at MARS as this provides a central location where an entire collection can be reviewed. The NCRN will provide Park curators with associated data required for cataloging each specimen. These data will be in comma-delimited format (.CSV) format for automated uploading into ANCS+.

10.3.3 Photographs

The NCRN maintains photographs in various formats but the majority by far are maintained in a digital format such as JPEG or TIFF. These formats are appropriate for most purposes; however, when dealing with photographic vouchers the NCRN has chosen to take a conservative approach. Original images are a high priority for placing in archival storage conditions. Accordingly, the NCRN encourages staff and contractors to provide photos as 35 mm slides (preferably Kodachrome or Ektachrome), which have a proven long-term stability (Wilhelm and Brower 1993). If contractors cannot provide slides, NCRN requests 4 x 6 color prints. For more information regarding photographic vouchers refer to the photo voucher SOP in appendix D. Also, appendix C contains the SOP describing the photographic tracking and cataloging system used by the NCRN.

Role of Curators in Storage and Archiving Procedures—Curators for parks within NCRN are an ongoing source of expertise, advice, and guidance on archiving and curatorial issues, and they have a role in almost every project undertaken by the Network. Project managers should involve Park curators when projects are in the planning stage, to ensure that all aspects of specimen preservation or document archiving will be considered, and that any associated expenses are included in project budgets.

Credits

Developed by:	Adapted from materials developed by Margaret Beer (NCPN)
Other contributions included:	Data Management Plan Working Group

Literature Cited

Edwards, Don. 2000. Data quality assurance. Pp. 70-91 *in* Michener, William K., and James W. Brunt (eds), 2000, Ecological Data: Design, Management and Processing, Methods in Ecology Series, Malden, MA: Blackwell Science Ltd.

Koenen, Marcus, Ed. 2005. National Capital Region Network of the National Park Service – Monitoring Plan.

National Park System Advisory Board. 2001. Rethinking the National Parks for the 21st Century: A Report of the National Park System Advisory Board. July.

National Park Service. 2001. Director's Order #19: Records Management.

National Park Service. 2002. Director's Order #11B: Ensuring quality of information disseminated by the National Park Service.

Olson, R.J. and R.A. McCord. 1998. Data Archival. In: Michener, W.K., J.H. Porter, and S.G. Stafford. Data and information management in the ecological sciences: a resource guide. LTER Network Office, University of New Mexico, Albuquerque, NM. pp 53-57.

Palmer, Craig J. 2003. Approaches to quality assurance and information management for regional ecological monitoring programs. Pp. 211-225 *in* Busch, David E., and Joel C. Trexler (eds), 2003, Monitoring Ecosystems: Interdisciplinary Approaches for Evaluating Ecoregional Initiatives. Washington, DC: Island Press.

Palmer, Craig J., and Eric. B. Landis. 2002 Draft. Lake Mead National Recreational Area Resource Management Division: Quality System Management Plan for Environmental Data Collection Projects.

U.S. Environmental Protection Agency. 2003. Guidance for geospatial data quality assurance project plans. EPA QA/G-5G. EPA/240/R-03/003. Washington, DC: Office of Environmental Information.

USDI National Park Service. 2003. NPS records disposition schedule. NPS-19, Appendix B – Revised, 5-03. 61 p.

Vogt-O'Connor, D. 2000. Planning digital projects for preservation and access. National Park Service Conserve O Gram 14:4. NPS Museum Management Program, Washington, DC. 4 p.

Wilhelm, H. and C. Brower. 1993. The permanence and care of color photographs: traditional and digital color prints, color negatives, slides, and motion pictures. Preservation Publishing Company, Grinnel, IA. 744 p.

Glossary

Cheat sheets — simple, sequential procedures for conducting specific data management practices (e.g., cheat sheet on linking digital images to Access databases); usually included as supplementary sections in overall data management plans, DMOPs or SOPs.

COTR or COR — contracts and agreements representative. Once trained as a COR, personnel can manage cooperative agreements and contacts.

Data management operating procedures (DMOPs) — detailed step-by-step instructions for carrying out data management practices (the 'How?' of data management planning); can be frequently updated; referenced by overall data management plan and monitoring protocol SOPs (e.g., naming conventions, photo management practices, data verification and validation techniques, deliverable specifications).

General guidance documents — data management plan sections, stand-alone guidelines and tip sheets (provide background and inform why things are to be done a certain way).

Guidelines — general descriptions of data management practices and procedures; guidelines can reference DMOPs, monitoring protocol SOPs and standards.

Inventory study/project — a field sampling and/or data mining effort designed to gather data for a natural resource inventory; may or may not be repeated.

Long-term monitoring study/project — a monitoring protocol designed to assess the status of a natural resource vital sign over time.

Policy — Rationale/guiding principles for data management practices and procedures (the 'Why?' of data management planning).

Project — a general term for a formal data gathering effort; 'formal' implies an organized, documented approach that includes, at minimum, a study plan/proposal detailing the protocol(s) followed; projects can be inventory studies, pilot studies, short-term/park-specific studies or longer-term monitoring projects

Monitoring protocol — the formal documents and sampling processes describing how a vital sign will be monitored; composed of a narrative section, standard operating procedures and supplementary information (databases, reports, tools, hardcopy materials); a monitoring protocol is a component of a vital sign monitoring study plan.

Protocol — a set of sampling procedures and/or experimental processes followed in a project.

Short-term monitoring study/project — pilot study designed to guide, test, and evaluate monitoring protocol development.

Specific guidance documents — DMOPs, standards, and cheat sheets, either as standalone documents or appendices to data management plan (inform the what and how of data management practices).

Standard operating procedures (SOPs) — detailed step-by-step instructions for carrying out sampling procedures in monitoring Protocols (the 'How?' of data management planning); dynamic in nature and frequently updated; can reference overall data management plan, standards and DMOPs.

Standard — a data management practice or policy implemented service-wide or network-wide (e.g., recommended naming conventions, FGDC metadata format); referenced by overall data management plan, monitoring SOPs and DMOPs.

Tips — small 'how-to' sections within larger data management documents (overall data management plan, DMOPs); may evolve into full cheat sheets and/or DMOPs.

Vital signs — physical, chemical or biological indicators that help assess the status of a park's natural resources.

Appendix A

Roles and Responsibilities for Personnel Involved in I&M Network Activity

Position Title	Information Management Responsibilities
Project Crew Member	Collect and record field data. Verify, validate and document all datasets.
Project Crew Leader	Supervise field crews and ensure that crews are adhering to all protocol related and Network standard operating procedures (SOPs).
Data/GIS Specialist	Perform a variety of Network data management and GIS tasks including data entry, data conversion, and data documentation. Work on overall data quality and stewardship
IT Specialist	Provide Network-wide IT support and helps support the Network's data management system.
Project Leader	Oversee the aspects and operations of one or more Network projects including data management. Act as a project liaison between the project crew, park staff and Network staff including the Network data manager.
GIS Manager	Manage regional GIS data. Coordinate and integrate local GIS and resource information management with Network, regional, and national standards.
Network Data Manager	Provide comprehensive data management guidance and support for all Network activities including planning, training, database development, GIS support, coordination and integration. Acts as the Regional Point of Contact (POC) for NPS databases.
Network Ecologist	Assist in the development of project sampling designs and ensures the proper collection and management of data by integrating natural resource science in Network activities. Responsibilities include data analysis and synthesis and reporting. Note: The Network Ecologist may also function as a Project Leader.
Network Coordinator	Ensure that programmatic data and information management requirements are met as park of overall network business. Note: The Network Coordinator may also function as a Project Leader.
I&M Data Manager (National)	Provide service-wide database design, support and services including receiving, storing and archiving data in service-wide databases.
Regional Coordinator	Assist with coordinating data management activities between regions as well as with non-NPS cooperators.
End Users	These information consumers' include park managers and superintendents, researchers and the general public. End users are responsible for the appropriate use and application of the data and data products and should provide feedback relating to data and data products.

Appendix B
Digital Data Storage
Standard Operating Procedure

Digital Data Storage
Standard Operating Procedures

Revision Log

Revision	Description of Change	Author	Effective Date
1	Original	G. Sanders	10/01/2005

Purpose

The purpose of this standard operating procedure (SOP) is to present the guidelines and overall process for tracking and storing digital data in the National Capital Region Network (NCRN).

Scope and Applicability

The standards described in this procedure pertain specifically to all digital data related to natural resources inventory and monitoring. Digital data refers to all data stored and manipulated on various digital media including computer files, cds, zip disks, DVDs, or other media. The data on the media may include documents such as annual reports, final reports, meeting notes, digital data files, digital photos, digital GIS data, spreadsheets, or scientific literature developed relating to the NCRN parks. Any files that fall under the guidelines of Records Management Handbook NPS-19, Appendix B, must be printed out and stored appropriately following Print Media Storage SOP.

Reference Documents

Records Management Handbook NPS-19, Appendix B. Document available online at:
http://data2.itc.nps.gov/npspolicy/DOrders.cfm

Dataset Catalog. Document available online at:
http://science.nature.nps.gov/im/apps/datacat/index.htm

SOP Print Media Storage Specifications.

Procedures and General Requirements

All electronic files connected with the NCRN Inventory and Monitoring Program are stored on the CUE File Server. I&M project data should always be stored under the I&M directory on the Shared drive. The specific storage location within the I&M directory depends on the project the data relates to as well as the status of the data (i.e., raw vs. certified). Only personnel employed by the NCRN I&M Program will have "write" privileges under this directory. Employees should not maintain any original I&M project data on either their local "personal" server directories or on their local hard drives.

Digital Data Storage
Standard Operating Procedures

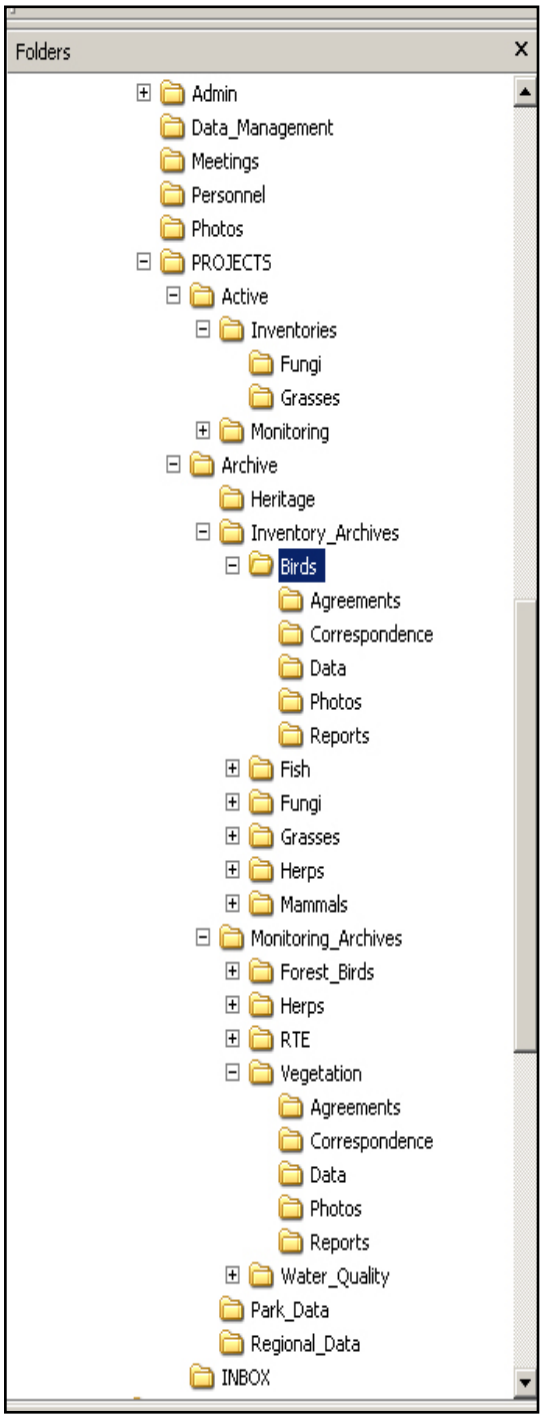


Figure 1. NCRN sample directory structure.

Directory Structure

Program Related Material

Digital files that relate to the program in general are maintained separately from project specific materials. Examples of program related materials include:

- Program reports (AARWP)
- Park information
- Program supplies and purchasing information
- Meeting minutes
- General photographs

Project Related Material

Project specific materials are stored under the Projects directory. This directory contains project tracking software as well as the following sub-folders:

ACTIVE DIRECTORY – Contains folders for those projects that are active. Files that are stored in the project specific folders are all files that are updated or changed on a regular basis. This directory primarily contains:

- Working project databases

ARCHIVE DIRECTORY – Contains files that are considered to be complete (i.e., files that will not undergo regular edits or changes. As an information security measure, this directory is formatted as "Read Only". Only the NCRN data manager will have write access for the Archive Directories. This directory contains:

- Scanned original data sheets
- Protocols and SOPs
- Master project databases

- Completed reports
- Digital copies of contracts or agreements

Field Data

All digital data relating to NCRN I&M projects are stored in a project specific folder either under the Active Directory or the Archive Directory depending on the level of completeness (i.e., whether the data has undergone quality assurance / quality control [QA/QC]). The data should be stored in a project specific database based on the National Resources Database Template (NRDT). The database should be designed specifically to meet the needs and requirements of the project.

Master project databases are stored in the project's archive folder while the working database resides in the Active Directory. Once the data has been certified, it can be transferred to the master database in the archive folder where it will be "read only."

The periodic transfer of certified data usually only applies to long-term projects such as vital signs monitoring projects that will continue indefinitely. Data from short-term projects such as biological inventories are usually certified all at once and the entire dataset archived.

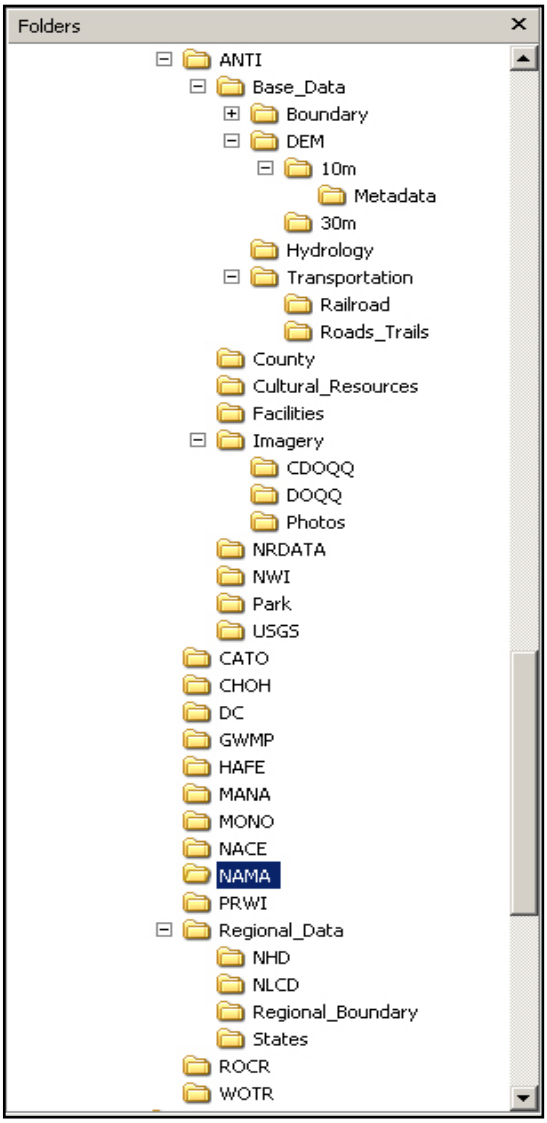


Figure 2. GIS directory structure.

Note: once data has been certified, all data relating to species occurrence should be entered into NPSpecies. In addition, all data products must be forwarded to the proper national repositories for dissemination to park resource managers as well as the public.

GIS Data

All NCRN GIS data is also stored on the shared drive on the CUE file server. Park and regional GIS data are used by regional natural resources staff at CUE in addition to the

I&M staff. As such, the GIS directory is not nested under the I&M Program directory (figure 2).

The files are organized by park and all coverages, shape files, grids or images located in these directories must be considered complete and accompanied by metadata. All folders containing data files have a 'Metadata' sub-folder where metadata files are stored for files in that directory. Given the metadata capabilities of ArcCatalog, the need for a separate metadata folder has diminished. We will maintain these folders until the metadata for all GIS layers has been updated through ArcCatalog.

Any GIS file that is undergoing edits should be stored in another location and copied to the public GIS directory only when editing and documentation have been completed. Files should also be maintained in the park specific projection (appendix A). Only the NCRN I&M Data Manager, regional GIS specialists and the NCR server manager have "write" privileges on this directory. All others accessing have "read only" privileges. For more information on Network and Regional GIS specifications refer to the GIS Specifications SOP.

Photos

All digital photos relating to I&M projects will be stored in a sub-directory under that project's directory. They will remain in the project's active directory until they have been added to the network's photo management system and received metadata documentation at which time they can then be archived.

Park or regional related photographs not necessarily associated with a specific I&M project are stored under the Center for Urban Ecology's photo directory on the Shared directory. These photos must also have complete documentation prior to being stored in this directory.

Please refer to the NCRN Photo Management SOP (*in development*) for more complete details regarding the network's photo management system.

Backup Procedures

The NCRN file Server not only supports the Network Inventory and Monitoring Program but also the entire regional natural resources and science staff located at the Center for Urban Ecology. Because of this, the amount of information requiring regular backups is great.

The NCRN completes a full backup of all server drives on a daily basis onto SuperDLT tapes. This includes the shared "public" drive as well all local "private" drives. Daily backups are maintained for a period of one month. A complete backup of all archive and GIS directories on the "Shared" drive is performed at the end of each month and are maintained for a period of six months. Any information maintained on someone's local hard drive is the responsibility of that individual. All backup tapes are maintained offsite by the I&M data manager.

Responsibilities

Project Leader: Be aware of all reports and documents that are being generated and file them properly in the project's 'Active' directory. When files need to be archived project leaders should notify the data manager and place the file to be archived in the 'Inbox' where the data manager can access it.

Project Manager or Data Manager: Determine which records are permanent and which are temporary. Ensures that all pertinent data are entered into NPSpecies, NatureBib and Dataset Catalog.

Data Manager: Archive all files once project leaders have completed all QA/QC and documentation steps. The data manager is also responsible for implementing the server backup procedures and maintaining all of the backup tapes.

Appendix C
Data Deliverables

General Guidance Document: Project Deliverables

Revision Log

Revision	Description of Change	Author	Effective Date
1	Original	G. Sanders	10/01/2005

Purpose

One of the primary goals of the National Park Service (NPS) Inventory and Monitoring (I&M) Program is to make natural resource information more available and useful to park managers, scientists, and educators for planning, management, research and education. The specifications presented here were developed to provide NPS staff and cooperators with a standard set of guidelines to use for the submission of all products produced during NPS sponsored natural resource studies.

This document is includes a suite of standard operating procedure documents and is intended to provide cooperators, contractors and staff with an understanding of the standards the NCRN requires for project deliverables. Numerous standard operating procedures are included in this document:

a. Project Database Standards SOP
b. GIS SOP
c. Metadata SOP
d. Reporting SOP (used for both annual and final reports)
e. Voucher SOP (Draft)
f. Photographs SOP

Scope and Applicability

The National Park Service and the Inventory and Monitoring Program have adopted certain word processing, database, and GIS software as standards to promote compatibility and sharing of data among parks. In addition to software standards, the NPS I&M program has adopted a set of products with standardized formats to be submitted along with all NPS, I&M sponsored studies. This document provides details on these required software and product standards.

This document applies to all projects conducted by the NCRN I&M Program and all those conducting work with or for the NCRN I&M Program should be familiar with this document and the guidance provided herein.

Definitions and Acronyms

NCRN – National Capital Region Network

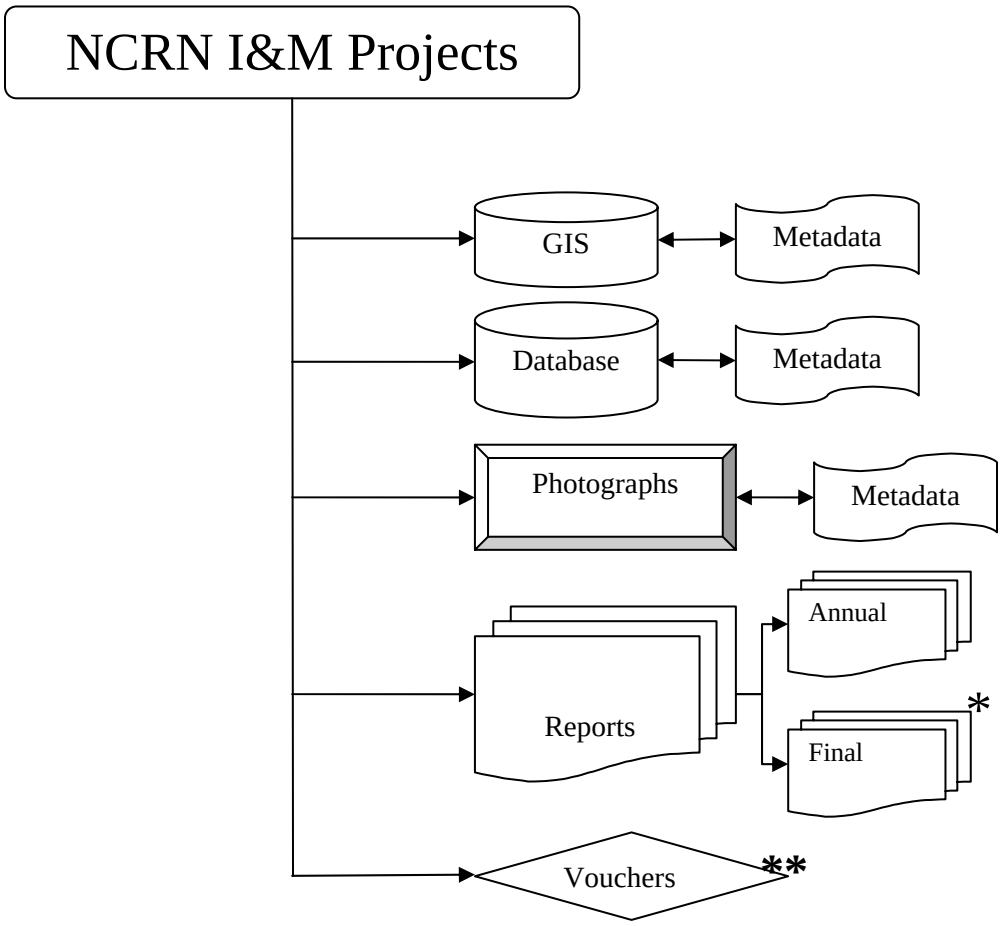
FGDC – Federal Geographic Data Committee

Deliverable – General term that describes products that are required during and at the completion of a project.

GIS – Geographic Information System

Metadata – Data about data. Must accompany all spatial and non-spatial datasets.

Flow Chart



* Only short-term projects (e.g. inventories or research projects) that have finite life span will normally require final reports.

General Requirements

Software Standards

Word Processing: Microsoft Word. All reports and other documents must be delivered in electronic format in the latest available version of MS Word unless otherwise specified in advance.

Databases: Microsoft Access. All database products must be delivered in electronic format in the latest available version of MS Access unless otherwise specified in advance. All database products will be compatible with the most recent version (v. 3) of the Natural Resource Database Template (described below under Database Specifications).

GIS Products: ArcInfo and/or ArcView. All GIS products must be compatible with ESRI's ArcINFO and/or ArcGIS 8.3 or greater and be accompanied by FGDC compliant metadata. Additional information regarding GIS and metadata specifications are described in the GIS Specifications and Metadata Specifications SOP documents.

Metadata: All data sets, both spatial and non-spatial, must be accompanied by FGDC compliant metadata. Please refer to the Metadata SOP and http://www.fgdc.gov for additional information.

Required Products and Standard Procedures

a. Reports
b. GIS Layers
c. Databases and Data Files
d. Digital Photos
e. Vouchers

Reports
Standard Operating Procedure

Revision Log

Revision	Description of Change	Author	Effective Date
1	Original draft	G. Sanders	10/01/2005
1.1	Added documentation	G. Sanders	12/06/2005

Purpose

The purpose of this SOP is to identify the different reports required by cooperators and outline the content standards.

Annual Reports: Principal investigators must submit a draft annual report to the designated NPS project manager. All appropriate comments from draft report reviews should be addressed and incorporated during the preparation of the final annual report – and must be approved by the NPS project manager before accepting the final report.

Final Reports: At the completion of a project, the investigator must submit a draft final report to the designated NPS project manager who will review the manuscript and may also seek additional comments from NPS regional staff, park personnel, and scientists well versed in the subject matter to ensure the technical quality and accuracy of information. All appropriate comments from draft final report reviews should be addressed and incorporated during the preparation of the final report – and must be approved by the NPS project manager before accepting the final report.

Procedures and General Requirements

All reports must include the following:

Hard Copy: Submit two (2) hard copies of the report to the Network Project Leader.

Electronic copy: Submit electronic copy in MS Word (version 2002 or higher) on cd or dvd,to the NCRN Inventory and Monitoring Coordinator or Project Leader.

Submit all other data as outlined in SOP Product Specification.

Format for reports should follow:

- *Title Page* – Should include: title, parks, study dates, report date, authors, and other credits (contract numbers etc.) as appropriate. Also include complete current mailing address (including zip + four), email, phone and fax number of the person to whom correspondence should be sent.

  Final Reports Only

  The title page should also include a 'TIC' Number which is an NPS assigned cataloging number. Prior to the submission of the final report, inform the NPS Project Leader and they will obtain the number and provide it to you.

Reports
Standard Operating Procedure

- *Report Checklist* - Use the checklist (Table 1) to make sure every section has been completed. Submit the completed checklist after the Title Page of the report.

- *Abstracts* - the third page should be an abstract that does not exceed 5% of the length of the paper. The abstract should explain the purpose of the study, describe the principal findings, and state the main conclusions. Many readers will rely heavily on the abstract, so it should be as descriptive as possible. Avoid uninformative sentences such as "The significance of these results is discussed." Generally, quantitative results (e.g., sample sizes, P-values) are unnecessary in abstracts and can be summarized in narrative fashion.

- *Keywords* - Below the abstract, provide 5-6 keywords that describe the subject of the paper; these need not duplicate words in the title.

- *Introduction* – Should include the hypotheses and purpose of the investigation, research objectives, conditions under which the study was conducted, the general plan of treatment of the subject, and summary of previous work (literature review) that relates to the project.

- *Study Area* – describe the area(s) and habitat(s) included in the study. A map can be used in this section to aid in defining the study area.

- *Sample Design* – describe the overall sampling design and explain how the sample locations were chosen.

- *Sample Locations* – include locations in the sampling design as well as any incidental locations (presented as GIS maps).

- *Methods* – Present a detailed explanation of the methods, materials and analytical techniques that were used in the field, laboratory, and office during the study. The methods should be documented so that the investigation could be exactly repeated if necessary. Be sure to include a description of data analysis methods and what statistical tests were used, if any. Describe the processes used for determining whether the data met data quality objectives and, if not, what corrective actions were taken. Detailed QA/QC procedures for the data collection, verification and validation should be placed in an appendix.

- *Results* – In a logical sequence (e.g. by park), present the findings of the study that either support or provide contrary evidence against the hypotheses or that answer questions posed in the introduction. Basic, descriptive statistics (sample size, mean, max/min, trends etc.) are appropriate when clearly presented. Avoid highly technical discussions of complex statistical testing; instead refer readers to a separate appendix for additional information. Consider presenting results separately for each park.

- *Discussion* – The “Discussion” and the “Conclusion” sections are the most important sections of the report. Present a clear interpretation of the data that addresses the hypotheses, objectives, or purpose for which the study was conducted. Be sure to include how this research is applicable to each park where the study occurred and to other studies that have been conducted in that specific area of research. Other findings may be reported that would be of general interest to the scientific community.

- *Conclusions* – Provide a specific and detailed summation of the conclusions of the project.

In addition the conclusion section should have three subsections that are of significant interest to the parks.

- Public Interest Highlights: This section has significant potential to create interest and support for this and other resource management projects. Provide a bulleted list of highlights from this project. Resource managers may use these as talking points with other park staff including management. Interpretation staff may use these to highlight aspects of interest to the general public.

- Management Recommendations: If management recommendations are warranted, highlight them here. If the study was initiated due to specific park management needs, management implications should be emphasized and discussed.

- Research Recommendations: Summarize additional research needs. This section may serve as a way to develop priority research projects for the parks and to support future funding.

- *Literature Cited* - List literature citations alphabetically by the first author's last name in a literature cited section following the acknowledgments. Literature cited entries should use a consistent format. Carefully double-check citations against citations in the text. See Attachment one for a style description and examples.

- *Appendices* - Provide additional data as necessary, including a hard copy printout of the data contained in the electronic spreadsheet or database.

- *Tables and Figures* - Tables and figures are to have complete but brief headings, and be numbered consecutively within the text.

- *Additional Manuscript Information*

Print must be in upper- and lowercase letters and of good quality, and paragraphs must be clearly delineated. Large blank spaces (such as between sections) are to be avoided. The English and scientific names of a species should be given the first time it is mentioned. The scientific name should be in *italic*. Indicate which taxonomic classification is being followed (eg AOU for bird species).

Metric measurements must be given unless English measurements are clearly more appropriate, in which case metric equivalents must be given in parentheses.

Define all symbols, abbreviations, and acronyms, but minimize their use.

Statistical terms and other measures are to conform to *Scientific Style and Format: the CBE Manual for Authors, Editors, and Publishers*, sixth edition.

Spelling should follow the *Webster's Third International Dictionary*, and other style points should follow *The Chicago Manual of Style*, fifteenth edition.

Pages, including tables and figures, should be numbered.

Reports
Standard Operating Procedure

Manuscripts must be in English; U.S. rather than British spelling should be used. We encourage authors whose first language is not English to have a native English speaker edit the paper before submission.

Table 1. Annual and Final Report Checklist.

	Report Checklist
	Cover page lists all authors and addresses.
	Document has been reviewed for content, spelling and grammar.
	Abstract and keywords are provided on page three of final report.
	Body of document includes introduction, description of study area, methods, results, and discussion.
	Conclusion Section includes Public Interest Highlights.
	Conclusion Section includes Management Recommendations.
	Conclusion Section includes Research Recommendations.
	Tables, figures, maps, and illustrations have descriptive titles and are sequentially numbered.
	Document is supported by scientific literature.
	Literature Cited has been checked against the text.
	Submission Checklist
	Two hard copies are being submitted.
	Electronic copy is in MS Word Format (2002 or later) and is submitted on cd, zip, dvd, or sent by email as an attachment.
	Data is being submitted electronically in MS Excel or MS Access format.
	All raw data sheets have been submitted.
	Metadata is provided for all electronic databases (Data fields are defined and database structure is discussed).
	Metadata is provided for any GIS data (following FGDC standards)
	Metadata is provided for raw data sheets (may include field sheet instructions or other documentation that defines fields).
	Photographs have been submitted along with permission for their use in brochures, websites, or other published materials.
	Final Check
	Checklist has been reviewed and all items checked off before final submission. Indicate YES or NO.

GIS Data
Standard Operating Procedure

Revision Log

Revision	Description of Change	Author	Effective Date
1	Original draft	G. Sanders	10/01/2005
1.1	Added additional documentation	G. Sanders	12/06/2005

Purpose

This SOP identifies the standards (data formats, coordinate systems, spatial scale and resolution) for all spatial data collected and managed by the National Capital Region I&M Network

Scope

The standards outlined in this SOP are applicable to all geographic datasets created or managed by Network staff or cooperators. This document does not provide guidance relating to the collection of GPS data but is applicable to how such data is formatted after collection. For detailed information regarding the collection of GPS data, refer to the NCRN SOP for GPS data collection.

Procedures and General Requirements

Data Formats:

- All vector data will be supplied as an ESRI Shapefile and ArcINFO coverage or an ArcINFO interchange file (*.e00) that is compatible with the current version of ArcGIS. All coverages should be created as double precision data sets. If the data was originally created in single precision, it should be converted to double precision before submitting the final product. Both coverages and shapefiles will also have defined projections (in the case of shapefiles, an accompanying *.prj file).

- All raster data will be supplied as an ArcINFO GRID or ArcINFO interchange file (*.e00) that is compatible with the current version of ArcGIS.

  Note: Be aware that when ArcINFO GRID files are converted to interchange files (*.e00) the interchange file size is often much larger than the original GRID. Therefore, this may not be the most appropriate method of transferring this data.

- Digital imagery (e.g. aerial photos) is to be supplied as tagged image file format (.TIFF) with the proper header file for geo-referencing purposes. Depending on the intended use of the imagery, it may be preferable to receive and store the imagery in an uncompressed format therefore, consult the project leader or data manager prior to storing imagery in a compressed format such as MrSID. All digital imagery will be properly orthorectified unless it is decided ahead of time that two dimensional georeferencing is appropriate. All imagery must be accompanied by appropriate header or world files.

- Data file names should not contain spaces. If a space is desired in a field name, use and underscore ("_"). In addition, field names should not exceed 10 characters due to limitations in ArcINFO and dbase.

- All GIS layers (rasters, vectors and imagery) should all be submitted with complete and FGDC compliant metadata. Please refer to the Metadata SOP found in this document for more details on.

Coordinate Systems:

- All spatial data collected and/or submitted for NPS programs will be geo-referenced and provided in the standard projection(s), noted below. The steps used to get the data in the proper projection must be documented in the accompanying FGDC compliant metadata. Prior approval from the project or data manager must be received to deviate from these specifications.

Table 2. Standard projections used by parks in the National Capital Region.

Park	Projection	Datum	Spheroid	FalseEasting	False Northing	Units
All	Geographic	NAD83	GRS 1980			Decimal Degrees
ANTI	UTM, Zone 18	NAD83	GRS 1980	500,000	0	Meters
CATO	UTM, Zone 18	NAD27	CLARK 1866	500,000	0	Meters
CHOH	State Plane MD	NAD83	GRS 1980	1,312,333.33	0	Feet
GWMP	State Plane VA	NAD83	GRS 1980	11,482,916.67	0	Feet
HAFE	State Plane WV	NAD27	CLARK 1866	500,000	0	Feet
NACC	State Plane MD	NAD83	GRS 1980	2,000,000	0	Meters
NACE	State Plane MD	NAD83	GRS 1980	400,000	0	Meters
MANA	UTM, Zone 18	NAD83	GRS 1980	500,000	0	Meters
MONO	UTM, Zone 18	NAD83	GRS 1980	500,000	0	Meters
PRWI	UTM, Zone 18	NAD27	CLARK 1866	500,000	0	Meters
ROCR	State Plane MD	NAD83	GRS 1980	400,000	0	Meters
WOTR	UTM, Zone 18	NAD83	GRS 1980	500,000	0	Meters

Spatial Scale and Resolution:

- Specific scale and spatial resolution requirements for image data will be specified in the contract or cooperative agreement for each project. For example, vegetation classification projects under the NPS/USGS vegetation classification and mapping program will use 1:12,000 color infrared aerial photographs (or better) with 60% overlap (endlap) and 30% sidelap.

- In general, new data should be compiled with an accuracy level better than US National Map Accuracy Standards (NMAS) for a 1:24,000 scale map; unless other requirements exist (which would then be included on a protocol-specific basis). However, all spatial data collected will be analyzed for their spatial accuracy and will meet or exceed NMAS for the appropriate scale (for more information, please see http://mapping.usgs.gov/standards).

    - Note: Adhering to these standards is very important especially when one considers that parks in the NCRN are not that large. Some parks contain very small disjunct parcels while others are long and narrow parks. If strict mapping accuracy standards are not followed, it is likely that mapped features may not fall inside the boundary of the parks when plotted in a GIS.

- For maps on publication scales larger than 1:20,000, not more than 10 percent of the points tested shall be in error by more than 1/30 inch, measured on the publication scale; for maps on publication scales of 1:20,000 or smaller, 1/50 inch. These limits of accuracy shall apply to positions of well-defined points only. Well-defined points are those that are easily visible or recoverable on the ground: monuments or markers, such as benchmarks and property boundary monuments; intersections of roads and railroads; and corners of large buildings or structures (or center points of small buildings). In general, what is well defined will also be determined by what is plot-able on the scale of the map within 1/100 inch.

- The following table provides the allowable horizontal error for some common scales:

Table 3. Allowable error based on USGS National Mapping Accuracy Standards.

Scale	Allowable Error
1:40,000	33.8 meters (111 feet)
1:31,680	16.1 meters (53 feet)
1:24,000	12.2 meters (40 feet)
1:20,000	10.1 meters (33 feet)
1:12,000	6.1 meters (20 feet)
1:9,600	4.9 meters (16 feet)
1:4,800	2.4 meters (8 feet)
1:2,400	1.2 meters (4 feet)
1:1,200	0.6 meters (2 feet)

- Any calculations done with location data should be done at double precision with the results rounded or truncated to the appropriate propagated error limits. All calculations and processing completed on the spatial data shall be reported in the FGDC compliant metadata that accompanies the GIS layer.

Database Specifications
Standard Operating Procedure

Revision Log

Revision	Description of Change	Author	Effective Date
1	Original draft	G. Sanders	10/01/2005

Purpose

To familiarize NPS staff and cooperators of the database standards used byt the NCRN I&M Program. The I&M Program developed a standard database template referred to as the Natural Resources Database Template (NRDT) to help standardize the overall design and functioning of databases used for I&M Program projects.

Scope and Applicability

The standards described in this document are relevant to all NCRN databases. While the specific fields included in the database may differ on a project by project basis, the general structure, organization of the database should be the same as the information presented here. Therefore, a separate data dictionary will be developed for each database to define all data elements; the data dictionary should be used in tandem with this document as a starting point and users guide to I&M Program databases.

Reference Documents

- Balter, Alison. 2002. Alison Balter's Mastering Access 2002 Desktop Development. Sams Publishing, Indianapolis, Indiana. 1348 pages.
- NRDT Version 3 Documentation

Procedures and General Requirements

- All databases from the NCRN I&M Program are based upon the model of the Natural Resource Database Template (NRDT, version 3), available online at: http://science.nature.nps.gov/im/apps/template/index.htm. The NRDT is a relational database in MS Access that may be used as a stand-alone database or linked to ArcView GIS software using other tools from the I&M Program such as ArcView to Access (available online at: http://www.nps.gov/akso/gis/).

Database Structure

Project databases are composed of a front-end forms interface (project_db.mdb) that is linked to one or more back-end tables that contain common tables (Common_db.mdb) used by multiple projects as well as the database where the actual project data is stored (project_be_db) (Figure 2). This format is not required but structuring the database in this manner assists in data organization and management. Incorporating tables that are common to many projects in one database eliminates the incorporate these common tables directly into each project database.

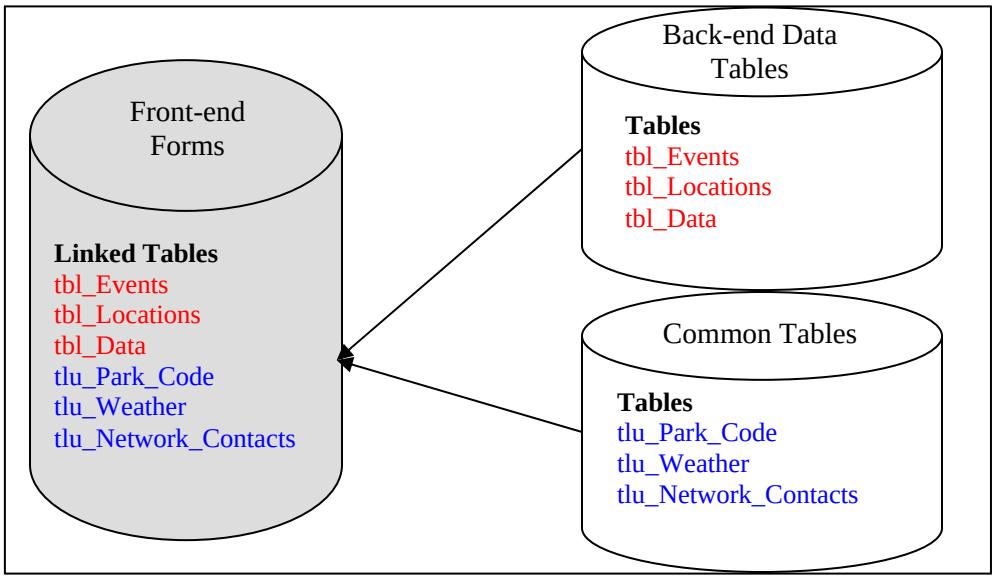


Figure 2. Diagram of NCRN project database structure.

Front-end

Forms database – This is the only part of the combined database structure that the users will work with. Data is entered into the forms and stored in the relevant tables in the back-end databases. The following are some common forms and sub-forms used during data entry.

- frm_Locations – Data entry form for locations information
    - sfrm_Sites – subform under the locations form where site information can be entered.
- frm_Events
    - sfrm_Event_Details – subform where additional event information such as weather information can be entered
    - sfrm_Meta_Events – used to enter information on when data was entered as well as who entered it and when QA/QC was performed and by whom.
- frm_Project_Data – field data
- frm_Contacts – form to enter new project or network contacts.

Two tables, tbl_Locations and tbl_Events, are the two core tables storing the sampling event and locations information for the project. These core tables are accompanied by project related tables that contain the project field data, voucher information and taxonomic look-up tables.

Note: Database users should not try and alter or change the data entry forms or the back-end structure in any way. If problems are encountered during data entry, please contact the NCRN data manager for assistance.

Data Entry

Data entry is initiated through the database switchboard which will open upon once the database application has finished opening. From this screen the user can enter locations data or field data. Note, if field data is selected the user can still link to the locations for from the data forms if new locations need to be entered. From the project data entry form, the user can enter location information (where), event information (when and what were the field conditions) and field data (what was collected).

The forms are designed to make data entry efficient through the use of pick lists and auto-populated fields.The data entry forms are broken into three tabs.

- Event Details Tab: The first tab contains fields for entering details about the sampling event including weather conditions, observers and field notes. This tab also contains a subform used to track when data was entered, checked and verified.

- Data Tab : The purpose of the data tab is self explanatory. All data collected in the field is entered here. Species look-up tables are pre-populated based on the current species lists from NCRN parks and the taxonomy is based upon ITIS (Integrated Taxonomic Information System, from the USDA). Species can be added to the look-up table and this is aided with a direct through a hyperlink to the USDA ITIS website. All species records in the look-up table are accompanied by the ITIS taxonomic serial number (TSN). The user can also indicate whether specimens were collected and if those specimens were prepared as vouchers.

- Vouchers Tab: The final tab is where voucher information is entered. Certain fields are prepopulated such as the taxonomic information. Ideally a set catalog numbers will be obtained prior to starting the project and they can also be automatically assigned.

Data Verification

Once data entry is complete for a set a of field data, the newly entered data must undergo quality control and assurance measures to make sure the data was entered correctly. Users can document when they have completed these steps through the fields on the data entry form.

Data Maintenance and Reports

The Reports screen allows the user to: (1) Generate Current Species Lists, (2) Preview and Print Herbarium Labels, and (3) Generate ANCS+ and Upload Spreadsheet – which combines the fields necessary for ANCS+ records into a spreadsheet which may be uploaded into ANCS+.

Responsibilities

- Project Scientist: The project scientist will work with the data manager at the outset of the project to develop a database that supports both project (protocol) and programmatic needs. Please review the

database specifications and the data dictionary that were developed specifically for your project, prior to starting field work, to ensure that the database will meet project needs and will capture all raw data. In addition, please use the database as provided or discuss necessary changes with the data manager. Changes to the database may result in problems with the database itself, or incompatibilities with other I&M Program data management needs. Finally, please work with data entry personnel (especially those not involved in data collection) to ensure they understand the database and your field data sheets before entering any data.

- Data or Project Manager: The data manager will work with the project scientist to develop a database that meets both project and I&M Program needs. Database development is typically a back-and-forth process that results in an implementation that works for everyone involved. Upon completion of the database, the data manager will develop a data dictionary that describes all data fields in the database, to ensure the greatest level of understanding by all database users and data entry personnel. Finally, the data manager will also work with data entry personnel to ensure that data are being entered properly and that quality assurance methods are being followed.

- Data Entry Personnel: Please review the data dictionary and the database specifications document before attempting to use the database you have been provided. Also, familiarize yourself with the project's field data sheet before beginning data entry. In general, the basic rule is: don't hesitate to ask any questions before beginning to enter data – you certainly don't want to go back and delete entries due to some misunderstanding.

Photo Management
Standard Operating Procedure

Revision Log

Revision	Description of Change	Author	Effective Date
1	Original draft	T. Paradis	10/01/2005

Purpose

This SOP provides logical standards and guidelines for storing, organizing, naming and retrieving photographs in electronic format in the Inventory and Monitoring Program, National Capital Region Network.

Scope and Applicability

The standards described herein pertain specifically to all digital photographs related to natural resources inventory and monitoring. In this SOP, *digital photograph,* refers to any photo in electronic format regardless of acquisition by scanner or digital camera. Most photos acquired and used by I&M personnel should fit into one of the general categories:

- Library Photos. These photos are final products that have been edited, documented, reviewed and added to the I&M Digital Photo Library that is accessible through the Library catalog located on the I&M server's 'S' drive. Photos in this library are of broad interest, and may be used for multiple purposes by I&M staff. They are public domain.

- Data Photos. Data photos are photos collected as field vouchers or project data and are part of a well defined data collection protocol. Examples may include: 1) site specific photos documenting a shoreline classification and, 2) photos taken as part of a maintenance facility assets inventory. Metadata for these photos is documented and stored in catalogs at project level databases within the project folder structure. Representative, unique and instructive data photos should be renamed and added to the I&M Digital Photo Library.

- Working photos. Photos in this category are works in progress. Working photos are documented, edited and ultimately cataloged with project data, or in the Photo Library or are deleted. Working photos are stored in centrally located employee specific folders located on the I&M server's 'S' drive.

Special Collections Photos. Special collection photographs, such as from a famous photographer, historic event or an existing archived collection are cataloged and archived in specified folders within the main I&M Photo Library.

References

Digital Photo Management Strategy for the Alaska Inventory and Monitoring Program, National Park Service, Alaska Inventory and Monitoring Program, April 23, 2004

Photo Management
Standard Operating Procedure

Procedures and General Requirements

Digital photo management flow typically follows these routine steps: Photos are acquired, cataloged and documented, edited, and archived. This SOP provides logical standards and guidelines to facilitate efficient digital image asset management.

Acquisition

Digital Camera Specifications

Digital photographs should be captured at a resolution and quality setting appropriate for the highest detail intended use. As a general rule images destined for printing or publication should be of a higher quality and resolution than images that will be used for email or internet use. If the camera will allow, the resolution should be set at 1760 x 1168 or higher. The quality should be set for "super fine" or "high". Uncompressed TIFF or RAW files retain the greatest amount of image information, but the trade-off for the highest image quality is that these image files are large, and fewer images can be saved to any single memory card.

Publication quality photos should be taken at a minimum of 5 megapixels. Because the destination of a photo is unknown at the time it is taken, all photos should be taken with this resolution, or for lower resolution cameras, the highest resolution possible. Any resizing or file compression should be done during post processing using image editing software.

Though most digital cameras can digitally imprint the date and time directly onto the photo image, this feature usually should not be used. Date and time data are automatically recorded in the EXIF metadata by most digital cameras. If the image is being cataloged and documented it has value – imprinting the image reduces the image quality and hence the image value.

Different brands of digital cameras name photos differently. Some cameras will have several file naming options that may include:

- Sequential numbering which resets each time a memory card is formatted or a new card is put in.
- Sequential numbering which loops from 0001 to 9999.
- Numbering based on date-photo sequence.

Project leaders must carefully review the naming options of the project cameras and ensure that the most useful convention is used by all cameras collecting data photos.

Scanning Specifications

Print Material Types

The resolution is selected based on the size of the original. The smaller the photo, slide, or other material, the higher the resolution should be used to acquire a detailed scan. Higher resolution scans will yield larger and better quality images.

Photo Management
Standard Operating Procedure

Minimum recommended scan resolutions for different formats:

- 35mm color slide or negative (scanner should be set up to scan transparent materials - both hardware and software)
    - Choose source size of approx. 1.3 x .85 inches (software should auto-detect this exact dimensions)
    - Choose target size same as original
    - Choose resolution 2400 dpi, 24 bit color (do NOT use 32 or 48 bit
    - color)
    - Save as uncompressed TIFF file
    - yields file size approximately 3120 x 2040 pixels, file size 15-20MB
- 3 1/2" x 5" color photograph
    - Choose source size 3.5 inch x 5.0 inch (software should auto-detect the exact size)
    - Choose target size the same as the original
    - For color photo, choose resolution 600-700 dpi, 24 bit color (do NOT use 32 or 48 bit color)
    - For black & white photo, choose resolution 600-700 dpi, 8 bit grayscale (do NOT use 16 bit grayscale)
    - Save as uncompressed TIFF file
    - yields file size approximately 3000 x 2100 pixels, file size 15-20MB
- 4" x 6" color OR black & white photograph
    - Choose source size 4.0 inch x 6.0 inch (software should auto-detect the exact size)
    - Choose target size the same as the original
    - For color photo, choose resolution 600 dpi, 24 bit color (do NOT use 32 or 48 bit color)
    - For black & white photo, choose resolution 600 dpi, 8 bit grayscale (do NOT use 16bit grayscale)
    - Save as uncompressed TIFF file
    - yields file size approximately 3600 x 2400 pixels, file size 15-20MB

- 5" x 7" color OR black & white photograph
    - Choose source size 5.0 inch x 7.0 inch (software should auto-detect the exact size) Photograph Management Specifications
    - Choose target size the same as the original
    - For color photo, choose resolution 600 (450 if controls allow it) dpi, 24 bit color (do NOT use 32 or 48 bit color)
    - For black & white photo, choose resolution 600 (450 if controls allow it) dpi, 8 bit grayscale (do NOT use 16bit grayscale)
    - Save as uncompressed TIFF file
    - scan to yield file size approximately: yields file size approximately 4200 x 3000 pixels, file size 20-25MB for color
- 8" x 10" color OR black & white photograph
    - Choose source size 8.0 inch x 10.0 inch (software should autodetect the exact size)
    - Choose target size the same as the original
    - For color photo, choose resolution 300 dpi, 24 bit color (do NOT use 32 or 48 bit color)
    - For black & white photo, choose resolution 300 dpi, 8 bit grayscale (do NOT use 16bit grayscale)
    - Save as uncompressed TIFF file
    - scan to yield file size approximately: yields file size approximately 4200 x 3000 pixels, file size 20-25MB for color
- 8 1/2" x 11" typewritten/printed paper
    - Scan at resolution 300 dpi or 400 dpi if the text has very small print
    - Save as uncompressed TIFF file

Cataloging and Documentation

After importing files from a digital camera or scanner, image workflow will usually follow these general steps:

A. Establish catalog destination

B. Copy original files

C. Archive original files

D. Attach metadata to work photos (e.g., project name, description, keywords, IPTC data, custom fields)

E. Rename

F. Edit and proof

G. Convert format and resize for use

Cataloging Image Types

Original and Edited Photos

Once the original files have been imported, the unaltered files should be copied and archived. The names of folders containing these photos should clearly indicate that the folder contains unedited originals. Original photo folders may include multiple photos of the same subject, blurry pictures, or other less desirable photos.

Working Photos

Photos that are awaiting editing, proofing or processing are considered working photos. These image files should be cataloged and stored in centrally located employee specific folders. Finished images are moved to the Photo Library or the appropriate project data folder.

Data Photos

Data photos are photos collected as part of a documented data collection protocol. The project data processing protocol should contain a detailed section on processing these photos. Photos taken as part of a project's data collection protocol are project data, so they must be organized, cataloged and stored with all other project data. Data photos of interest to a greater audience should be copied, renamed and cataloged in the Digital Image Library. For more information on the storage of digital data, see the NCRN SOP for Digital Data Maintenance.

Miscellaneous Project Photos:

Incidental or opportunistic photos taken by project personnel are not data photos and can be managed as miscellaneous photos. Miscellaneous photos taken as part of a project should be stored in the project miscellaneous photos directory. This allows the photos to stay with the project, but does not confuse them with data photos. These photos may be further processed to become Library photos. Photos of interest to a greater audience should be copied, renamed and cataloged in the I&M Digital Image Library.

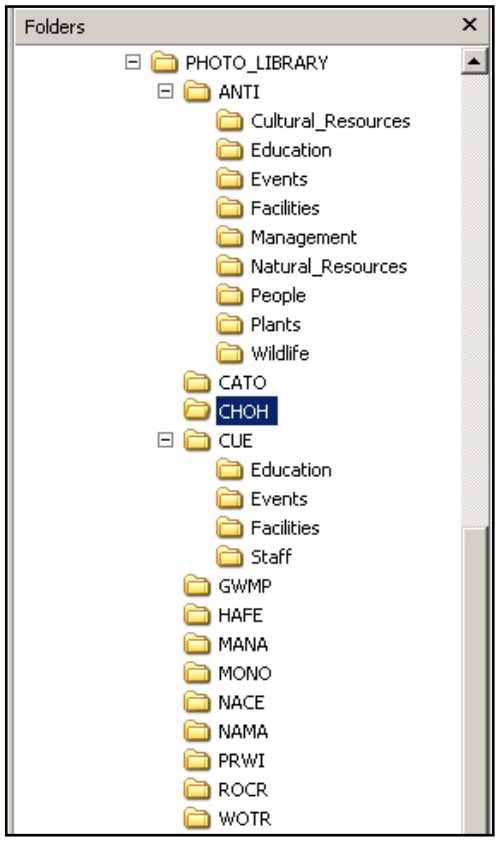


Figure 3. Sample directory structure for digital photos.

Library Photos

Library photos are public domain photos available for a variety of purposes to all I&M personnel. These photos have been processed, documented and reviewed for inclusion in the I&M Photo Library. Generally these images are high quality, high resolution, but lower quality images may be included if the subject matter has broad appeal and is special, rare or unique.

4.2b Documentatoin

All images should be documented and cataloged after importing or scanning. Documentation, in the form of metadata, provides the minimum information a user will need to appropriately use the photo. Cataloging provides a collective means for organizing, searching, and retrieving photos.

Software

Various software packages are currently under review at the time of this writing. Procedures specific to the selected program will be documented in the future.

Metadata

Specific data keywords and fields will be defined conforming to the requirements of the selected software program, and will vary with the specifics of individual projects and images. Following are typical metadata fields used for I&M photos :

- Description
- Photographer, Contributors, or Archive Institution (all that apply)
- Collection Name, if applicable (historic and archives)
- Project
- Publisher
- Park
- Location
- GPS Data
- Format/Size/Source
- Subject
- Genus
- Species
- Other taxonomic information
- Keywords
- Credits
- Distribution Restrictions, such as copyright or sensitivity

Photo Management
Standard Operating Procedure

Digital cameras attach some metadata to the image as EXIF data usually including:

- File name
- Aperture
- Date Digitized
- Date Taken
- Exposure Bias
- Exposure Time
- F-Number
- Flash (yes, no)
- Focal Length
- ISO Speed
- Light source
- Shutter Speed

Image Naming Standards

Since image asset management software possess powerful viewing and search capabilities utilizing user defined fields and metadata, detailed descriptive file names are not required to efficiently identify, catalog, or use digital photos. In order to easily differentiate Photo Library images from data photos, all I&M photos will be renamed in accordance with one of two standardized conventions.

Image management software can batch rename any number photos as part of an automated cataloging sequence saving the user time and guarding aqainst typographical errors . In some cases, renaming should be done initially as a batch process for information that is common to all images, and then in a second step to add information specific to the individual photo.

In all cases, photo names should not use spaces or special characters. As a general rule file names should be less than 20 characters.

Photo Library Naming Standards

The image file name consists of three or four parts:

1. The park code.

2. A brief description of the image.

3. The year the photo was taken, written as YYYY.

4. A numeric extender for multi-part photos or sequences of the same subject.

Each of the first three parts is separated with a hyphen and the forth, when required, with an underscore.

Examples:

ANTI-SpanishMoss-2001.jpg

CHOH-Sunset-2002.tif

MANA-Canon-1999_1.tif

MANA-Canon-1999_2.tif

Data Photo Naming Standards

Since specific project metada can be attached to the image file itself, data photo names merely provide a simple chronological connection to rest of the project data.

The image file name should consist of three or four parts:

1. The date the photo was taken, written as YYMMDD.

2. Project Code, Park Code or description

3. Chronological photo number.

4. A numeric extender for multi-part photos or sequences of the same subject.

Examples:

990829_Lichens_262.jpg

050704_Deer_097.jpg

020112_Bats _020-1.jpg

020112_Bats _020-2.jpg

4.5. Image Viewing and Editing

Viewing

Software for image viewing is under review at the time of this writing. The selected viewing software will be part of an integrated Digital Image Asset Management program.

Editing

Editing can be accomplished with any digital photo editor (e.g. Photoshop CS, Photoshop Elements)

Basic photo editing:

- Poor quality photos should be deleted, except where the subject is highly unique.
- Medium quality photos should be assessed against existing photos of the same subject in the I&M Photo Library. If the photos duplicate the subject with no enhancement of quality or perspective, the photo should be deleted.
- Photos should be rotated to portrait or landscape.
- Photos of people should have 'red eye' removed.
- Photos should be cropped to remove edge areas that grossly distract from the subject.

Metadata
Standard Operating Procedure

Revision Log

Revision	Description of Change	Author	Effective Date
1	Original draft	G. Sanders	10/01/2005

Purpose

The purpose of this standard operating procedure is to familiarize NPS staff and cooperators with the requirements that the I&M Program has for dataset documentation. This document also serves to assist personnel through metadata generation process

Scope and Applicability

The National Park Service requires that all datasets, both spatial and non-spatial, be accompanied by metadata that meets the Federal Geographic Data Committee (FGDC) metadata standards.

This document is not intended to serve as step-by-step instructions for metadata generation, but should serve as an overview of metadata requirements and provide sources that can guide the reader through the process of generating FGDC compliant metadata.

Reference Documents

Code of Federal Regulations, Executive Order 12906 (1994) Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure. Document available online at http://www.fgdc.gov/publications/documents/geninfo/execord.html

Federal Geographic Data Committee (1998) Content standard for digital geospatial metadata. Document available online at: http://www.fgdc.gov/metadata/csdgm/.

Gregson, Joe (2001) Chapter 5 of what? – Metadata. Document available online at http://science.nature.nps.gov/im/datamgmt/docs.htm

National Park Service (2004) Posting Digital Geospatial Data to the NPS GIS Data Clearinghouse. Document online at http://www.nps.gov/gis/data_info/posting.html.

National Spatial Data Infrastructure (2004) Metadata Factsheet. Document available online at http://www.fgdc.gov/publications/documents/metadata/metafact.pdf.

Procedures and General Requirements

Overview

- Metadata is really just data about data that describes the content, quality, condition and other data characteristics, using a standardized format.

Metadata
Standard Operating Procedure

- Executive Order 12906 (EO 12906) mandates that each federal agency document all new geospatial data that it collects or produces (directly or indirectly), using the FGDC Content Standard for Digital Geospatial Metadata. Please see http://www.fgdc.gov for additional information.

- In addition, metadata and related datasets must be made electronically available through the National Spatial Data Infrastructure (NSDI) Clearinghouse (http://www.nbii.gov/datainfo/metadata/clearinghouse/) or other FGDC clearinghouse nodes (with certain exceptions for sensitive information).

Tools for Metadata Generation

There is no single, standard software package used for metadata generation within the NPS I&M Program. Three programs that are typically used are: the NPS Metadata Editor, Dataset Catalog, ArcCatalog and they are described below.

- The I&M Program has developed a Metadata Editor that can function as a stand alone application or as an extension of ArcCatalog. The NPS Metadata Tools and Editor is a metadata management and editing application that implements two separate interfaces: either 1) as an extension within ArcCatalog (v8.3 or v9.0), or 2) as a standalone desktop interface. The application is intended to be the primary editor for metadata that will be uploaded to the NR-GIS Data Store. The NPS Metadata Tools and Editor application integrates with the NR-GIS Data Store information system by producing XML metadata based on the NPS Metadata Profile. This XML metadata file can then be uploaded to the NR-GIS Metadata and Data Store application.

- The Dataset Catalog was developed by the NPS I&M Program in Microsoft Access for non-spatial data that requires documentation and for cataloging all park datasets. Dataset Catalog is a tool for non-spatial data, providing a means for parks to keep an inventory of various data files, notebooks of field data forms, photographs, etc. and generate basic (non-compliant) metadata for documentation. Dataset Catalog is available online at http://science.nature.nps.gov/im/apps/datacat/index.htm, and action is currently underway to integrate the Dataset Catalog with a master online metadata database as part of the NR/GIS Metadata System. This effort will result in an NPS Integrated Metadata System for spatial and natural resources metadata.

- ArcCatalog is part of ERSI's ArcGIS software system. With ArcCatalog, you can browse, manage, create, and organize your geographic and tabular data. In addition, ArcCatalog comes with support for several popular metadata standards to allow you to create, edit, and view information about your data. There are editors to enter metadata, a storage schema, and property sheets to view the data. ArcCatalog may be used for metadata generation for all geospatial data.

Table 4. Comparison of the three main metadata generation tools.

Tool/Features	Dataset Catalog	ArcCatalog	NPS Tools & Editor
Minimal Metadata	Y	Y	Y
Comprehensive Metadata	N	Y	Y
Biological Metadata	Some	N	Y
Create metadata templates?	Y	Y	Y
FGDC Section 1	Y	Y	Y
FGDC Section 2	Some	Y	Y
FGDC Section 3	Some	Y	Y
FGDC Section 4	Some	Y	Y
FGDC Section 5	Some	Y	Y
FGDC Section 6	Some	Y	Y
FGDC Section 7	Y	Y	Y
Import	.txt	.sgm, .txt, .xml	.sgm, .txt, .xml
Export	.txt, .xml	.txt, .xml	.txt, .xml, .sgm, .html
Parsing	N	N	Y

Metadata Generation

- Metadate must be created for every GIS datalayer as well as all tabular datasets. The NCRN strongly recommends that ArcCatalog, accompanied by the NPS Metadata Tools extension, be used to generate all spatial dataset metadata. All NPS metadata should be created using the NPS Metadata Profile which includes additional fields for entering NPS specific information such as park codes.

  Parse the metadata with the Metadata Parser tool (online at http://geology.usgs.gov/tools/metadata/) which generates FGDC-compliant output in ASCII text, HTML and SGML formats – as required for inclusion on the NPS GIS clearinghouse server. Recent releases of ArcCatalog have improved parsing capabilities which may eliminate the need for the Metadata Parser

Metadata
Standard Operating Procedure

Metadata Outlined

- Identification Information – the basic information about the data set. Examples include title, geographic area covered, current-ness, and rules for acquiring or using the data.

- Data Quality Information − assessment of the quality of the data set. Examples include positional and attribute accuracy, completeness, consistency, sources of information, and methods used to produce the data.

- Spatial Data Organization Information − the mechanism used to represent spatial information in the data set. Examples include the method used to represent spatial positions directly (such as raster or vector) and indirectly (such as street addresses or county codes) and the number of spatial objects in the data set.

- Spatial Reference Information − description of the reference frame for, and means of encoding, coordinates in the data set. Examples include the name of and parameters for map projections or grid coordinate systems, horizontal and vertical datums, and the coordinate system resolution.

- Entity and Attribute Information − information about the content of the data set, including the entity types and their attributes and the domains from which attribute values may be assigned. Examples include the names and definitions of features, attributes, and attribute values.

- Distribution Information − information about obtaining the data set. Examples include contact information for the distributor, available formats, information about how to obtain data sets online or on physical media (such as cartridge tape or CD-ROM), and fees for the data.

- Multi-use Sections - the standard has multi-use sections, or templates, that allow one to "re-use" metadata elements in various sections of the standard. For example: contact information for organizations or individuals that developed the data, in Identification Information, and distribute the data set, in Distribution Information, provide for the entry of the same type of information in two distinct sections of the standard.

- Extensibility - the standard provides a methodology and process for data producers or the user community to profile and extend the metadata standard beyond the base standard to meet individual organizations and discipline metadata requirements.

❖ Note: - The metadata standard does not specify how the dataset is to be organized in a computer system, or in data transfer, or the means by which the data is transmitted or communicated to the user. The metadata standard provides information structure and content to describe the characteristics of the dataset.

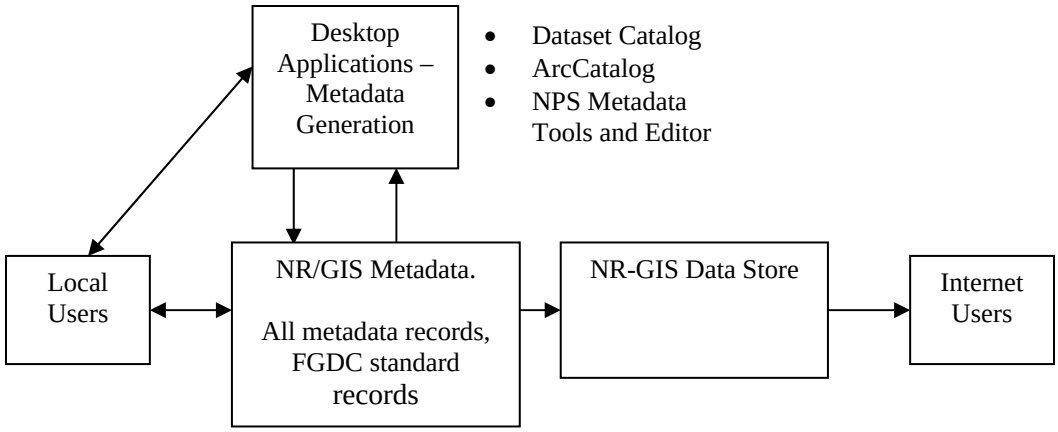
- Appendix 1 in this document contains metadata interview questions that help the person generating metadata gather all of the information required for fully FGDC compliant metadata.

- Appendix 2 is a metadata quality assurance/quality control checklist to help insure the information supplied during metadata generation is as complete and correct as possible.

Metadata
Standard Operating Procedure

NPS Integrated Metadata System

- The NPS GIS program uses NPS GIS Clearinghouse Servers at North Carolina State University and the NPS Denver Service Center Technical Information Center to store and serve vector and raster geospatial data, respectively. These clearinghouse servers require that metadata be submitted as ASCII text, HTML and SGML formats which allow the metadata to be indexed and searched by the NSDI Clearinghouse and NPSFocus for public access to geospatial data and metadata (see diagram below).



- The integrated Natural Resource GIS (NR/GIS) metadata database system reduces metadata generation to three applications: Dataset Catalog (catalogs all datasets using abbreviated metadata), ArcCatalog (for GIS-integrated geospatial metadata), and SMMS (for FGDC Biologic Profile and other geospatial metadata).

- The NPSFocus BAT Metastar Gateway is linked to the NR/GIS metadata database system and provides Z39.50 server search capabilities for the entire system.

- In order to post the data, it should be placed on a CD-ROM that contains (1) an ASCII text file listing the data set name and a one line description of the data, (2) all metadata (txt, html, and sgml formats), and (3) the digital geospatial data (Coverages should be exported as an ArcInfo interchange file (.e00) and shapefiles (shx, shp, dbf etc.) should be zipped into a single *.zip file.

- Raster data sets created by the park or contractor, such as ArcInfo GRIDs and value added USGS products and accompanying metadata should be sent to Joe Gregson at Colorado State University (really? Not Ft. Collins?) on CD-ROM. Again, the CD should contain: (1) an ASCII text file listing the data set name and an one line description of the data, (2) all metadata (txt, html, and sgml formats), and the digital geospatial data.

- Please consult with the NCRN data manager to obtain the latest information about (and assistance in) posting data and metadata to the NPS GIS Data Clearinghouse. Documentation is available online at: http://www.nps.gov/gis/data_info/posting.html.

❖ Note: if any of your data are copyrighted (e.g. Landsat scenes or LandInfo tiled DRG data) check your license agreement to see if it may be made available for public distribution via the clearinghouse.

❖ Note: do not send sensitive or classified data to the clearinghouse servers for distribution.

Responsibilities

- Project Scientist (NPS staff or cooperator): The person creating geospatial data is also responsible for FGDC compliant metadata development for any geospatial data set generated or modified as part of an NPS natural resource project. FGDC compliant metadata must accompany all geospatial datasets and is due when the final report and deliverables are submitted to the NCRN and parks included in the study. If you have not generated metadata in the past, or if you have any questions, please contact the data or project manager for assistance.

- Data or Project Manager: The data or project manager must ensure that FGDC compliant metadata is received for all geospatial data received from the project scientist (Cooperator or NPS staff). The data or project manager should provide guidance and assistance to staff and cooperators seeking assistance with metadata creation.

Appendix D
Voucher Guidance

General Guidance Document: Vouchers

Version 1.0

Revision Log

Revision	Description of Change	Author	Effective Date
1		G. Sanders	10/01/2005

Purpose

The purpose of this document is to provide Network staff, cooperators and contractors with guidance on how to collect and process voucher specimens and photos collected as part of I&M projects.

Scope and Applicability

The procedures outlined in this document pertain to all those (network staff, contractors or cooperators) who might be collecting voucher specimens or photos as part of I&M related projects taking place in any park with in the National Capital Region. These procedures apply to all projects that the NCRN I&M Program is involved in not only those sponsored or managed by the network.

Definitions and Acronyms

NPSpecies – National I&M biodiversity database.

RPRS – Research Permit and Reporting System – must be used by PI's or project leaders to obtain a permit from each park involved in a study.

ANCS – Automated National Catalog System – software used to track all NPS Voucher specimens. Maintained by park and regional curators.

Reference Documents

- Cited under individual SOP's

1. Flow Chart

   Figure 1

2. Procedures and General Requirements
   SOP – Specimen Collection Processing
   SOP – Photographic Voucher Standards

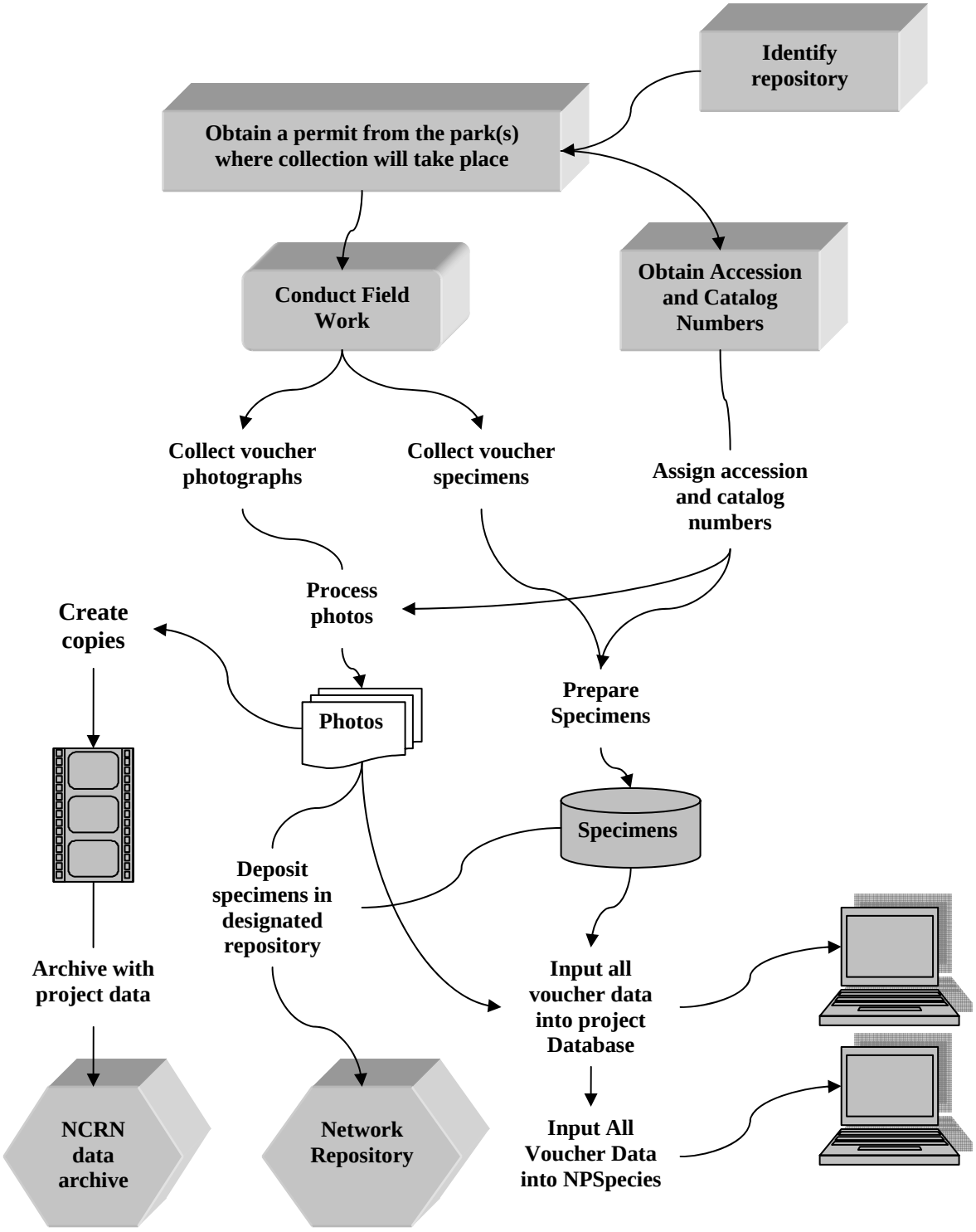


Figure 1. Voucher collection and processing flow diagram.

Specimen Collection
Standard Operating Procedure

Revision Log

Revision	Description of Change	Author	Effective Date
1		G. Sanders	10/01/2005

Purpose

This SOP was developed to establish consistent and comprehensive guidance for processing and administering voucher specimens collected during scientific studies conducted by the National Capital Region Network (NCRN) Inventory and Monitoring (I&M) Program.

Scope and Applicability

The standards described in this procedure pertain to all single or multiple park projects involving biological specimen collecting activities within the National Capital Region that in some way involve the I&M Program (via personnel time, funding, data storage and maintenance, etc.).

This document does not describe the specifics involved with photographic vouchers. For more information and the requirements involved with collecting photographic vouchers, see the SOP for photo vouchers.

This document does not describe the procedure for preparation or curation of voucher specimens. Voucher preparation and curation are activities that require great skill and professional knowledge that varies by taxa and even by species. Curation should be addressed in detail through the scope of work and budget for each project.

Reference Documents

- Code of Federal Regulations (2003), Title 36 (http://www.access.gpo.gov/nara/cfr/page1)
- National Park Service (2000) Director's Order 24, NPS Museum Collections Management (http://www.nps.gov/policy/DOrders/DOrder24.html)
- National Park Service (2003) Museum Handbook (http://www.cr.nps.gov/museum/publications/handbook.html)
- National Park Service Natural Resource Management Reference Manual #77 (draft: http://www.nature.nps.gov/RM77/)
- National Capital Region Network I&M Program (2003) SOP Product Specifications (draft)

- Research Permit and Reporting System Frequently Asked Questions (2004, http://science.nature.nps.gov/research/ac/html/RPRSFAQ).

Procedures and General Requirements

Research Permit and Reporting System

- Principle investigators must obtain a research permit from NPS prior to beginning any scientific study on a national park. The National Park Service has developed the Research Permit and Reporting System (RPRS: http://science.nature.nps.gov/research/ac/ResearchIndex) to help streamline this process.

- The purpose of the RPRS system is to help investigators: (1) apply for permission to conduct natural resource or social science field work within a specific unit of the National Park System, (2) review permit application requirements and field work restrictions before applying for permission to conduct a study, (3) review the objectives and findings of previously conducted scientific studies before formulating plans for new studies, (4) grant permission to conduct studies within parks (5) provide an avenue for submitting annual accomplishment reports, and (6) search and review the types of research activities park managers are most interested in attracting.

Specimen Collections and Guidelines

- Investigators may collect specimens upon successful completion of the RPRS application process. To the extent that specimen collection may impact a population, it is important to weigh the destructiveness of sampling against the value of the information gained, especially when collecting rare species. Photographs, sound recordings, or specimen fragments should be considered as alternatives to collecting biological specimens.

- The permit application must specify what kinds of specimens will be collected, whether those specimens will be exhausted during the investigation or permanently retained. The proposed repository must be identified for specimens that will be retained. To facilitate the permit approval process, it is required that the applicant contact the park resource manager (Appendix 1) and discuss the details of specimen retention, what data will be collected from specimens, how specimens or records will be submitted, and the proposed repository for biological collections.

- Voucher preparation is the responsibility of the principle investigator and must follow the guidelines outlined in 36 CFR and the Museum Handbook (Chapters 3 & 4 and Append. H), NPS Management Policies, RM77, Natural Resource Reference Manual #77, and Director's Order 24 for Museum

Collections. NOTE: Specimens that are to be consumed in the course of research should NOT be cataloged.

Voucher Repositories

In order to increase the efficiency of processing and tracking voucher specimens, the National Capital Region Museum Resource Center (hereafter MARS) will be identified as the repository for ALL biological specimens collected during I&M related projects. Accessioning the collection as a whole in one repository will greatly increase the efficiency of voucher processing; simplify collection management and the accountability of collections and enable researchers to access specimens from multiple parks at a single location.

Other Repositories

There are cases when the use of repositories other than MARS is appropriate

- In the case that a park wishes to maintain certain collections, that park can request to have the vouchers transferred from MARS to the park as long as the park has the proper facilities to house and maintain the specimens.
- Often, repositories other than the parks or MARS are often identified as sensible locations for certain types of collections (e.g. the USDA Mycology Collection might be the most logical and useful location to house mycology vouchers). Note all voucher specimens collected on NPS lands are the property of the National Park Service, regardless of the final repository.

    - Any time NPS collections are loaned to other agencies or institutions, NPS Form 10-127 and 10-127a, which describes the terms and conditions of the loan agreement, must be completed. All long-term loan agreements will be processed and tracked by MARS.

Cataloging Specimens

- Individual parks will no longer provide numbers (accession and catalog numbers) required by the NPS Automated National Catalog System (ANCS+) for the voucher specimens collected in their parks. All accession and catalog numbers will be requested from MARS by NCRN I&M Program Staff. This will create a "one-stop-shop" for all required ANCS+ information instead of having to contact up to eleven different parks.

    - It is recommended that accession and catalog numbers be obtained prior to data collection so that they can be incorporated directly into the project database. All NPS accession and catalog numbers will be assigned to the specimen when it is added to the project database. Doing this prior to collecting specimens will expedite the processing of vouchers at the completion of the project. The number of catalog numbers requested should be based on an educated estimate of

the number of vouchers that might be collected. Any extra numbers not used should be returned to the park that issued them.

- Studies that require multiple specimens will be assigned one accession number (ANCS) and will have individual catalog numbers (CATALOG) assigned to each specimen.

- The principal investigator must provide a project database that is compatible with NPS standard databases upon project completion containing detailed information related to each voucher collected. The information describes whole or partial specimens, audio recordings, photographs and must include the following fields of information:

    - Park Code
        - Scientific Name (following Integrated Taxonomic Information System convention)
        - Taxonomic Serial Number (TSN, according to ITIS)
        - Common Name
        - Locality
        - Latitude/Longitude or UTM X/UTM Y, Zone (for UTM coordinates)
        - Geographic Datum
        - Habitat Description
        - Collected By
        - Accession Number*
        - Catalog Number*
        - Date
        - Other supporting information may be required as determined by the individual needs or each project.

    *NOTE – the project database will only contain this information if these numbers are obtained prior to the beginning of specimen collection.

- All voucher information from project databases will be entered into the NPSpecies database (the NPS biodiversity database) by NCRN staff including the associated accession and catalog numbers obtained from MARS.

- Once the appropriate voucher information has been entered into NPSpecies, a text file (*.txt) can be automatically generated and exported from NPSpecies for importing into ANCS+. This text file contains all of the information required by ANCS+ and will automatically populate the required database fields. This will reduce the burden on park curators from having to manually enter this information into the ANCS+ system.

Specimen Collection
Standard Operating Procedure

Responsibilities

- Principle Investigator: Submit applications for research or collecting permits as soon as possible to ensure that processing does not delay the start of field work. Work with the park research coordinator or resource management staff – keeping them apprised of all park visits and research activities. Follow all rules and regulations applicable to voucher specimen collection on NPS property as outlined above. The PI is also responsible for ensuring that all of their staff are have copies of this SOP and of the collection permit and are aware of the specifics of these documents. Contact park resource managers and NCRN I&M program staff with questions regarding permitting or specimen collection. Any deviations from the guidelines described above (or in the contract/cooperative agreement) must be approved in advance by the I&M program monitoring coordinator in conjunction with park research coordinators.

- Park Resource Manager and Park Curator: Work with the principle investigator to issue research or collecting permits and to stay informed about all park visits and research activities. It is imperative that park personnel in charge of permitting ensure that all research and collecting permits are completed properly. This includes making sure that the permit meets the requirements of 36 CFR 2.5:

  > 36 CFR 2.5: *Field data, objects, specimens, and features obtained for preservation during inventory, monitoring, research, and study projects, together with associated records and reports, will be managed over the long term within the museum collection. Specimens that are not authorized for consumptive analysis will be labeled and cataloged into an appropriate cataloging system in accordance with applicable regulations.*

  Park personnel are expected to communicate all problems relating to the I&M project or data manager and ensure that appropriate txt files are imported into ANCS+ upon completion of the project and that specimen labeled properly

- Project Manager or Data Manager: Provide all applicable product and database specification documents to the principle investigator at the beginning of all contracts or cooperative agreements (prior to the start of fieldwork). Work with the project scientist and park resource management staff to ensure good communication throughout each project. Provide data management support to develop a project and voucher database which meets NPS NCRN standards. Acquire the necessary ANCS+ (Accession and Catalog) numbers from MARS and ensure that all data required to document specimen collections are collected and input into NPSpecies. Export the ANCS+ text file and forward said file to the appropriate park curator as well as to MARS for importing into the ANCS+ database.

- NCR Museum Resource Center (MARS): Issue all of the accession and catalog numbers required by I&M projects to the NCRN I&M data manager. Receive and maintain all voucher specimens collected during projects involving the I&M Program. Coordinate and track all outgoing loan agreements involving vouchers specimens collected during I&M projects with other agencies or institutions. Arrange for the transfer of voucher specimens to the park in which the specimens were collected if the park desires to house the specimens and has the proper facilities to do so.

Photographic Voucher
Standard Operating Procedure

Revision Log

Revision	Description of Change	Author	Effective Date
1		G. Sanders	10/01/2005

Purpose

The purpose of this document is to establish guidelines for collecting photo vouchers for the parks of the National Capital Region Network (NCRN)

Scope and Applicability

The following discussion provides an overview of NPS NCRN I & M guidelines related to collecting photo vouchers. The information is presented in the form of a Frequently Asked Questions (FAQ) document. These guidelines should be followed for taking photo vouchers. The process of archiving vouchers, including photo vouchers, is presented in a separate SOP (SOP Specimen Collection).

Reference Documents

- NCRN SOP Specimen Collection
- Museum Handbook, Part II, Appendix K, p. 3 Title?
- Wilhelm, Henry and Carol Brower. 1993. The Permanence and Care of Color Photographs: Traditional and Digital Color Prints, Color Negatives, Slides, and Motion Pictures, Preservation Publishing Company, Grinnell, IA.

Procedures and General Requirements

I. Film Media Requirements
II. Film Processing and Copy Format
III. Documentation
IV. Curating Developed Slides

Photographic Voucher
Standard Operating Procedure

I. Film Media Requirements and Copy Format

Media Requirements

NCRN requires staff and contractors to provide photographic vouchers as 35mm slides (preferably Kodachrome or Ektachrome), which have a proven long-term stability (Wilhelm and Brower 1993). Fuji and Kodak are both acceptable films, given that the best quality paper is used to produce the final print copy, such as Kodak's 'Duralife'. See http://www.photographic.com/film/143/ for a review of different high quality slide films.

For documenting inanimate objects (e.g. buildings, monuments, or signs), black-and-white film is the format of choice for permanence. For voucher photographs of biological specimens, however, color is necessary in order to record details of both faunal and floral. This helps determine species, gender, seasonal stage, and age.

II. Film Processing

Processing

If Kodak film is used, Kodak processing should be used, to ensure quality control in temperature and chemical composition of the processor baths. In addition to creating the color balance appropriate for the film of choice, these procedures will contribute to stability, thus longevity, of the processed film. Either Fuji or Kodak processing may be used for Fujichrome and Fujicolor film.

If a print is made with the intent to store permanently, the same emphasis on proper storage and care must be given. Fuji and Kodak are both acceptable films, given that the best quality paper is used to produce the final print copy, such as Kodak's 'Duralife'. Again, the film should be an E-6 process film.

Film should not be sent in for overnight or 1-hour processing, as the temperature and quality controls are not as stringent. However, with proper processing and subsequent storage, color print film now boasts the same 75-100 year stability as black-and-white. When you submit slide film for processing, specify that the lab use archival-quality paper mounts. If the slides return with plastic mounts, they will need to be remounted with paper mounts to meet archival standards.

If you use prints are archived instead of slides, all should be printed on Kodak Duralife paper or Fujicolor Crystal Archive paper.

Copy Format

Once images have been selected for the voucher collection, .duplicate hard copies should be made in the form of a duplicate transparency and a print for general viewing. The originals should be cataloged into the museum collection using the procedures outlined below. Original images should never be used for general purposes, but should be stored and treated as museum artifacts. The only acceptable use for originals is to

produce duplicates. Make plans to accomplish this quickly, in order that the originals spend as little time as possible outside a controlled environment.

A second hard copy should be made of each high quality original image (properly exposed, good focus, identifiable field marks, ie, your 'keepers'), for general viewing. A high-resolution, first-generation digital image should be made from the first film image, on Kodak Photo CD-R's (considered by NPS as archival as a CD can be until obsolete) for digital documentation. All of the voucher images from a single park can be stored on recordable CDs or DVDs. A second copy can be saved for general use.

III. Documentation

Field Log (appendix 2)

The field photographer should keep a field log, recording each photograph taken and the time of exposure. Appendix 2 contains a sample photo log. This is the same log as that used for photographing museum artifacts, with a field for locality data added.

Many cameras have an automatic date feature which can be programmed to insert a digital date in one corner of the negative and print but this does not replace the field log required for voucher documentation. The field log is the record for each roll. Assign a roll number when the film is loaded, placing a temporary label on your camera with the roll number. Prepare the field log, including roll number, exposure number, date, f-stop, shutter speed, (or autoexposure) subject, provenience, and lighting or other conditions.

IV. Curating Developed Slides

Label Roll

Once pictures are returned from the lab, they need to be labeled. Label each envelope or box with its date range and roll number. For this, a permanent ink pen (rapidograph or Pigma©) that will not bleed through to the pictures, or a soft lead (graphite?) pencil, will do.

Store Negatives or Slides

Negatives—Dust negative strips with an anti-static brush and place in order in archival sleeves. Label the sleeve sheets with roll numbers. Label corresponding prints with roll and exposure numbers.

Slides—Label each original slide with roll number and exposure number, using an archival plastic marker such as PEC or Kaiser plastic marking pens, or using laser-printed archival-quality slide labels and place in archival polyethelene slide sheets. Slide labels should include: a unique ID, project name, photographer, photo date, a brief identification of contents (e.g., species name, plot ID), and geographic location (UTMs or description). All slides will be stored in polypropylene slide sleeves at the NCRN office until transferred to MARS.

In addition, all slides will be scanned and saved as TIFF files. These electronic copies are used as the primary means of distributing or reproducing the images. Label copy slides or prints with roll number, exposure number and subject information. After labeling, select the best images as vouchers. Other worthwhile slides may be processed to this point as well, but be sure to note any exposures you dispose of on the Object Photograph Log. Also note the exposures selected as vouchers on this log.

Prints—Label each print you plan to keep with roll number and exposure number, and place the best ones in mylar sleeves within archival boxes. Each photo will be labeled on the back, using archival-quality labels that are either laser-printed or hand-labeled, with the same information elements required for slides. These will be used for scanning. Others may be placed in acid-free file folders, interleaved with acid-free paper. As with slides, note any exposures you dispose of on the Object Photograph Log and note the exposures selected as vouchers on the log. Photographic archival equipment can be purchased at http://www.hollingercorp.com/.

V. Summary

In summary, the following are general procedures for producing, processing, curating and storing (What's the difference between curating and storing?) voucher photographs:

1. Take pictures and record log sheet
2. Label undeveloped roll with roll number: PARK 2002-000X or PARK 2002-0X NH

3. Browse the contents of each box or folder to determine its contents, and see that you have the correct photo log information with the correct roll. If handling prints or negatives, always wear curator's gloves!

Negatives and Prints:

4. Dust negative strips with anti-static brush and place in negative strip pages in order. Do not cut strips!
5. Label strip pages with roll number, picture date and general subject matter, using a plastic marking pen (black).
6. Place hanger in strip page and file in box. Archival negative pages, slide pages, hanger strips and boxes are available through University Products and Light Impressions and Hollinger Corporation.
7. Confirm print order against negatives, slides or prints against photo log, and label print w/ roll, exposure number: "2002-00X" #XY

Slides and Prints:

8. Using a fine point radiograph ink pen with India ink, label slide or print in upper left corner with roll and exposure number. Allow time to dry completely, and check to see that ink is dry and smear-proof before stacking or placing in mylar envelope.
9. Select best exposure of each relevant view of each specimen, and place in mylar storage sleeve. Mark selected photos on photo log with asterisk, keyed at bottom of each sheet.

Photographic Voucher
Standard Operating Procedure

10. Store slides, by roll, in polyethelene slide pages.
11. Using black film marker pen, label sleeves in upper right corner with roll and exposure number. Allow to dry completely before stacking.
12. Store in photo storage boxes, filed by object type.
13. Enter photo information in Visual Inventory database.
14. Produce digital scans of selected images (remove prints or negatives from sleeves while scanning).
15. Prepare originals of these images for shipment to a processor for duplication. Be sure to pack safely and send via certified mail. Upon return of originals and duplicates, verify the return of all originals and place originals in museum storage.
16. Store unselected prints in acid-free folders with acid-free, unbuffered, lignin-free paper interleaving. Do not use glassine or kraft paper envelopes.
17. Store all photographic negatives and prints in a controlled climate:
    a) Temperature < 68 F
    b) RH 40 – 50 % fluctuation < 5 %
    c) Clean, filtered air, with acetic acids and oxidic gases removed
    d) Color prints 35 – 50 F; 20 – 30 % RH (ea –10 F doubles life expectancy, if RH stable) [1]

[1] National Park Service, Museum Handbook, Part I, Appendix R p. 14

Appendix 1. Contact information for Resource Manager/Curators at each NCRN park.

Park	Resource Manager/Curator	Work Phone	E-mail
ANTI	Ed Wenschof/	301-432-6236	ed_wenschof@nps.gov
	Alann Schimdt	301-432-5124	alann_schmidt@nps.gov
CATO	Becky Loncosky/	301-416-0536	sally_griffin@nps.gov
	Sally Griffin	301-663-9399	becky_loncosky@nps.gov
CHOH	Maria Frias Sauter	301-714-2224	marie_frias@nps.gov
GWMP	Brent Steury	703-289-2541	brent_stuery@nps.gov
	Marry Troy	703-235-1535	mary_troy@nps.gov
HAFE	Bill Hebb/	304-535-6223	bill_hebb@nps.gov
	Nancy Hatcher	304-535-6163	nancy_hatcher@nps.gov
MANA	Bryan Gorsira	703-754-1859	bryan_gorsira@nps.gov
MONO	Andrew Banasik	301-662-3515	andrew_banasik@nps.gov
	Gloria Swift		Gloria_Swift@nps.gov
NACC	Steve Lorenzetti	202-485-9680	steve_lorenzetti@nps.gov
NACE	Susan Rudy	202-690-5167	susan_rudy@nps.gov
	Mike Antonioni	202-692-6049	mike_antonioni@nps.gov
PRWI	Jennifer Lee		jennifer_lee@nps.gov
	Judy Volonoski	703-221-2176	judy_volonoski@nps.gov
ROCR	Ken Ferebee	202 895-6221 202-895-6011	ken_ferebee@nps.gov
	Perry Wheelock		perry_wheelock@nps.gov
WOTR	Duane Erwin	703-255-1822	duane_erwin@nps.gov

Appendix 2 – Photographic record and tracking forms.

PHOTOGRAPHIC RECORD

Event/Collection/Project:

National Park Service Fort Union Trading Post N H S

Photographer: Date: _____

Roll/Negative No. _____ Film Size/Type: _____

References _____

Purpose _____

DATE TAKEN	EXP. NO.	UTM COORDINATES	DESCRIPTION	EXP. INFO.	CATALOG NO.

Notes:

Vouchers:

US Department of the Interior
National Park Service

Fort Union Trading Post N H S

PARK

FILM ROLL LOG SHEET

ROLL/NEG NO.	*FILM SIZE/TYPE:*	*EVENT / COLLECTION / PROJECT / ACTIVITY*	*PHOTOGRAPHER(S)*	*DATE(S)*

P:\Curatorial\Curatorial Forms\FOUS Film Roll Log Sheet.doc

Page _____ of _____

Appendix E

Twelve Basic Inventories Conducted by the Inventory and Monitoring Program

Inventory	Description
Natural Resource Bibliographies	Bibliographies of existing research are made available to all parks to help them identify inventory needs.
Base Cartography Data	Digital cartographic products that park managers need to prepare maps and perform spatial analyses and assessments are being acquired.
Species Occurrence Inventory	Lists of the vertebrates and vascular plants currently known to occur in parks have been and continue to be compiled and verified. New field inventories are documenting additional species, especially those in plant and animal groups left out of previous inventories.
Species Distribution Inventory	New field inventories are also focusing on the distribution of species of concern to managers, including threatened and endangered species and exotics.
Vegetation Maps	All parks will be provided maps of their vegetative communities based on recent aerial photography and following a standard classification.
Soil Resources Inventory	Soils maps are being created for parks through a partnership with the Natural Resource Conservation Service. Additional products include data about physical and chemical properties of those soils and information derived from those data about potentialities and problems of use on each kind of soil.
Geologic Information Inventory	Geologic maps and digital products for parks are being completed through partnerships with the U.S. Geological Survey and state geologic agencies. Also included are an on-site evaluation of park geologic maps, resources and issues and a geologic report with basic geologic information on geologic setting and history, geologic hazards, and other geologic related issues.
Water Resources Inventory	The locations of streams, lakes, and wetlands are being documented digitally.
Water Chemistry Inventory	Water quality information is being collected for all “key” water bodies found in the parks.
Air Quality Inventory	Where the National Park Service does not have its own monitoring stations, data from U.S. Environmental Protection Agency air quality monitoring stations near parks are being summarized into an air quality atlas to assess air quality conditions in parks.
Air Quality-Related Values Assessment	Basic air quality-related information includes identification of visibility and other park resources that may be affected by air quality. The information will be available through a Web-based computer program.
Climate Data Inventory	Basic meteorological parameters such as precipitation and daily temperature are being collected.

Appendix F

Vital Signs Monitored by the National Capital Region Network

Vital Signs Monitored by the National Capital Region Network			
Category	**Vital Sign Name**	**Vital Sign Measures**	**Possible Management Application**
Air and Climate	Ozone*	Atmospheric ozone concentration	Minimize ozone exposure to visitors and staff.
	Wet deposition*	Wet deposition chemistry (pH, NO3-, SO4-), NADP	Rigorous determination of air quality impacts to watersheds within and surrounding parks.
	Visibility and particulate matter*	Visibility (PM 2.5 mass fraction)	Minimize air pollution exposure to visitors and staff.
	Mercury deposition*	Mercury Deposition Network	Supports regional air quality monitoring and serves to augment local health advisories.
	Weather*	Ambient temperature, precipitation	Supports understanding of local climate change and land use impacts.
Geology and Soils	Shoreline features	Rate of shoreline change	Determine need for shoreline restoration.
	Physical Habitat Index (PHI)	Stream habitat structure, river depth, vegetation composition on adjacent lands	Identify stream restoration areas; identify threats to stream habitat.
Water	Surface water dynamics	Flow, discharge (CFS), gauge/stage height	Identify threats to maintaining stream flow and corresponding benchmarks for restoration efforts.
	Water chemistry	Core parameters (pH, DO, specific conductance, temp), acid neutralizing capacity	Characterize the range of chemical constituents and identify excesses or deficiencies to particular core water parameters.
	Nutrient dynamics	Nitrate, ammonia, total phosphate	Identify potential pollution sources.
	Aquatic macroinvertebrates	Species composition and abundance	Identify impacts of pollution and habitat loss to benthic communities.
Biological Integrity	Invasive/exotic plants	Detection of new species in sensitive areas	Identify where exotic plant eradication efforts are needed.
	Forest insect pests	Presence of new pest species	Control new infestations before they become a resource threat.
	Forest vegetation	Species diversity, age and size classes	Assess need for deer management; trigger invasive species or pest species management; evaluate effectiveness of exotic plant treatment.
	Fishes	Species composition	Identify pollution sources; trigger restoration efforts.
	Amphibians	Species composition and proportion of area occupied, malformations	Trigger wetland or upland management for significant amphibian habitats; identify pollution sources.
	Landbirds	Species composition and abundance	Recommend and evaluate forest and grassland management efforts; inform regional assessments of significant bird areas.
	White-tailed deer	Deer density	Assess deer management efforts; protect herd health and vegetation communities.
	T&E species and communities	Presence and absence, status, threats	Trigger site restoration; meet legal/policy requirements.
Landscapes (Ecosystem Pattern and Processes)	Land cover/land use	Area of dominant land cover types, connectivity, core/edge ratio of dominant forest communities, weighted average patch size, adjacency matrix	Evaluate ecosystem conversion; inform regional assessments of land use; derive importance of parks in urbanizing landscape.
	Landscape condition	Landuse intensity, disturbance status	Evaluate cumulative habitat change; trigger habitat restoration.

* Vital signs that are already monitored by partner agencies.

Appendix G
Data Collection

General Guidance Document: Data Loggers

Revision Log

Revision	Description of Change	Author	Effective Date
1	Original	G. Sanders	Oct. 1, 2005

Purpose

The purpose of this guidance document is to provide a broad overview and general guidance regarding the use of automated data collection devices. As technology advances new and innovative methods of collecting field data are developed and many of these new technologies offer added benefits such as helping to increase data integrity and reducing the need for intensive QA/QC. The use of these devices does not come without new concerns that need to be addressed. This document will describe the general issues that users need to be aware of when using any kind of computerized data collection device. More specific guidance with regard to the use of these devices can be found in the individual protocol data collection SOPs.

Scope and Applicability

This guidance document applies to all computerized or electronic devices that are used for data collection on I&M or regional projects. The specific audience that should refer to this document includes I&M Network staff as well as cooperators, contractors and regional park staff who may be involved with the use of these devices.

Definitions and Acronyms

- Data Loggers – devices that record data regarding environmental conditions with very little, if any, user input. Examples of such devices include stream flow meters, water chemistry meters and atmospheric monitoring devices.
- PDA – Handheld computer such as a Palm OS or Windows CE device.

1. Flow Chart

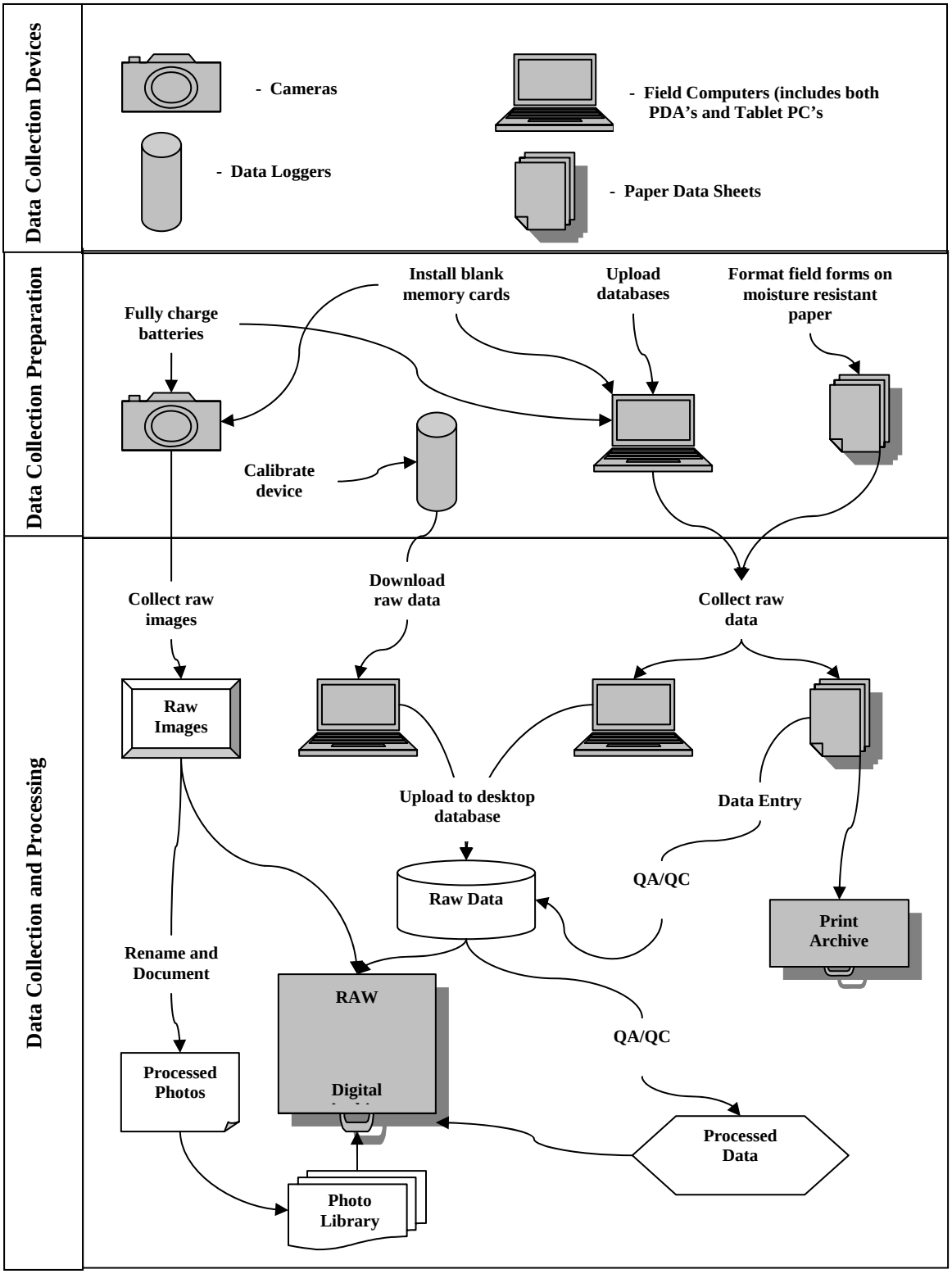


Figure 1. Flow chart indicating data collection methods and the appropriate steps to take during instrument preparation, data acquisition, and data processing.

Procedures and General Requirements

Data Collection Guidance

There are numerous methods for collecting data in the field ranging from paper field forms to automated computers and sensors that are permanently deployed and collect information independent of user input. The collection technique used will depend on the project and the specific data collection requirements. Listed below are some of the tools available for field data collection. Refer to the protocol specific SOP's for details on how the following tools may apply to individual projects.

- Field Forms – the most common method of recording field data. Use a formatted, project-specific data sheet as opposed to a field notebook. We strongly recommend acid-free paper to prevent fading and subsequent data loss as well as moisture resistant paper such as Rite in the Rain. Some circumstances may warrant the use of paper and writing implements that can withstand, dust, and other extreme environmental conditions. Standardized data sheets that identify the pieces of information to be recorded and forms that reflect the design of the computer data entry interface will help ensure that all relevant information is recorded and subsequent data entry errors are minimized. Data sheets should contain as much basic preprinted project information as possible and sufficient space for recording relevant metadata such as date, collectors, weather conditions, etc. They should clearly specify all required information, using examples where needed to ensure that the proper data are recorded. Data recorders should adhere to the following guidelines:

    1. All information added to the data sheet must be printed and clearly legible.

    2. If alterations to the information are necessary, the original information should be crossed out with a single line and the new information written next to the original entry. Information should never be erased and old information should not be overwritten.

    Upon return from the field, copies of all original data sheets should be made and checked for legibility and completeness (i.e., no data cut off at the edges). The original datasheets should be used for data entry after which they will be stored in the fire proof cabinet. Copies of original datasheets should be archived offsite.

    Original forms are archived.

1. Field Computers – increase data collection and data entry efficiency by eliminating the need for paper field forms. Data can be downloaded directly from the field computers to the office desktops thereby eliminating the data entry step. Fewer chances for error as QA/QC checks can be built into the field database. Could be inefficient if copious amounts of notes or comments need to be recorded in the field.

    These portable units, however, are subject to environmental constraints such as heat, dust, and moisture. When handheld computers are used for data entry in the field, the data should be downloaded daily to avoid potential loss of information. No field computer should be used unless it is equipped with a removable flash memory card to store field data in case of a catastrophic failure of the field unit. Batteries should be checked prior to a data collection trip, and they should be charged at the end of every field day. Also, in case the unit becomes inoperable in the field, printed data sheets should always accompany field teams on data collection trips.

    - Handheld computers (PDA's) –The small size and relative low cost of these devices makes them attractive options for collecting field. Good for small field projects but not powerful enough for large data intensive field projects. PDA's can be ruggedized fairly easily and at a relatively low cost. Most run either Windows CE or Palm operating systems which may require additional processing/programming to transfer/create the database structure in the field units. Software applications such as Pendragon Forms® can assist with this task on both Palm and Windows based devices

- Tablet PC's – same properties as most laptops and provide the user with the convenience of a touch screen interface. They are bulkier, more expensive and harder to ruggedize than PDA's but are more powerful as well. Good for field projects that are very data intensive. Because these units run Windows XP (Tablet Edition) the project database can be directly transferred from desktop units to field units without additional programming steps.

1. Data collection should not take place until the unit and the database application are fully tested.

2. After data collection is complete, data should be offloaded from the device or memory card to an office computer immediately.

3. All raw data should be archived.

4. Data must undergo QA/QC.

5. Once data has been validated, the validated dataset should also then be archived.

2. Automated Data Loggers – mainly used to collect ambient information such as air or water quality information. These devices can be calibrated and programmed to automatically record data and store them for later download directly to a computer, thereby eliminating the possibility for manual data entry errors. Must be properly calibrated and maintained so field crews must receive proper training and SOP's outlining the calibration/unit maintenance procedures. SOPs are currently in development and will be distributed to all field crews using these devices.

   - Permanently deployed devices – often cost prohibitive. Data from these devices must be retrieved and batteries changed on a regular basis. These intervals should be defined in the protocol.

   - Portable hand-held devices – deployed for sampling only during site visits. Generally less expensive than units that are permanently deployed in the field. Prior to and following field visits device components should be inventoried to ensure that all necessary equipment is accounted for.

   1. Data is downloaded to a mobile PC or PDA.

   2. Data is uploaded to a desktop PC immediately and raw data is archived.

   3. Data is examined to determine if there are any anomalies or indications that the device did not properly record conditions.

   4. Once the data has been validated it is incorporated into the desktop database application.

3. Cameras – Photographs provide an excellent visual record of field visits. Cameras are useful for capturing photo point records of long-term study sites. They also serve well for automated data collection by remotely recording information using web cams or trip cameras. If photographs are intended to serve as vouchers, refer to Appendix J for specific details regarding photographic vouchers. Refer to the NCRN Photo SOP for more specifics on digital photography standards.

Appendix H

Non-Programmatic External Data that Support NCRN Vital Signs Monitoring

Vital Sign	Subcategory	Agency	Website (if available)	Contact[a]	Site location (CODE)	Parameter(s) of Record (units)
Ozone	Day by day and site by site exceedences	MWCOG	http://www.air-watch.net/exceedances.php?pollutant=0&Hour=11&Day=20&Month=10&Year=2004	Jen Desimone	Aggregated	Number of 8-hour standard exceedances (days)
	Hourly concentrations (site specific)	MWCOG	http://www.air-watch.net/exceedances.php?pollutant=0&Hour=11&Day=20&Month=10&Year=2004	Jen Desimone	TBA	Ozone (ppb)
	Spatially aggregated SUM06 (EPA-NAAQS standard for vegetation effects)	NPS	Air atlas (ArcIMS retrieval website) http://massive.natnet.du.edu/website/ozone/	John Ray	Aggregated	SUM_06
	Hourly concentrations (site specific)	EPA	CASTNet (Airs data) http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm	John Ray	TBA	Ozone (ppb)
	Calculated Sum06trends	TBA	TBA(?)	TBA	Aggregated	SUM_06
Sulfur wet deposition	Weekly deposition	NADP	http://nadp.sws.uiuc.edu/	John Ray	TBA	SO4 deposition (kg/ha/yr)
	Annual/Decadal scale trends	NADP	TBA--Chris Lehmann or NPS ARD	Chris Lehmann		SO4 deposition (kg/ha/yr)
Nitrogen wet deposition	Weekly deposition		http://nadp.sws.uiuc.edu/	John Ray		N deposition (Kg N/ha/yr)
	Annual/Decadal scale trends		TBA--Chris Lehmann or NPS ARD	Chris Lehmann		N deposition (Kg N/ha/yr)

NOTE:

a. For data not available on web, include phone and email.

Vital Sign	Subcategory	Agency	Website (if available)	Contact[a]	Site location (CODE)	Parameter(s) of Record (units)
$PM_{2.5}$	$PM_{2.5}$ concentration	IMPROVE	http://vista.cira.colostate.edu/improve/	John Ray	WASH	PM2.5 concentration
	IMPROVE calculated visibility trends	IMPROVE	http://vista.cira.colostate.edu/improve/Data/GraphicViewer/Trends.htm	John Ray	WASH	worst 20%; best 20% of days
Mercury Wet deposition	MDN deposition	NADP-MDN	http://nadp.sws.uiuc.edu/mdn/	John Ray; Alan Kolker	TBA	Hg (ug/m2/yr)
	MDN spatial and temporal trends	NADP-MDN	TBA	John Ray/Chris Lehmann	TBA	Hg (ug/m2/yr)
Fish Tissue Mercury	Body burden (in common fish, e.g., bass) equivalents	USGS/EPA	http://emmma.usgs.gov/mapping.aspx	TBA	variable	ppm
	Trends in body burden (in same equivalents)	USGS/EPA	TBA	TBA	Aggregated	ppm
Climate	Precipitation	NOAA	TBA	TBA	TBA	Temp (averaging time and spatial consideration TBA)
	temperature	NOAA	TBA	TBA	TBA	Temp (averaging time and spatial consideration TBA)
Remote Sensing	Land cover/Land Use Change	TBA	TBA	TBA	Various	Various
	Shoreline Change	TBA	TBA	TBA	Various	Various

NOTE:

a. For data not available on web, include phone and email.

Appendix I
NPSpecies Data Entry
Standard Operating Procedure

NPSpecies Data Entry
Standard Operating Procedure

Revision Log

Revision	Description of Change	Author	Effective Date
1	Original	G. Sanders	Oct. 1, 2005

Purpose

This standard operating procedure (SOP) provides a quick introduction to entering biological inventory data into the online and desktop version of NPSpecies.

Scope and Applicability

NPSpecies is the National Park Service database to store, manage, and disseminate scientific information on the biodiversity of organisms in National Park Service units throughout the United States and its territories. The database is available in an online version and a desktop version for parks or offices that have slow connections to the internet. Data should always be entered into the online version of NPSpecies if possible except to add ANCS+ data as indicated. The off-line version may be helpful to generate reports and perform data analyses that are not possible using the online version.

Reference Documents

NPSpecies: http://science.nature.nps.gov/im/apps/npspp/Discover.htm

NPSpecies Desktop Application:
http://science.nature.nps.gov/im/apps/npspp/DesktopApp.htm

Data Dictionary and Training Documents:
http://science.nature.nps.gov/im/apps/npspp/Documents.htm

Procedures and General Requirements

Entering Data into NPSpecies (Online)

Log into the NPSpecies webpage (https://science1.nature.nps.gov/npspecies); you will need a user name and password (available from NCRN Data Manager: Geoff Sanders 202-342-1443x230).

Select “Secure Utilities” from list at left.

Select “Edit Records.”

Choose the park for which you wish to enter data and click “Submit.”

Select the group for which you will be entering data (e.g., mammal, bird, vascular plant) under the correct heading (e.g., references, vouchers, observations).

NPSpecies Data Entry
Standard Operating Procedure

Add References: click on "Add New Reference" if the reference has not been entered into NatureBib already. You can then add the information for the reference (e.g., title, author, date, keywords, abstract).

To search for existing references: click the "Search For and Link to a Reference Not Listed Below" button. You can then search for the reference by title, author, Bibkey ID#, etc. If the reference already has some information associated with it in NPSpecies, it will show up in the list of references in the left column and you can simply click on it to bring up the current list and then add to it or edit it.

To add species to an existing reference: click on the "Add Additional Names to This Reference" button once the reference information is displayed in the right-hand portion of the screen (i.e., after you have added a new reference or clicked on a pre-existing one). This will bring up a pick-list of all species linked to the park. You can add any species on this list to the list for that reference simply by clicking on the species name (sorted alphabetically within taxon group).

To add multiple species at a time: hold down the Control key when selecting species. Selected species will appear highlighted on the list. Click "Submit Edits" once you have selected the species you wish to enter. The species should then appear below the bibliographic information for the reference (bottom of the right side of the screen). You can add as many species as you would like at a time, but be careful about trying to select too many at once because if you forget to hold the Control key down when you select one, all other highlighted species will become un-highlighted, and then you have to go through and select them all again.

To delete species from a reference: click on the "Delete Names From This Reference" button. This will bring up a pick-list of all the species currently entered for the reference. You can select any species desired (use Control key to select more than one), and then click "Delete."

To add a new species: If you need to enter a species name that is not already linked to the park (doesn't show up on the pick-list), click on "Species" in the left column toward the top. This will bring you to the overall park species list for the taxon group you are currently under. If you need to enter a species for a group other than the one you originally entered the reference under, click on "Other Species Categories" and choose the desired taxon group from the far left list ("Park Species List"). You can then simply click on the "Add New Species Name" button and type the species you would like to add into the top-most box in the "Name Search" pop-up box. You can enter only a portion of the species name as well (e.g., just the genus) and it will search for any species containing that portion. After you enter the name and click on the "Search" button, the matches in ITIS will show up in the bottom portion of the pop-up box. You can then select the name you wish to add (you may click on "Taxonomic Profile" to get more information about your selection before adding it), and then click on the "Add to Park-Species List" button. Then just fill in the information you

have (you can enter park status and such if you know it, but you can usually leave these blank) and click on the "Submit Edits" button. The species is now on the park species list and can be linked to a voucher, observation or reference. If the species you wanted had not shown up in the list after you did your ITIS search, you can add a new species into ITIS. To do this, click on the "Add New Name" button and fill in the new information, and then follow the same procedure as above.

Add Vouchers: click on the "Add New Voucher" buttons at the top left of the screen and then add your information (enter as much information as you have; it will prompt you if you have skipped any required fields).

Put specimen identification number from PI into Comments box.

Also see adding ANCS+ data in Section III below.

Add Observations: click on "Add New Observation" then add your information (enter as much information as you have; it will prompt you if you have skipped any required fields).

Downloading NPSpecies Data for the Desktop Version

To download the desktop version of NPSpecies, go to http://science.nature.nps.gov/im/apps/npspp/DesktopApp.htm

From the NPSpecies home page (On-line version), select "Secure Utilities" from the list at left.

Select "Download Data."

Choose the park you would like to download data for and click on the "Submit" button (you can only choose one park at a time). The date next to the park name is when the data was last changed. You do not have to download data if it has not changed since last time you downloaded.

Choose to save the data to a disk or to your hard drive.

Using the Desktop Version of NPSpecies

When you start up the MS ACCESS program, a window will open telling you which files the program is linked to (or asking you to link files, if that has not already been done). The Working Data File is the file with the specific park data that you downloaded from the online version of NPSpecies. If the program is not linked to the Working Data File for the park you wish to work on, select the "Tools" button from the main screen and browse to the file for the desired park. Once you have browsed to and selected the park file (Working Data File) you would like to work on, click on the "Link Files" button. The rest of the functionality is fairly self-explanatory.

For Adding ANCS+ data: click on the vouchers tab. All the vouchers that have already been entered on the online version should be listed.

Click on the tree icon next to a voucher record.

Under View, select ANCS+. The entire record will appear with ANCS+ data fields.

Enter the appropriate Catalog data and ANCS+ data. Then hit the Specimen ID button to auto-generate a Specimen ID.

When all the voucher data has been entered, click the ANCS Export button to create a txt file. Save the txt file and send it to the park museum specialist.

NOTE: *DO NOT DELETE ANCS DATA from NPSpecies when asked. Delete only after ANCS data has been successfully imported by the park museum specialist.*

Appendix J
GPS Specifications
Standard Operating Procedure

GPS Specifications
Standard Operating Procedure

Revision Log

Revision	Description of Change	Author	Effective Date
1	Original	G. Sanders	Oct. 1, 2005

Purpose

This standard operating procedure (SOP) describes the procedures used to collect Global Positioning System (GPS) data of acceptable quality in the field. This document also provides information on instrument settings, field operations, and standards in the recording of positional data.

Scope and Applicability

GPS is a constellation of 24 Department of Defense satellites that orbit the earth approximately every 12 hours, emitting signals to earth at precisely the same time. The position and time information transmitted by these satellites is used by GPS receivers to trilaterate a location coordinate on the earth using three or more satellites.

The standards described in this SOP pertain to all data collected with hand-held or backpack GPS receivers. This document focuses upon field operations and instrument settings to be used when collecting GPS data, and includes some recommended standards for collecting positional data. This document is not intended to serve as a training manual for those new to GPS data collection. For more detailed information regarding the collection of GPS data, please consult the Reference Documents.

Reference Documents

- A reference manual for creating quality GIS data with GPS technology is available at http://www.nps.gov/gis/gps/gps4gis/
- Operating manuals for each individual device.

Procedures and General Requirements

All GPS Units

Always do differential corrections. Post-process all data, even if real-time differential correction is used in the field.

Check the graphics data collection screen regularly to see if you are getting multi-path or other apparent distortions to the data. Garmin and PLGRs require the user to monitor the screen and stop data collection during poor PDOP or SNR windows. Trimble receivers set to the appropriate mask (see above) will stop collecting data automatically.

Be aware of the possibility of multi-path interference, and use offsets or other methods to keep the antenna away from building overhangs, tall fences or walls, and heavy canopy whenever possible.

If maximum accuracy is required, it is important to sync the collection rate with the base station logging rate. Stations log anywhere from 1 to 30 second data. It is recommended that logging rates be in multiples of 1 or 5 for best differential corrections. Setting logging rates other than 1 and 5 may reduce the number of positions that are in sync with base data and reduce accuracy.

Map all features in a single area on a single day or on consecutive days.

Be aware of the ambient conditions and how they may affect the unit. The LCD screen may be hard to see in bright sunlight or if the user is wearing polarized lenses. Extreme temperatures may affect the display as well. Most LCD displays will "black out" due to extreme temperatures.

Garmin

Estimated Horizontal Error (EHE) should be less than or equal to 12 meters. This will keep you just within National Map Accuracy Standards for a 1:24,000 scale map – which defines the maximum acceptable error.

Use a minimum of 4 satellites (3D) for every position.

For point data, collect 90-120 positions at 1-2 second intervals and average.

For line or polygon data, use a 2-5 second interval for walking and road driving, depending upon the road type and speed of the vehicle. Force (e.g., wait for) a position at each corner and use a minimum of 3 positions to define any curve or change in direction.

GPS projection settings should match the projection requirements used by the park where the data is being collected.

PLGR

Estimated Horizontal Error (EHE) should be less than or equal to 12 meters. This will keep you just within National Map Accuracy Standards for a 1:24,000 scale map – which defines the maximum acceptable error.

Use a minimum of 4 satellites (3D) for every position.

For high accuracy point data use the average mode (AVG). Collect data at a single point for at least 60 seconds but preferably two minutes. Be sure to switch to continuous mode (CONT) when moving the unit.

GPS projection settings should match the projection requirements used by the park where the data is being collected.

GPS Specifications
Standard Operating Procedure

Trimble Pathfinder Unit (PRO XRs, Recons, and GeoExplorers) Settings

Positional dilution of precision (PDOP) should be less than or equal to 6. Start with a maximum PDOP setting of 4, and shift to 5 if data collection is not successful. This will meet National Map Accuracy Standards for a 1:5,000 scale map.

Use a minimum of 4 satellites (3D) for every position.

Signal to Noise Ratio (SNR) should be less than or equal to 5.

Set Elevation Mask to 15.

Be sure to check the Antenna Height setting – which should be the typical height at which the antenna will be carried. If the antenna is attached to a pole, it must be located above the user's head and the antenna height setting should be the height at the top of the pole. Whenever possible, the antenna should be clear of any obstructions.

For point data, collect a minimum of 30 positions at 1-second intervals and average.

For line or polygon data, use a 2-5 second interval for walking and road driving, depending upon the road type and speed of the vehicle. Force (e.g., wait for) a position at each corner and use a minimum of 3 positions to define any curve or change in direction.

Data should be exported in the projection used by the park where data was collected.

Park Unit Data Standards

All digital geospatial data should reference the coordinate system appropriate for its use and it should be documented in the metadata. When collecting GPS data in the field, the projection used will depend on the park where data is being collected. A list of standard park projections is included at the end of this SOP.

Horizontal/Vertical Accuracy and Precision

All spatial data collected will be analyzed for spatial accuracy and will meet or exceed the National Map Accuracy Standards for the intended scale and use (for more information, please see http://mapping.usgs.gov/standards). Any calculations done with location data should be done at double precision with the results rounded or truncated to the appropriate propagated error limits. All calculations and processing completed on the spatial data will be reported in the metadata.

Positional coordinate data should NOT be recorded in NAD27 in the field. Datum conversions should be done as an office, post-processing activity using software that applies a full NADCON datum conversion in order to assure accuracy and precision.

Real-time differential correction techniques should be employed whenever possible for efficiency and time savings. The distance between the base station and the remote GPS receiver should be kept to a minimum, preferably less than 150 miles.

GPS Specifications
Standard Operating Procedure

Responsibilities

Data Collectors

Be certain that you understand how to use the GPS unit, data logger, and any study-specific data collection forms prior to going into the field and collecting data. Ask questions or request training as needed. There is nothing worse than wasting a day, or more, of field work due to improper use of the equipment. Plan your trips to the field in advance, and target times of the day when the most satellites will be overhead (check the latest almanac at http://www.trimble.com/gpsdataresources.html) to get the most of your time in the field. Also, spend time (at the beginning of the project) going over the data you are collecting and differentially correcting to make sure you are collecting everything that is needed, in a proper way. Don't wait until the end of the project to do this or an entire field season may be wasted when errors could have been corrected early on. If something doesn't seem right to you, please ask.

Data or Project Manager

Provide data collection personnel with training and/or instruction in the use of all GPS units, data loggers, and any study-specific data collection forms prior to sending them out into the field to collect data. If possible, accompany all new data collectors in the field until you both are certain about and comfortable with all data collection procedures. Be available to answer questions from data collection personnel, and require that early data collection efforts be thoroughly analyzed to minimize systemic mistakes that could greatly diminish the usefulness of the data collected. If possible, examine data from initial data collection efforts as soon as possible, to catch any systemic errors that need correction.

GPS Specifications
Standard Operating Procedure

Standard Park Projections:

ANTI
UTM – meters
GRS 1980
Zone 18N
NAD 83

CATO
UTM – meters
CLARK 1866
Zone 18N
NAD 27

CHOH
MD STATE PLANE
(feet)
GRS 1980
Zone 4126 (FIPS zone
1900)
NAD 83

GWMP
VA STATE PLANE (feet)
GRS 1980
Zone 4501 (north)
NAD 83

HAFE
WV STATE PLANE
(feet)
CLARKE 1866
Zone 5651N
(FIPS Zone 4701)
NAD 27

MANA
UTM – meters
GRS 1980
Zone 18N

MONO
UTM –meters
GRS 1980
Zone 18N
NAD 83

NACC
MD STATE PLANE (meters)
GRS 1980
Zone 4126
NAD 83

NACE
MD STATE PLANE (meters)
GRS 1980
Zone 4126
NAD 83

PRWI
UTM – meters
Clarke 1866
Zone 18N
NAD 27

ROCR
MD STATE PLANE (meters)
GRS 1980
Zone 4126
NAD 83

WOTR
UTM – meters
GRS 1980
Zone 18N
NAD 83

Appendix K
Park Status Report
Standard Operating Procedure

Park Status Report
Standard Operating Procedure

Revision Log

Revision	Description of Change	Author	Effective Date
1	Original	G. Sanders	Oct. 1, 2005

Purpose

The purpose of this standard operating procedure (SOP) is to provide parks within the NCRN a summary/update detailing inventory and monitoring activities occurring in each individual park. The information for the Park Status Reports (PSR) is based on the AAWPR.

Procedure and General Requirements

The report lists bullets for every inventory and monitoring project related to each of the 11 I&M parks. Standard topic headings for the bullets include:

I. Inventories

1. Completed Inventories funded by NCRN I & M Program
2. Ongoing Inventories funded by NCRN I & M Program
3. Inventory Projects Planned

II. Monitoring Projects.

1. Ongoing Monitoring Projects

III. Data Management

IV. GPRA Goals

V. Research Questions

VI. Summary Highlights

See the example in attachment 1.

Note following points:

1. For sections I and II – Budgets represents total project budget and not the amount spent on the park.
2. For section II. Draft Vital Signs – This is table 3.4 in the Draft Monitoring Plan.

3. For section III. The NPSpecies table is based on the Standard Report generated from the Desktop Version of NPSPecies.

    a. Under Analyses: The "summary stats" provides species counts for all taxonomic groups.

    b. Under All Data: The number of references, vouchers, and observations are gathered by clicking on the icons and looking at the number of records generated.

Attachment 1. Example Park Status Report

Inventory and Monitoring Information – CATO

CATO - Park Status (Last Update: 7/5/04)

I. Biological Inventories

1. Completed Inventories funded by NCRN I & M Program

- Bird Inventory – PI John Sinclair and volunteer effort. Completed Fall 03. Total Project Budget: None.
- Small/Medium Sized Mammals – PI Bill McShea (Smithsonian Inst.). Completed Fall 03. Total Project Budget: $53,500.

2. Ongoing Inventories funded by NCRN I & M Program

- Bats – PI Ed Gates (UMCES). Expected Completion Fall 04. Total Project Budget: $110,994.
- Fish – PI Rich Raesley (Frostburg State University). Expected Completion Fall 04. Total Project Budget: $70,457.
- Paleo. Resources – PI Vincent Santucci (GWMP). Expected Completion Fall 04. Total Project Budget: $7,000
- Herps – PI Tom Pauley (UMCES). Expected Completion Fall 04. Total Project Budget: $114,000.
- Vegetation Mapping – NPS Lead: Diane Pavek. Partial funding received from I & M. Total Project Budget from IM Program (FY01-FY03 only): $210,854 (more may be added FY04). Expected completion?

3. Inventory Projects Planned time frame?

- Complete QA/QC and data certification for all completed inventories in FY04 and FY05.

II. Monitoring

Park Status Report
Standard Operating Procedure

1. Ongoing Monitoring Projects

i. Vital Sign Selection

The I & M planning process has identified a variety of potential vital signs based on input from park managers and subject matter experts. Final selection of vital signs will be based on prioritization once monitoring protocols and costs are known. The final selection of vital signs must also be approved by the Science Advisory Committee and Board of Directors. See Appendix 1 for list of draft vital signs.

ii. Protocol Development

Current monitoring projects focus on the development of monitoring protocols. The I & M staff is working with a variety of principal investigators to identify sampling methods, sampling sites, reporting and analyses mechanisms. Note that sample sites have not yet been selected, and they may or may not occur in CATO.

- Amphibians – IAA with Patuxent Wildlife Research Center. PI is Dr. Robin Jung. Focus will be on species diversity and distribution. Total Project Budget: $51,842.
- Forest Bird Species – IAA in development with Patuxent Wildlife Research Center. PI is Deanna Dawson. Focus will be on species diversity and abundance. Total Estimated Budget: $25,000.
- RTE Species – CA with Virginia Polytechnical Institure. PI is Jeff Waldon/Allison Wells. Focus will be on species meeting I & M RTE criteria. Total Project Budget: $77,146.
- Remote Sensing – CA with Frostburg. PI is Phil Townsend and Bob Gardner. Focus is on monitoring change of major landcover types in park and region. There will also be a classification of land use intensity. Total Project Budget: $150,000.
- Vegetation Monitoring – IAA with USDA Enterprise Unit. PI is David Chojnacky. Focus will be on forest health indicators, including deer browse and cover of invasive and exotic species. Total Project Budget: $130.000.
- Vital Sign Conceptual Model – CA with Frostburg. PI is Phil Townsend and Bob Gardner. Focus is on developing scientific basis for conceptual models that describe ecosystems in NCRN parks. Total Project Budget: $71,792.
- Water Quality – CA with Bob Hildebrand (UMCES). Focus is on core water quality parameters required by WRD. Monitoring will also include macroinvertebrate and nutrient monitoring, fish index of biological integrity. Total Project Budget: $61,207.

- White-tailed Deer – Distance Sampling. PI is Scott Bates (CUE).

III. Data Management

I & M Inventory database (NPSpecies) contains 1,121 records (Table 2). The bird data has been certified and QA/QC will continue during FY04 and FY05.

Table 2. Summary statistics of data managed by I & M in NPSpecies (Updated 7/5/04).

Taxonomic Group	Number of Species Entered
- Amphibians	33
- Birds*	173*
- Fish	17
- Mammals	59
- Reptiles	29
- Plants	818
- Non-Vertebrates	171
- Fungi	101
- Monera	0
- Protista	0
References (also in NatureBib)	32
Observations	2120
Vouchers	258

*Data met QA/QC requirements and has been certified by CATO.

Park Status Report
Standard Operating Procedure

IV. GPRA Goals Summary

The I & M program generates data that can be used in part to report GPRA goals. Only two GPRA goals are reported by WASO (Ib1and Ib3). Table 3 summarized natural resource related GPRA goals.

Table 3. GPRA goals (FY03) for each park that pertain to information generated by the Inventory and Monitoring program of the National Capital Region Network. This is Table 1.2 from the Draft Monitoring Plan.

GPRA Goal	Goal #	Parks with this goal										
		ANTI	CATO	CHOH	GWMP	HAFE	MANA	MONO	NACE	PRWI	ROCR	WOTR
Disturbed lands restored	Ia1A	X	X	X	X		X		X			
Disturbed lands (other)	Ia01A										X	
Exotic vegetation contained	Ia1B	X	X		X	X	X	X	X	X	X	X
Species of concern populations have improved status	Ia2X								X			
Natural resource inventories*	Ib1											
Natural Resource Inventories (park based)	Ib01	X		X	X			X	X			X
Vital signs for natural resource monitoring identified*	Ib3	X	X	X	X	X	X	X	X	X	X	

Park Status Report
Standard Operating Procedure

GPRA Goal	Goal #	Parks with this goal										
		ANTI	CATO	CHOH	GWMP	HAFE	MANA	MONO	NACE	PRWI	ROCR	WOTR
Water quality improvement	Ia04		X							X		

*Indicate GPRA goals tracked by WASO I & M Program.

V. Research Questions

Every year, the I & M Program identifies research questions as they come up through ongoing inventory and monitoring efforts. The questions below are not prioritized.

1. What is the long-term impact of deer browse on forest regeneration?
2. What is the significance of the Table Mountain Pine and the role of fire ecology in sustaining it?

VI. Summary Highlights (FY04)

- Bird inventory is complete and data is has been certified in NPSpecies.
- I & M Program has summarized vital signs (See Appendix 1) for the region, including CATO. Next steps include approval from SAC and BOD. Protocol development is already underway.
- Total budget for I & M projects involving CATO: $1,022,798.00 (FY01-FY04)

Appendix 1. Summary of priority vital signs for the NCRN (Catoctin Mountain Park highlighted).

Level 1 Name	Level 2 Name	Level 3 Name	Network Vital Sign Name	VitalSignMeasures											
Air and Climate	Air Quality	Ozone	Ozone*	Atmospheric ozone concentration	X	X	X	X	X	X	X	X	X	X	X
		Wet and dry deposition	Wet and dry deposition*	Wet deposition chemistry (pH, NO3-, SO4=), NADP	X	X	X	X	X	X	X	X	X	X	X
		Visibility and particulate matter	Visibility and particulate matter*	Visibility (PM 2.5 mass fraction)	X	X	X	X	X	X	X	X	X	X	X
		Air contaminants	Mercury deposition*	Mercury Deposition Network	X	X	X	X	X	X	X	X	X	X	X
	Weather and Climate	Weather and Climate	Weather*	ambient temperature, precipitation, wind speed, wind direction	X	X	X	X	X	X	X	X	X	X	X
Geology and Soils	Geomorphology	Coastal / oceanographic features and processes	Shoreline features	Rate of shoreline change			X	X				X			
		Stream / river channel characteristics	Physical Habitat Index (PHI)	Stream habitat structure, river depth, vegetation composition on adjacent lands	X	X	X	X	X	X	X	X	X	X	X
Water	Hydrology	Surface water dynamics	Surface water dynamics	Discharge (CFS), guage/stage height, lake elevation	X	X	X	X	X	X	X	X	X	X	X
	Water Quality	Water chemistry	Water chemistry	core parameters (pH, DO, conductance, temp), cations (Ca, Mg, Na, K), anions (PO4, NO2, Br, SO4, Cl, acid neutralizing capacity), turbidity, suspended sediments, BOD, COD, alkalinity, Secchi disc	X	X	X	X	X	X	X	X	X	X	X
		Nutrient dynamics	Nutrient dynamics	Nitrate, Amonia, DON, Nitrite, Orthophosphate, Total K	X	X	X	X	X	X	X	X	X	X	X

Level 1 Name	Level 2 Name	Level 3 Name	Network Vital Sign Name	VitalSignMeasures											
		Aquatic macroinvertebrates and algae	Aquatic macroinvertebrates	Species composition and abundance	X	X	X	X	X	X	X	X	X	X	X
Biological Integrity	Invasive Species	Invasive/Exotic plants	Invasive/exotic Plants	Early warning predictive models of incipient invasions	X	X	X	X	X	X	X	X	X	X	X
	Infestations and Disease	Insect pests	Gypsy Moths*	Gypsy moth egg mass counts	X	X	X	X	X	X	X	X	X	X	X
	Focal Species or Communities	Forest vegetation	Forest and grassland vegetation	Species diversity, age and size classes, woody debris	X	X	X	X	X	X	X	X	X	X	X
		Fishes	Fish Index of Biological Integrity (FIBI)	Spp. Compostion and abundance	X	X	X	X	X	X	X	X	X	X	X
		Amphibians and Reptiles	Amphibians	Spp. Composition and abundance, malformations	X	X	X	X	X	X	X	X	X	X	X
		Birds	Landbirds	Spp. Composition and abundance	X	X	X	X	X	X	X	X	X	X	X
		Mammals	White-tailed Deer	Deer density	X	X	X	X	X	X	X	X	X	X	X
	At-risk Biota	T&E species and communities	T&E species and communities	Spp. Abundance, presence and absence, status, threats	X	X	X	X	X	X	X	X	X	X	X
Ecosystem Pattern and Processes	Land Cover / Land Use	Land cover / Land use	Land cover/Land use	Area of dominant land cover types, connectivity, core/edge ratio of dominant forest communities, are weighted average patch size, adjacency matrix	X	X	X	X	X	X	X	X	X	X	X
	Productivity	Productivity	Landscape condition	Landuse intensity, disturbance status	X	X	X	X	X	X	X	X	X	X	X

* Already being monitored

Appendix L
Print Media Storage
Standard Operating Procedure

Print Media Storage
Standard Operating Procedure

Revision Log

Revision	Description of Change	Author	Effective Date
1	Original	G. Sanders	Oct. 1, 2005

Purpose

The purpose of this standard operating procedure (SOP) is to present the guidelines and overall process for tracking and storing analog data in the National Capital Region Network.

Scope and Applicability

The standards described in this SOP specifically pertain to all analog data related to natural resources inventory and monitoring. Analog data refers to all hard copy documents including paper files and photographs. This includes annual reports, final reports, meeting notes, photo vouchers, and scientific literature developed relating to the NCRN parks. All documents meeting these criteria will be stored either permanently or temporarily as dictated by the Records Management Handbook NPS-19, Appendix B. Topics in the Handbook related to the I&M Program are generally listed under Code N. Analog data do not include specimen collections or digital data.

Reference Documents

Records Management Handbook NPS-19, Appendix B. Document available online at: http://data2.itc.nps.gov/npspolicy/DOrders.cfm

NatureBib. Download desktop version, access online version and view documentation including help files. Available online at: http://www.nature.nps.gov/nrbib/index.htm.

Dataset Catalog. Document available online at: http://science.nature.nps.gov/im/apps/datacat/index.htm

SOP Digital Data Storage Specifications.

SOP Photo Records Specifications.

Print Media Storage
Standard Operating Procedure

Flow Chart

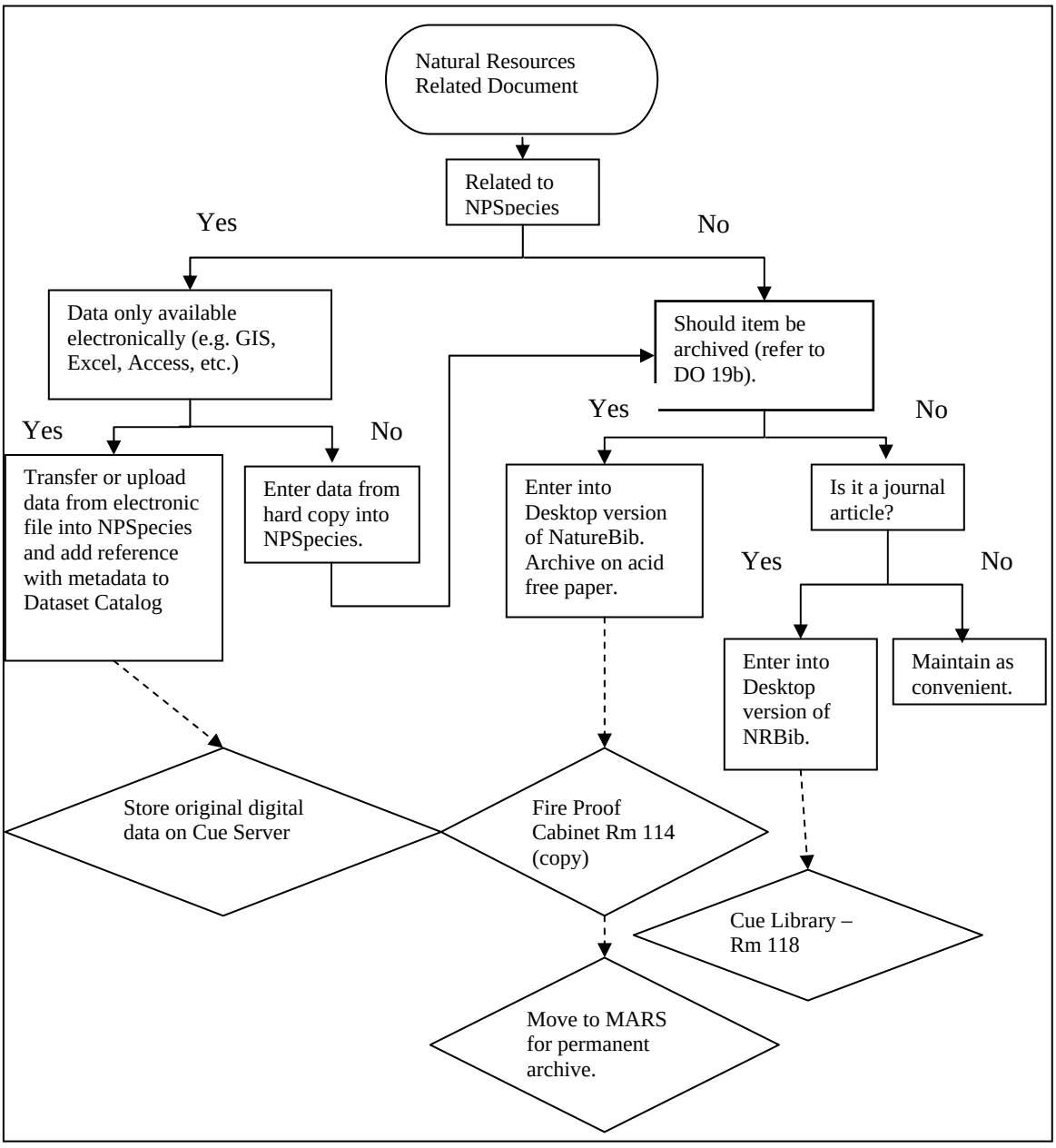


Figure 1. Flow chart for tracking hard copy documents.

Print Media Storage
Standard Operating Procedure

Procedures and General Requirements

The NPS Inventory and Monitoring (I&M) Program must manage a variety of documents including memos, meeting notes, cooperative agreements, and annual reports. All documents must be managed according to the guidelines established by the Records Management Program. Records are generally defined as those files that document NPS policies, activities, or other transactions; the creator or recipient acted in official capacity during creation or receipt of material; the subject matter related to an authorized activity of the agency, and made it available to others. All documents meeting these criteria will be managed as dictated by the Records Management Handbook NPS-19, Appendix B. Topics related to the I&M Program are generally listed under Code N. Among these records are documents that can be stored temporarily or ones that must be stored permanently.

Temporary Records: The I&M program will maintain working documents for an unspecified period of time. The working documents include documents related to the monitoring planning process but that do not meet the strict definition of permanent records (see below). These temporary records will be maintained in file cabinets located in the I&M office (Room 118) at the Center for Urban Ecology. In addition to working documents, I&M will store scientific literature pertaining to the parks and the resources in the region, including those documents cited in NPSpecies. These will be stored in a file cabinet in the I&M Office at CUE.

Permanent Records: Records that require permanent archiving according NPS-19 Appendix B will be coordinated with the National Capital Region's Museum Resource Center (MARS). MARS provides temperature and humidity controlled facilities suitable for permanent archiving and tracks its archives through ANCS+. MARS meets all NPS archiving guidelines. All documents requiring archiving will be delivered to MARS on archival acid free paper. MARS assigns an ANCS code and permanently stores documents.

Examples of documents that will require permanent archiving include:

- Annual and Final reports generated by either the I&M program or cooperators/contractors.
- Completed Inventory Study Plan, Monitoring Plan and the Data Management Plan
- Original data sheets
- Slides and photographs
- Contracts or Cooperative Agreements

Prior to archiving printed materials at MARS, permanent records will be maintained at CUE in Fire Proof Filing Cabinets. Copies of all project final reports will also be provided to the parks for archiving on site. NCRN IM Program, however, is not responsible for archiving functions in the parks.

Storing Temporary and Permanent Documents

All paper documents managed by the I&M program will be housed one of the following locations:

I&M Library

File cabinet housed in Office 118 at CUE. Contains scientific documents and reports that are cited in NPSpecies. Also contains scientific articles pertaining to the resources in the region.

Files in the CUE Library are stored by a code plus a sequential number. The ID is written on the top right corner of a document and filed in alphabetical and numerical order.

Code	Description
Park CODE	The four letter park code is used for any document pertaining to a specific park. Park codes include: ANTI, APPA, CATO, CHOH, GWMP, HAFE, MANA, MONO, NAMA, NACE, PRWI, ROCR, and WOTR
MULTI	Used for any document pertaining to more than one park
OTHER	Pertains to general science documents related to inventory and monitoring

Fire Proof Cabinets

These cabinets are housed in Office 115 at CUE. They contains temporary files following the guidelines presented in NPS-19, Appendix B. The cabinets also contain copies of permanent records that have been archived at MARS. Contents of the cabinets are filed alphabetically by topic (table 1). The cabinets also contain original field data sheets from I&M projects.

Electronic Copies of Paper Documents: In addition to maintaining paper records, the I&M Program will maintain electronic files whenever possible. Archived electronic files will be maintained on the CUE Server in the Archive directory. This directory is formatted as "READ ONLY" with only the NCRN data manager and the regional IT staff having access "WRITE" access. The Archive directory contains subfolders for each project the NCRN has been involved with or is presently involved where digitally archived data is stored.

Photo Records: Photos taken by the I&M Program include slides, prints, and digital photos. All hard copy photos will be filed in the Fire Proof Cabinet according to table 1. Digital photos will be filed in the PHOTOS directory on the file server. All photos will be tracked using photo management software. The NCRN is currently evaluating a number of difference packages to determine what will work best for the Network's needs. See Photo Management SOP for more details.

Tracking Paper and Electronic Copies of Temporary and Permanent Archives

All documents and their electronic copies will be tracked using reference management software. In the past the network used the desktop version of NatureBib to track references. The NCRN is currently searching for an alternative software package or application that will improve our ability to

track and disseminate reference materials. Until a replacement is found, the NCRN will continue to use NatureBib to track documents. A brief description of the desktop version of NatureBib follows:

NatureBib: NatureBib is an NPS database developed to list documents including reports and publications that are pertinent to the National Parks. The database is available in both an online version that is linked to NPSpecies, and as a stand-alone desktop version. The desktop version of NatureBib follows the structure of the online version. It will be used to track all reference material, records, and photographs housed at the Center for Urban Ecology. The file location can be entered by call number or physical location into the "location" data field. Support and information about the database including tutorial versions is available online (http://www.nature.nps.gov/nrbib/index.htm). See also NRBib Help Files (http://www.nature.nps.gov/nrbib/helpfile.htm) and Data Dictionary for details.

Table 1. Major filing categories and their content descriptions for documents stored in the I&M Fire Cabinet (Office 115 at CUE).

Category	Description
AAWRP	Each year's AAWRP
Abstracts	All abstracts submitted for presentations or posters
Cooperative Agreements	Original Cooperative agreement or task agreement including scope of work
Inventory	All annual reports submitted by cooperators
Fact Sheet	All I&M Fact sheets developed by NCR I&M Program and WASO
Meetings (BOD)	All agendas and minutes
Meetings (Project Planning)	Agenda and minutes; includes updated Project Plans
Meetings (SAC)	Agenda and minutes; includes notes from I&M Steering Committee
Meetings (Stakeholder Analysis)	Agenda and minutes
Memos	All WASO Memos for I&M
Monitoring Workshop	Agenda and Design, and Report
NCR Status Update	Final NCR Status Report by TNC and NPS 1999
Newsletters (Online)	One Minute Updates
Newsletters (Quarterly)	The Monitor
Proposals	All funded proposals by I&M program
RFP (Inventories)	All RFPs developed and advertised
RFP (Monitoring)	All RFPs developed and advertised

Print Media Storage
Standard Operating Procedure

Responsibilities

Project Leader or Data Manager

Be aware of all reports and documents that are being generated and deliver to Data Manager for tracking and storage. Determine which records are permanent and temporary. Follow figure 1 and enter all appropriate data into NPSpecies, NatureBib, and/or Dataset Catalog. Store document accordingly.

Museum Specialist

Handle storage of permanent files at MARS. Accession permanent files into museum collection and tracks with ANCS+. Provide the necessary catalog and accession numbers.