Southern Colorado Plateau
Inventory and Monitoring Network

National Park Service
U.S. Department of the Interior

Natural Resource Program Center
Fort Collins, Colorado

Data Management Plan for the Southern Colorado Plateau Network

Nicole J. Tancreto and Marguerite N. Hendrie

National Park Service, Southern Colorado Plateau Network,
Northern Arizona University, P.O. Box 5765, Flagstaff, AZ 86011

October 2006

Data Management Plan for the Southern Colorado Plateau Network

Nicole J. Tancreto and Marguerite N. Hendrie

National Park Service, Southern Colorado Plateau Network,
Northern Arizona University, P.O. Box 5765, Flagstaff, AZ 86011

October 2006

Recommended Citation: N.J. Tancreto and M.N. Hendrie. 2006. Data management plan for the Southern Colorado Plateau Network. National Park Service, Flagstaff, AZ. 77 p. plus appendices.

U.S. Department of the Interior
National Park Service, Southern Colorado Plateau Network
Northern Arizona University
P.O. Box 5765
Flagstaff, AZ 86011

Acknowledgements

This plan was written with considerable help, inspiration and ideas gleaned from data management plans written and developed by other data managers throughout the NPS Inventory & Monitoring program.

EXECUTIVE SUMMARY **1**

CHAPTER 1 INTRODUCTION **6**

1.1 Overview and Scope 6
1.2 Plan Revisions 6
1.3 Goals and Objectives 7
1.4 Network Organization and Management 7
1.5 Sources of Natural Resource Data 9

CHAPTER 2 ROLES AND RESPONSIBILITIES **10**

2.1 Data Management Roles 11
2.2 Personnel Resources and Responsibilities 12
2.3 Data Management Coordination 14

CHAPTER 3 INFRASTRUCTURE & SYSTEM ARCHITECTURE **16**

3.1 Infrastructure 16
3.2 SCPN System Architecture 17
3.3 LAN and WAN Infrastructure 19
3.4 IT Network Security 19

CHAPTER 4 DATA MANAGEMENT PROCESS AND WORKFLOW **21**

4.1 Types of Data and Information 21
4.2 Data Management Framework 21
4.3 Phases of Data Processing 23
4.4 Water Quality Data 24

CHAPTER 5 DATA DESIGN **25**

5.1 Data Modeling 25
5.2 Database Standards 25
5.3 Data Sheet Design 26
5.4 Database Application Design 26

CHAPTER 6 DATA ACQUISITION AND VERIFICATION **28**

6.1 Network Data 28
6.2 NPS Data 30
6.3 External data 32

CHAPTER 7 DATA VALIDATION, SUMMARY AND EXPORT FOR ANALYSIS **33**

7.1 Data Validation 33
7.2 Data Summaries 33
7.3 Export for Analysis 34

CHAPTER 8 DOCUMENTATION **35**

8.1 Metadata 35
8.2 Data Quality Documentation 36
8.3 Database Application Documentation 36

CHAPTER 9 DATA OWNERSHIP AND SENSITIVITY **38**

9.1 Data Ownership 38
9.2 Data Sensitivity 39

CHAPTER 10 DISSEMINATION AND ACCESS **41**

10.1 Distribution Mechanisms 41
10.2 NPS Access and Products 42
10.3 Public Access and Products 42

10.4 Tracking and Feedback 43

CHAPTER 11 ARCHIVING AND STORAGE OF DIGITAL PRODUCTS 44

11.1 NPS and Network Standards 44
11.2 Electronic Archiving and Storage 44
11.3 Electronic Archiving Process and Workflow 44
11.4 Quality Control for Converted Data 44
11.5 Digital Data Backups 45

CHAPTER 12 ARCHIVING AND STORAGE OF PHYSICAL PRODUCTS 46

12.1 NPS Standards 46
12.2 SCPN Standards 46
12.3 Archiving Process 46
12.4 Network Working Files and Library 47

LIST OF TABLES

Table 1-1. Establishment Purpose and Size of Southern Colorado Plateau Network Park Units. 8
Table 1-2. SCPN staff, data management role and position type. 9
Table 2-1. Data management tasks and the positions with principal responsibility. 10
Table 2-2. Roles and responsibilities for data stewardship 11
Table 3-1. SCPN infrastructure service or support providers. 16
Table 3-2. General directory structure and their roles for Master Libraries. 19
Table 4-1. Categories of data 21

LIST OF FIGURES

Figure 1-1. Overview of Southern Colorado Plateau Network park unit locations 7
Figure 2-1. Data stewardship responsibilities of the core data management team. 12
Figure 3-1. SCPN local, and national infrastructure. 17
Figure 3-2. Directory organization of working files. 18
Figure 4-1. Data management activities within the context of a monitoring project. 22

APPENDICES 48

Appendix A. Data Management Plan Revision Log 49
Appendix B. Backup Procedures 50
Appendix C. Operating Procedures for Uploading and Downloading Data to and from GPS Units (draft) 55
Appendix D. GIS Data Organization and Standard Operating Procedures for Data and Metadata Creation (draft) 59
Appendix E. Digital Photo Management (draft) 62
Appendix F. File Naming Conventions (to be developed) 67
Appendix G. Data Modeling 68
Appendix H. Standard Data Summary Reports (to be developed) 71
Appendix I. Exporting Data for Analysis (to be developed) 72
Appendix J. Metadata Creation for Non-spatial Data Sets 73

Executive Summary

Data become information through the process of analysis, synthesis, modeling, or other types of interpretation. Data management provides a means for organizing, documenting and archiving data so that the original information potential is maintained through time. This is particularly important for long-term programs where the lifespan of a data set will likely be longer than the careers of those who developed it. A data management system that can effectively produce, maintain and distribute monitoring results is central to the success of the Inventory & Monitoring (I&M) Program.

This data management plan outlines the general data management standards; expected roles and responsibilities; and data processing, storage and distribution guidelines for the Southern Colorado Plateau Network (SCPN). Detailed data management procedures for specific SCPN monitoring projects will be based on these general guidelines.

Goals and Objectives

The SCPN approach to data management is user oriented, the user ranging from SCPN staff to park managers. The primary goal of the SCPN data management program is to ensure the quality, clarity, security, longevity and availability of SCPN I&M data.

Quality – SCPN I&M data will be used by park staff to inform management decisions regarding park natural resources; it is essential that these data be accurate and complete. Appropriate quality assurance measures will be employed throughout the process of collecting, processing and maintaining data. Good data stewardship habits and attitudes will be encouraged.

Clarity – Confusing and cryptic data sets are of little use and can be easily misinterpreted. All data and information products will be accompanied by complete documentation so that users will be aware of the applicability and limitations of the data.

Security – All information products will be maintained so that appropriate levels of access are provided to SCPN and park staff. Existing technologies will be utilized to protect I&M data from corruption or loss, ensuring the long-term security and integrity of the data.

Longevity – Many factors combine to increase the longevity of a data set: proper documentation, organization, and standardization to modern technologies. SCPN will ensure that all data sets are completely documented. Data sets will be organized in a logical and consistent manner so that nothing is lost over time. As software and hardware technology changes, data sets will be updated so that they remain readable and accessible using new technology.

Availability – I&M data can only be useful to park managers if it is easily available in a timely manner and in a useful form. Information products will be distributed to park management on a regular schedule and, when appropriate, will be made available to a broader audience.

Sources of Natural Resource Data

The existence of numerous potential sources of ecological data about park natural resources requires SCPN to make some priorities about managing data. Some sources of natural resource data include:

- Inventories
- Monitoring
- Protocol development studies
- Special focus studies completed by parks
- External research projects
- Studies by other land management agencies on adjacent lands
- Resource impact evaluations related to park planning and compliance regulations
- Resource management and restoration work

SCPN will be able to maintain the highest level of control over data collected through the I&M Program and the data management program will focus on these data. However, the goal is to apply the same standards, procedures and attitudes about data management to other sources of natural resource data over the long-term, working toward raising the level of data management for projects originating outside the I&M Program.

Roles and Responsibilities

Data management and stewardship is the responsibility of all participants in SCPN Network I&M activities; it requires true collaboration among many people with a broad range of tasks and responsibilities. Good habits and attitudes are as important as standards and procedures. Although primary responsibility resides with the data manager, GIS specialist, and project managers who make up the core data management team, all SCPN staff and cooperators are responsible for ensuring data stewardship is practiced throughout the life of a monitoring project.

Infrastructure and System Architecture

Management and dissemination of monitoring data is made possible by information technology infrastructure and system architecture. Infrastructure refers to the network of computers and servers that information systems are built upon. System architecture refers to the application, database system, repositories, and software tools that make up the framework of the SCPN's data management enterprise.

SCPN relies on cooperative agreements with Northern Arizona University (NAU) and the USGS Colorado Plateau Research Station (CPRS) as well as NPS regional and national information technology personnel and resources for maintenance and support of computer and networking infrastructure.

SCPN System Architecture

Working files, master libraries and digital archives will be stored on SCPN file and data servers. A template project directory for databases, files, and project documentation will be used for each monitoring project. This directory will contain working files for which all project team members will have read/write access. Master libraries are repositories for final information products and certified databases. Master libraries will be organized according to type – databases will be stored together, documents will be stored together, etc. – and only certain SCPN personnel will have full access. Digital archives will be repositories for packages of data and information that can easily be redistributed and will have extremely limited write access.

Rather than developing a single integrated database system, SCPN will develop stand-alone project databases that share design standards and centralized lookup tables for data shared across projects. These modular databases allow for greater flexibility to accommodate each project's needs and sufficient standardization can ensure the ability to aggregate and summarize data across multiple projects. SCPN currently uses Microsoft Access for all project databases and is investigating the need to move to a client-server relational database management system such as Microsoft SQL Server.

National System Architecture

The national I&M Program provides several repositories for hosting SCPN information products and applications for summarizing park data at a national level. The applications are available online and allow users to access basic natural resource information for SCPN parks:

- *NatureBib* – master database for natural resource bibliographic references
- *NPSpecies* – master database for species occurrence records and evidence (voucher specimens, references, observations or data sets) at each park
- *NR-GIS Metadata and Data Store* – master database of metadata for GIS and natural resource data sets and a repository for that data

Data Management Process and Workflow

Within the context of a monitoring project, SCPN data management tasks can be divided into several types of activities. These activities provide the backbone for planning and executing data management procedures for each monitoring project. Specific procedures and guidelines for each of these activities are explained in the data management plan. Data design refers to the design and development of project data sheets, the database and database application. During data acquisition, data are collected in the field or acquired from other sources, entered into the project database, and verified. Data will then be validated to ensure they are within normal ranges, summarized and exported for analysis. Documentation will be completed, data and information products will be distributed, and both digital and analog products will be archived.

Data and documentation take different forms and are maintained in different places throughout the phases of a project. These phases can be modeled as a sequence of events and tasks which involve interaction with the following items:

- *Raw data* – Analog data recorded by hand on field data sheets 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 time period). This may be the only database for short-term projects with no need to distinguish current season data from the full set of validated 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 manager that the data have passed all quality assurance requirements and are complete and ready for distribution. Metadata records include detailed information about project data needed for proper use and interpretation.
- *Master database* – Project-specific database for storing the full set of validated project data, used for viewing, summarizing, and analysis. Current season data from the working database must pass all quality assurance steps prior to upload into this master project database.
- *Reports and data products* – Information that is derived from certified project data.
- *Edit log* – A means of tracking changes to certified data.
- *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.
- *Archived data* – Digital and hard-copy data stored and maintained for the long-term

Water Quality Data

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; http://www.nature.nps.gov/water/infoanddata/index.cfm). NPS WRD has selected the Environmental Protection Agency's national STOrage and RETrieval water quality database (STORET) as the standard for archiving NPS water quality data. WRD maintains a copy of STORET and requires that all I&M chemical, physical and biological water quality data be archived in this copy. SCPN will use the Microsoft Access database application (called NPSTORET) developed by WRD to fulfill this requirement.

The USGS/WRD-Colorado developed a water quality database that enables assessment of the temporal and spatial distribution characteristics of water quality data available for the 19 SCPN park units. For the purpose of developing the Water Quality Vital Signs Monitoring Plan, this database provides a useful tool for evaluation of historical water quality of the waters in and surrounding SCPN park units, and determination of historical and current water quality conditions in and near these parks. SCPN plans to maintain and regularly update the water quality database through data retrievals from NPSTORET or STORET, enabling it to serve as a dynamic tool for the network's long-term water quality monitoring and analysis needs.

Data Design

The data manager and project manager will collaborate on design of field data sheets, database structure and database application for each monitoring project. Databases will be standardized where possible following I&M recommended guidelines for database structure and naming conventions developed in the Natural Resource Database Template (NRDT) and the Recommended Naming Standards. SCPN will also develop standardized lookup tables for data elements shared across many monitoring projects. Database design will also be guided by a data modeling process involving the creation of three types of data models: conceptual, logical and physical.

Data Acquisition and Quality Control

Data managed and utilized by the network will originate from three types of sources: within the network, other NPS data collection efforts, and outside the NPS altogether.

- **Network Data** – any data produced from projects that are initiated (funded) by the SCPN I&M Program or projects that in some way involve the I&M Program.
- **NPS Data** – any data produced by the NPS that did not involve the SCPN I&M Program.
- **External Data** – any data produced by agencies or institutions other than the National Park Service.

SCPN staff are responsible for the acquisition and quality control of network data. Project crew leaders and members are primarily responsible for data collection, data entry and verification of data acquired from field data collection. Each monitoring project protocol will detail procedures for these data acquisition steps based on guidelines outlined in this plan. As data are collected and entered into a database, quality control procedures will be used to increase accuracy and limit transcription mistakes. A verification procedure will be used to check for and correct any transcription mistakes.

NPS and external data will be acquired only with complete documentation and metadata. These data will undergo limited processing and quality control by SCPN staff to ensure compatibility with SCPN project databases, if necessary. Legacy data from parks will be evaluated and prioritized for digitizing or converting to modern database formats. Some external ancillary data, such as climate data, will be acquired when needed in subsets rather than stored by SCPN in its entirety.

Quality Assurance, Data Summarization, and Export for Analysis

Quality assurance, data summary and data analysis are the responsibility of the project managers; however, the data manager will provide tools to project managers to facilitate these three activities. Data validation procedures (ensuring measures are logical and within normal ranges) will be detailed in each monitoring protocol and will generally include outlier detection and other exploratory analyses.

Routine data summaries will be produced after data have been verified on a schedule specific to each project. Summaries will generally be automated within the database application, but park-specific data reports can be produced for management needs. Automated exports will also be included in each database application to enable project managers to export subsets of data in a format ready for import into specific statistics or other analytic software programs.

Documentation

Data set documentation is the responsibility of the project manager and data manager. All data sets will be documented with formal metadata, using Federal Geographic Data Committee and USGS National Biological Information Infrastructure standards. Documentation accompanying database applications will include a manual with instructions for using the application, an entity relationship diagram, a data dictionary and programming code documentation.

Access and Archiving

Data Ownership and Sensitivity

SCPN data products are owned by the National Park Service provided under OMB Circular A-110, Section 36. The Freedom of Information Act (FOIA) establishes that the federal government, the NPS

included, must provide access to non-protected data and information of interest to the public through reading rooms or the Internet.

The NPS is directed to protect information about the nature and location of sensitive park resources under one Executive Order and four resource confidentiality laws:

- Executive Order No. 13007: Indian Sacred Sites
- National Parks Omnibus Management Act (NPOMA; 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)

All monitoring information products will be vetted for sensitive data prior to making them available to the general public. Classification of sensitive I&M data will be a shared responsibility that includes network staff, park resource management staff, park superintendents, and investigators working on individual projects. Park management has ultimate responsibility for deciding which information is sensitive and should not be released to the public. The network has ultimate responsibility for ensuring that sensitive data are not released to the public.

Dissemination and Access

Dissemination of monitoring and information products from SCPN will follow these guidelines:

- data will be easily located and acquired
- only data subjected to full quality control and quality assurance measures will be released
- data will be accompanied by complete metadata
- sensitive data will be identified and protected from unauthorized access

Information products will be made available primarily through websites and clearinghouses which will allow users to search for and download reports, summarized data, maps and metadata and other associated information. Distribution means will include (but may not be limited to):

- SCPN public website
- NR-GIS Metadata and Data Store
- Service-wide databases, such as NPSTORET, NPSpecies, and NatureBib
- Regional, Network, or Park data servers protected with read-only access
- FTP sites, CDs, DVDs, or hard drives, as appropriate

Archiving and Storage

Digital and analog information products will be stored, archived and maintained in a variety of repositories. Digital products resulting from monitoring projects will be archived on SCPN file servers and national file and data servers and protected from catastrophic loss by regular, automated backups to external media. Analog products will be archived to NPS standards by individual park facilities or approved non-NPS institutions. At the termination of a project or at regular milestones, an archival package will be prepared and delivered to the desired location.

Chapter 1 Introduction

The Inventory and Monitoring Program (I&M) 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 a framework for the I&M Program, which fully integrates natural resource monitoring and other scientific activities into the management processes of the National Park system.

A distinction is usually made between the term data, which refers to assemblages of raw or un-interpreted facts, records, or observations; and information, which is created from data through the process of analysis, synthesis, modeling, or other types of interpretation. Data management provides a means for organizing, documenting and archiving data so that the original information potential is maintained through time. This is particularly important for long-term programs where the lifespan of a data set will likely be longer than the careers of those who developed it. A data management system that can effectively produce, maintain and distribute monitoring results is central to the success of the I&M Program.

This data management plan outlines general data management standards; expected roles and responsibilities; and data processing, storage and distribution guidelines for the Southern Colorado Plateau Network (SCPN). Detailed data management procedures for specific SCPN monitoring projects will be based on these general guidelines.

1.1 Overview and Scope

This data management plan establishes general data management standards and guidelines for the SCPN. The primary audience of the plan is network staff, and it is the document to which staff and cooperators can refer for guidelines and standards relating to data acquisition, quality control & assurance, documentation, storage, maintenance, and distribution. The main focus is on the SCPN monitoring program, including: collection, processing and maintenance of raw data; acquisition of external data; and the dissemination of data and information to park management, the scientific community, and the general public. This plan will not cover project-specific aspects of data analysis and summarization or park-specific data management needs. The Data Management Plan will inform all SCPN staff of general data management standards, expected roles and responsibilities, and data processing, storage and distribution guidelines.

The main body of this document will provide general guidance and standards for data management tasks and activities. SCPN will develop specific data management procedures and protocols for each individual monitoring project. Commonly used data processing or management procedures will be specified in appendices to this plan, facilitating revision as procedures evolve with technology and time.

1.2 Plan Revisions

SCPN will update this plan regularly as the network's data management standards and practices change with time. Appendices may be added or updated more often as they outline detailed instructions and procedures. The main body of the plan will undergo a full review and revision on a minimum schedule of every five years.

Document versions will be designated by using version numbers in parentheses that are appended to the heading of the modified section. For example, the first version of Chapter 4 will have the heading of “Chapter 4.” A revision to this chapter will have the heading “Chapter 4. (v2).” A section of this chapter would be “Chapter 4. 1 (v4)” for the fourth version of Chapter 4 section 1. Appendix A of the plan (Data Management Plan revision history log) contains a log that will identify Data Management Plan versions and summarize changes. Previous versions will be archived in their entirety for reference. The latest version of the plan will be available on the SCPN website and will include the revision log as an attachment.

1.3 Goals and Objectives

The SCPN approach to data management is user oriented, possible users being park managers, SCPN staff, other agency partners, researchers and the general public. The primary goal of the SCPN data management program is to ensure the quality, clarity, security, longevity and availability of SCPN I&M data.

Quality – SCPN I&M data will be used by park staff to inform management decisions regarding park natural resources; it is essential that these data be accurate and complete. Appropriate quality assurance measures will be employed throughout the process of collecting, processing and maintaining data. Good data stewardship habits and attitudes will be encouraged.

Clarity – Confusing and cryptic data sets are of little use and can be easily misinterpreted. All data and information products will be accompanied by complete documentation so that users will be aware of the applicability and limitations of the data.

Security – All information products will be maintained so that appropriate levels of access are provided to SCPN and park staff. Existing technologies will be utilized to protect I&M data from corruption, ensuring the long-term security and integrity of the data.

Longevity – Many factors combine to increase the longevity of a data set: proper documentation, organization, and standardization to modern technologies. SCPN will ensure that all data sets are completely documented. Data sets will be organized in a logical and consistent manner so that nothing is lost over time. As software and hardware technology changes, data sets will be updated so that they remain readable and accessible using new technology.

Availability – I&M data can only be useful to park managers if it is easily available in a timely manner and in a useful form. Information products will be distributed to park management on a regular schedule and made broadly available.

1.4 Network Organization and Management

The SCPN is composed of 19 parks located in northern Arizona, northwestern New Mexico, southwestern Colorado and southern Utah (Figure 1-1).

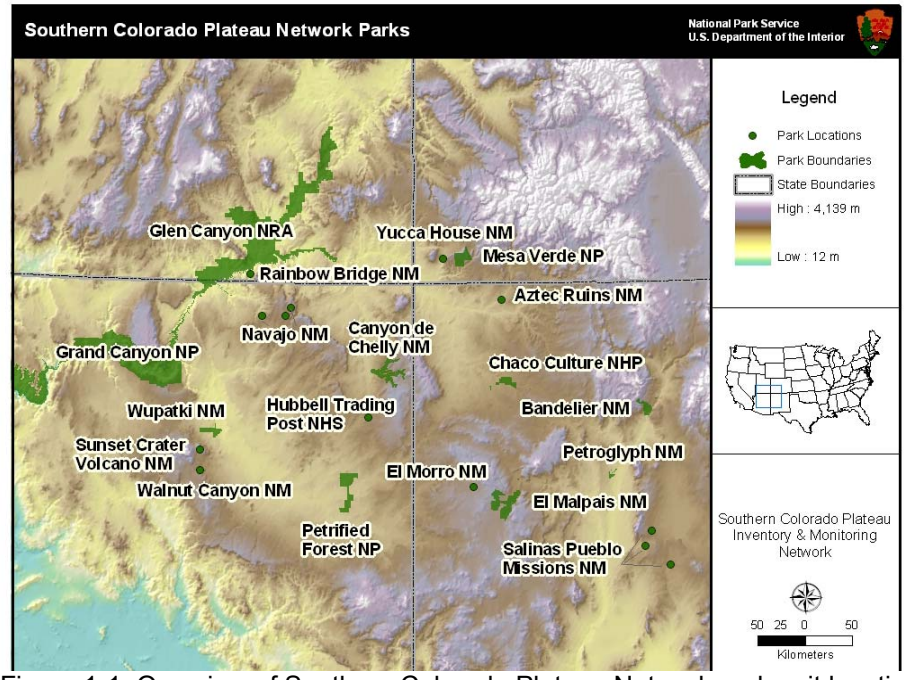


Figure 1-1. Overview of Southern Colorado Plateau Network park unit locations

The parks range in size from 14 to more than 500,000 hectares (Table 1-1), with more than 750,000 hectares within the network designated or proposed as wilderness. The majority of the SCPN parks were designated primarily to protect cultural resources.

Table 1-1. Establishment purpose and size of Southern Colorado Plateau Network park units.

				Originally Established For	
Park	**Abbreviation**	**State**	**Hectares**	**Cultural Resources**	**Natural Resources**
Aztec Ruins National Monument	AZRU	NM	130	X	
Bandelier National Monument	BAND	NM	13,367	X	
Canyon De Chelly National Monument	CACH	AZ	37,448	X	
Chaco Culture National Historical Park	CHCU	NM	13,929	X	
El Malpais National Monument	ELMA	NM	47,352	X	X
El Morro National Monument	ELMO	NM	420	X	
Glen Canyon National Recreation Area	GLCA	AZ/UT	505,909	*	*
Grand Canyon National Park	GRCA	AZ	488,551	X	X
Hubbell Trading Post National Historic Site	HUTR	AZ	65	X	
Mesa Verde National Park	MEVE	CO	21,546	X	
Navajo National Monument	NAVA	AZ	243	X	
Petrified Forest National Park	PEFO	AZ	37,852**		X
Petroglyph National Monument	PETR	NM	2,923	X	
Rainbow Bridge National Monument	RABR	UT	66	X	X
Salinas Pueblo Missions National Monument	SAPU	NM	432	X	
Sunset Crater Volcano National Monument	SUCR	AZ	1,230		X
Walnut Canyon National Monument	WACA	AZ	1,465	X	X
Wupatki National Monument	WUPA	AZ	14,388	X	
Yucca House National Monument	YUHO	CO	13	X	

* Glen Canyon NRA was established to "…provide for public outdoor recreation use of Lake Powell…" and to "…preserve the scenic, scientific, and historic features…of the area".

** Recently approved boundary addition to Petrified Forest NP will bring the total area to 88,439 ha pending additional funding.

SCPN's offices are located on the campus of Northern Arizona University (NAU) in Flagstaff, Arizona, and are co-located with the Colorado Plateau Cooperative Ecosystems Studies Unit (CPCESU) and the USGS Colorado Plateau Research Station (CPRS).

1.4.1 Network Staff

A cooperative relationship with NAU through the CPCESU allows SCPN to meet recurring needs for discipline-specific monitoring crew leaders and members through NAU agreements. All SCPN staff, whether NPS or NAU, will be directly involved with data management in some form (see Chapter 2 for more details). The operational staffing planned for SCPN for the next five years is shown in Table 1-2.

Table 1-2. Current and projected SCPN staff, data management role and position type.

Position	Data Management Role	Position Type
Program Manager	Network Program Manager	NPS Permanent
Hydrologist	Project Manager	NPS Permanent
Vegetation Ecologist	Project Manager	NPS Permanent
Quantitative Ecologist	Project Manager, Statistician	NPS Permanent
GIS Specialist	Network GIS Specialist	NPS Permanent
Cartographic Technician - projected	Network GIS Specialist	NAU Temporary
Data Manager	Network Data Manager	NPS Permanent
Data Management Assistant	Network Data Manager	NPS Permanent
Botanist - projected	Project Crew Leader	NPS Permanent
Water Quality Lead - projected	Project Crew Leader	NAU Temporary
Biological/Hydrological Science Technicians (5) - projected	Project Crew Members	NAU Temporary

1.5 Sources of Natural Resource Data

The existence of numerous sources of ecological data about park natural resources requires SCPN to make some priorities about managing data. Some sources of natural resource data include:

- Inventories
- Monitoring
- Protocol development studies
- Special focus studies completed by parks
- External research projects
- Studies by other land management agencies on adjacent lands
- Resource impact evaluations related to park planning and compliance regulations
- Resource management and restoration work

SCPN will be able to maintain the highest level of control over data collected through the I&M Program and the data management program will focus on these data. However, the goal is to apply the same standards, procedures and attitudes about data management to other sources of natural resource data over the long-term, working toward raising the level of data management for projects originating outside the I&M Program.

Chapter 2 Roles and Responsibilities

Data management and stewardship is the responsibility of all participants in SCPN Network I&M activities; it requires true collaboration among many people with a broad range of tasks and responsibilities. Although primary responsibility resides with the data managers, GIS specialists, and project managers who make up the core data management team, all SCPN staff and cooperators are responsible for ensuring data stewardship is practiced throughout the life of a monitoring project. Each network staff member must understand the data management responsibilities of his or her position in the design, acquisition, QA/QC, summarization, documentation, access and archiving of data (Table 2-1). Cooperators and contractors also have responsibilities for data stewardship. This chapter discusses general data management roles and responsibilities of all network staff and cooperators.

Table 2-1. Data management tasks and the positions with principal responsibility.

Data Management Activity	Tasks	Principal Responsibility
Data Design	• Database design • Datasheet design • Data entry design • Data domains and dictionaries • Database documentation	Data Manager Project Manager Project Crew Leader Network Coordinator
Data Acquisition & Quality Control	• Train field crew in data collection protocols • Test for consistency in field measurements • Conduct & record field measurements • Enter measurements into database • Data verification (ensure data are correctly transcribed from data sheets to database)	Project Manager Project Crew Leader Project Crew Data Manager Data Management Asst. GIS Specialist
Quality Assurance, Data Summarization & Export for Analysis	• Data validation (check for missing data, out-of-range values, and logical errors) • Design and automate standard data summaries • Design and automate export routines • Produce and review data summaries	Data Manager Project Manager
Documentation & Metadata	• Database documentation • Data quality documentation • Formal FGDC/NBII compliant metadata	Data Manager Project Manager GIS Specialist
Access & Archiving	• Complete metadata • Archive data • Catalog data/reports • Distribute data/reports to park managers • Provide appropriate access to data/reports	Data Manager Data Management Asst. GIS Specialist

2.1 Data Management Roles

A *role* is a function or position (e.g., Project Manager)
A *responsibility* is a duty or obligation (e.g., review data records)

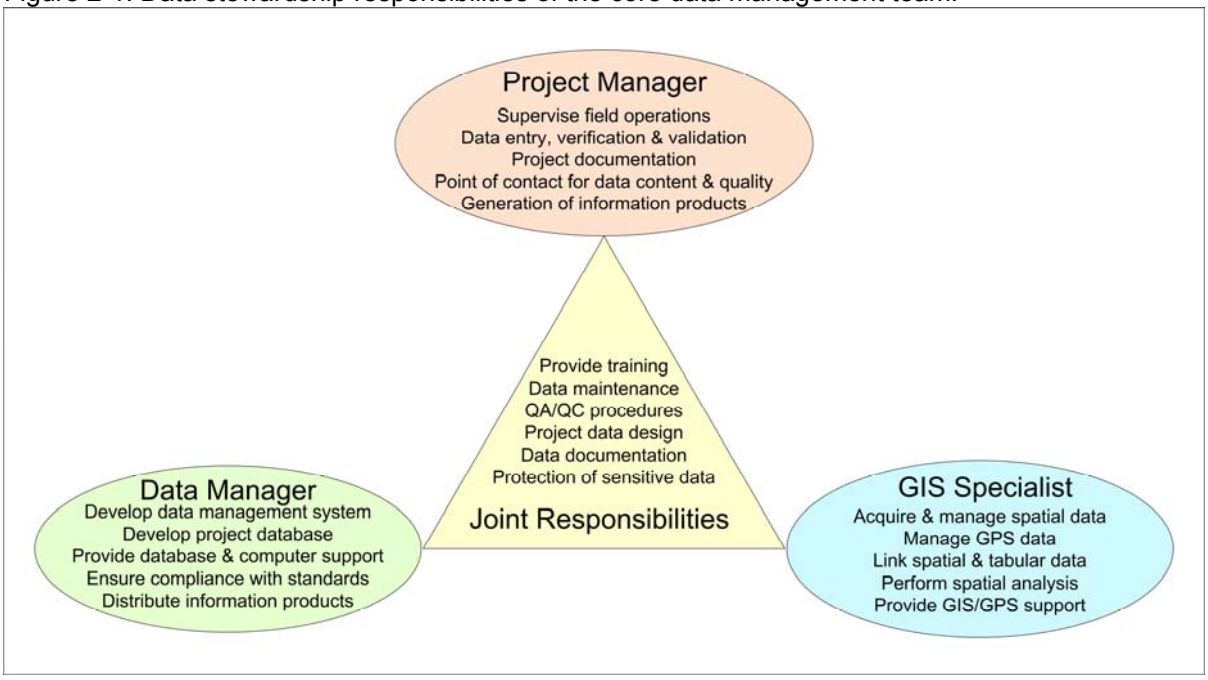
Table 2-2 summarizes the roles played by SCPN staff, cooperators, and park staff and the responsibilities of each. These roles are listed 'from the ground up' to help demonstrate the hierarchy of responsibilities. For example, a project manager is ultimately responsible for the activities listed in the field level roles of crew leader and crew member. Also, the network program manager ensures that the network data manager and ecologist achieve the required performance level.

Table 2-2. Roles and responsibilities for data stewardship

Role	Data Stewardship Responsibilities
Project Crew Member	• Collect, record, enter and verify data.
Project Crew Leader	• Lead field crew in data collection. • Organize and verify data.
Project Manager	• Supervise and train project crew in proper data collection techniques. • Validate data. • Provide data set documentation and metadata. • Perform statistical analysis and interpret results.
Network Data Manager	• Develop and support network data management system. • Ensure project data are organized, compliant, and safe. • Provide for dissemination of project data to end users.
Network GIS Specialist	• Process, manage, validate and document spatial data. • Provide spatial data to support monitoring projects. • Conduct spatial analyses. • Work with Data Manager to integrate spatial and tabular data. • Manage GPS data collection, correction and storage.
Network Quantitative Ecologist (or consulting statistician)	• Collaborate with Project Managers to analyze project data.
Network Program Manager	• Coordinate and oversee all network activities.
Information Technology Specialist	• Provide IT support for hardware, software, and networking.
Park Resource Managers and Superintendents	• Inform scope and direction of monitoring information needs. • Integrate monitoring data into park planning and management decisions.
Park Curator	• Oversee all aspects of specimen acquisition, documentation, and preservation, and manage the park collections.
Park Research Coordinator	• Facilitate data acquisition by external researchers. • Communicate NPS requirements to permit holders.
I&M Data Manager (National Level)	• Provide Service-wide database availability and support.

The core data management team will consist of the project manager, data manager and GIS specialist. Each position has core data stewardship activities for which he/she is responsible, and there are also shared responsibilities (Figure 2-1).

Figure 2-1. Data stewardship responsibilities of the core data management team.



2.2 Personnel Resources and Responsibilities

The data manager, GIS specialist and project manager share primary responsibility for network data management, but cooperators may also play a large role. It is the responsibility of network staff, primarily project managers and the data manager, to ensure that cooperators are aware of their responsibilities and network guidelines and standards.

2.2.1 Data Manager

The network data manager has primary responsibility for data management within the network. The data manager develops policies, guidelines, standards and procedures for data management. Primary responsibilities include:

- develop and maintain the data management system for project data management and metadata creation.
- create and support project databases in accordance with best practices and current standards.
- establish and implement procedures for collecting, working with, and storing digital data.
- provide training in the theory and practice of data management tailored to the needs of project personnel.
- provide appropriate access to and protection of digital data and information.
- collaborate with GIS specialists to integrate tabular data with geospatial data in a GIS system in a manner that meets project objectives.
- provide basic computer support to network staff.

Data managers will also work closely with project managers to:

- define the scope of project data and create a data design and database application that meets project needs.
- develop quality assurance aspects of project protocols and standard procedure documents.
- identify elements that can be built into the database structure to facilitate quality control, such as required fields, range limits, pick-lists and conditional validation rules.
- create a user interface that streamlines the process of data entry, review, validation, and summarization that is consistent with the capabilities of the project staff.
- develop automated database procedures to improve the data summarization and reporting process.
- make sure that project documentation is complete, complies with metadata requirements, and enhances the interpretability and longevity of the project data.
- ensure regular archiving of project materials.

2.2.2 Project Managers

The network project managers are responsible for coordination and supervision of individual projects and are involved in all aspects of project data management. This requires close interaction and coordination with the data manager and project cooperator(s). Primary responsibilities of project managers are:

- develop, document, and implement standard procedures for field data collection and data handling.
- enact and supervise quality assurance and quality control measures for the project.
- supervise and certify all field operations, including staff training, equipment calibration, species identification, and data collection. Supervise or perform data entry, verification and validation.
- maintain concise explanatory documentation of all deviations from standard procedures.
- ensure documentation of important details of each field data collection period.
- maintain hard copies of data forms and assemble original data forms for archiving on a regular basis.
- work with program coordinators to identify analysis and reporting mechanisms, and to establish a schedule for regular project milestones such as data collection periods, data processing target dates, and reporting deadlines.
- produce regular summary reports and conduct periodic trend analysis of data, store the resulting reports, and make them available to users.
- act as the main point of contact concerning data content.

The project manager works closely with the data manager to:

- develop quality assurance and quality control procedures specific to project operations.
- identify training needs for staff related to data management philosophy, database software use, quality control procedures, etc.
- coordinate changes to the field data forms and the user interface for the project database.
- fully document and maintain master data.
- identify sensitive information that requires special consideration before distribution.
- manage the archival process to ensure regular archiving of project documentation, original field data, databases, reports and summaries, and other products from the project.
- create data export and summary procedures to automate and standardize data products needed for secondary use or analysis.
- identify priority legacy data sets and convert to current formats.
- complete project documentation describing the who, what, where, when, why and how of a project.
- increase the interpretability and accessibility of project-related products.

2.2.3 GIS Specialist

The network GIS specialist provides fundamental program and project support by acquiring and maintaining geospatial data, performing spatial analyses, and documenting all spatial data sets. The GIS specialist also has an essential communication role in the network: maps and graphic spatial analyses are one of the principal means of conveying complex geographic or quantitative data to wide audiences. The GIS specialist obtains, creates, and manages spatial data themes associated with network inventory and monitoring projects, as well as other spatial data related to the full range of park resources. The GIS specialist works in collaboration with the project manager and data manager to:

- determine GIS data and analysis needs for the project.
- develop procedures for field collection of spatial data including the use of GPS and other spatial data collection techniques.
- conduct spatial analyses of project and related data.
- create maps and other means of displaying spatial data to meet project objectives.
- document data in compliance with spatial metadata standards.
- design databases and other applications for the network.
- create relationships between GIS and non-spatial data and create applications for their analysis.
- establish and implement procedures to protect sensitive spatial data according to project needs.
- develop and maintain an infrastructure for metadata creation and maintenance.
- ensure that project metadata are created and comply with national and agency standards.
- integrate tabular data with spatial data in a GIS system.

2.2.4 Cooperators

Cooperating researchers have played and will continue to play a significant role in I&M projects and associated data acquisition and data development. The SCPN is currently employing cooperating researchers to develop monitoring protocols. Cooperators may also be significantly involved in implementing some of the monitoring projects once protocols are completed. Cooperators serving as project managers and their field crews will assume the same responsibilities as described for network staff. In addition, cooperator responsibilities include:

- communicate regularly with the program manager and the data manager during the course of the project.
- communicate with park contacts regarding planned field work in their park units.
- follow data management procedures for field data collection, data entry, QA/QC, GPS data collection and metadata production.

2.3 Data Management Coordination

The Natural Resource Challenge states that collaboration among the National Park Service, 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 in parks are properly developed, maintained, and made available for management decision-making, research, and education.

The network data managers work with I&M Program data management staff and regional resource information management personnel to maintain a high-level of involvement in Service-wide and regional databases and data management policy. The data managers also work locally with network personnel, park staff, and cooperators to promote and develop workable standards and procedures that result in the compatibility and availability of data sets.

Key contacts include park GIS specialists and the project managers for each monitoring or inventory project. Involvement and input from park resource management staff is also essential. We rely on everyone in the network for the successful development of planning materials, inventory study plans,

and monitoring protocols. Consistent and productive communication among these individuals leads to common understanding and better synchronization of data management activities. This communication may take the form of personal visits, phone calls, email, joint meetings and training sessions, and participation in meetings and work of the SCPN Technical Advisory Committee and Board of Directors.

Chapter 3 Infrastructure & System Architecture

One of the most important functions of the I&M Program is to ensure that data and information produced by the program are disseminated and shared among network staff, network cooperators, park staff, and the larger scientific community. This chapter discusses the components of information system architecture and infrastructure needed to reach these diverse groups. It includes all infrastructure components required from project development through information dissemination via a clearinghouse.

Infrastructure refers to the network of computers and servers that information systems are built upon. System architecture refers to the applications, database system, repositories, and software tools that make up the framework of the SCPN's data management enterprise.

3.1 Infrastructure

The data management infrastructure needs to support these required functions:

- Provide a central repository for master data sets.
- Provide controlled subsets of data for local computing.
- Provide a means for uploading and downloading data for both NPS and public.
- Support desktop and internet applications.
- Provide security, stability, and backups.

SCPN relies on partnerships with Northern Arizona University (NAU) and the USGS Colorado Plateau Research Station (CPRS), as well as the national and regional information technology personnel and resources, to assist with maintenance of computer infrastructure (Table 3-1).

Table 3-1. SCPN infrastructure service or support providers.

Service or Support Provided	NAU	CPRS	NPS	SCPN
Telecommunications hardware and service	X			
Networking hardware	X			
Networking services	X	X		
Networking security		X		X
Computer hardware support and maintenance		X		X
Computer software installation and support				X
Email administration and support			X	
Web services			X	
Clearinghouse repositories			X	
Data backup			X	X

The computer server infrastructure used by SCPN includes both local and remote resources (Figure 3-1). Local resources are primarily maintained and provided by SCPN staff, CPRS, and NAU cooperators. Remote resources are primarily maintained and provided by national and regional NPS staff. Computer servers can be divided into five primary functions:

- *Work File Server* – A local server available for read/write access to all network staff where commonly shared working files are stored and backed up.
- *Data File Server* – A read-only repository of data files. Files stored here do not change often and need a consistent location, such as in the case of GIS data, imagery, or final reports. Write access is under strict control, while all NPS employees have read access.
- *Database Server* – Dedicated to running an enterprise level relational database management system (RDBMS) such as SQL Server, Oracle, or ArcSDE. Read and write access is under strict control. Data are accessed through a front-end application which may be running on the user's local machine or from a separate application server.
- *Application Server* – Provides applications used to access data stored in a RDBMS on a database server. Files on these servers change frequently and backups need to support these changes. Application servers are often Internet servers as well.

- *Internet Server* – Provides access to web pages, applications and files via the Internet. Acts as a gateway to data and information products and controls access to those products. Special software may be used on internet servers, such as ColdFusion for website pages and database applications, Blue Angel for clearinghouse search engines, and ArcIMS for internet interactive mapping.

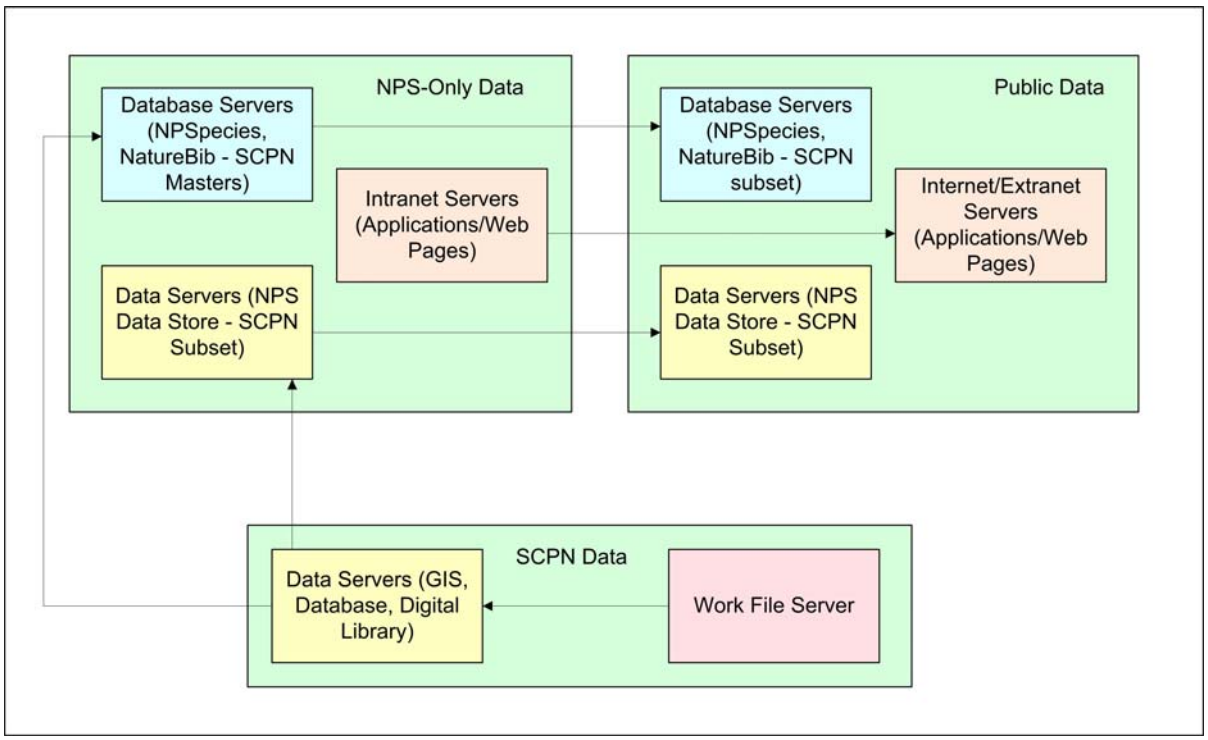


Figure 3-1. SCPN local, and national infrastructure.

3.2 SCPN System Architecture

SCPN computer resources and users are treated as a separate workgroup under the USGS domain on the NAU Wide Area Network. The NPS workgroup has access to the SCPN file and data servers, printer/scanner, and plotter. NPS users do not have access to any USGS resources, and USGS users do not have access to any NPS resources (except for the USGS systems administrator).

3.2.1 Working Project Files

Initial project files, for both short and long term, are typically stored on the work file server. The project files must conform to naming and documentation standards.

The data manager creates a working directory for each project. Project managers may modify this standard structure as needed, such as deleting directories that are not being used. Template project directory structures for short-term projects and protocol development have been developed (Figure 3-2).

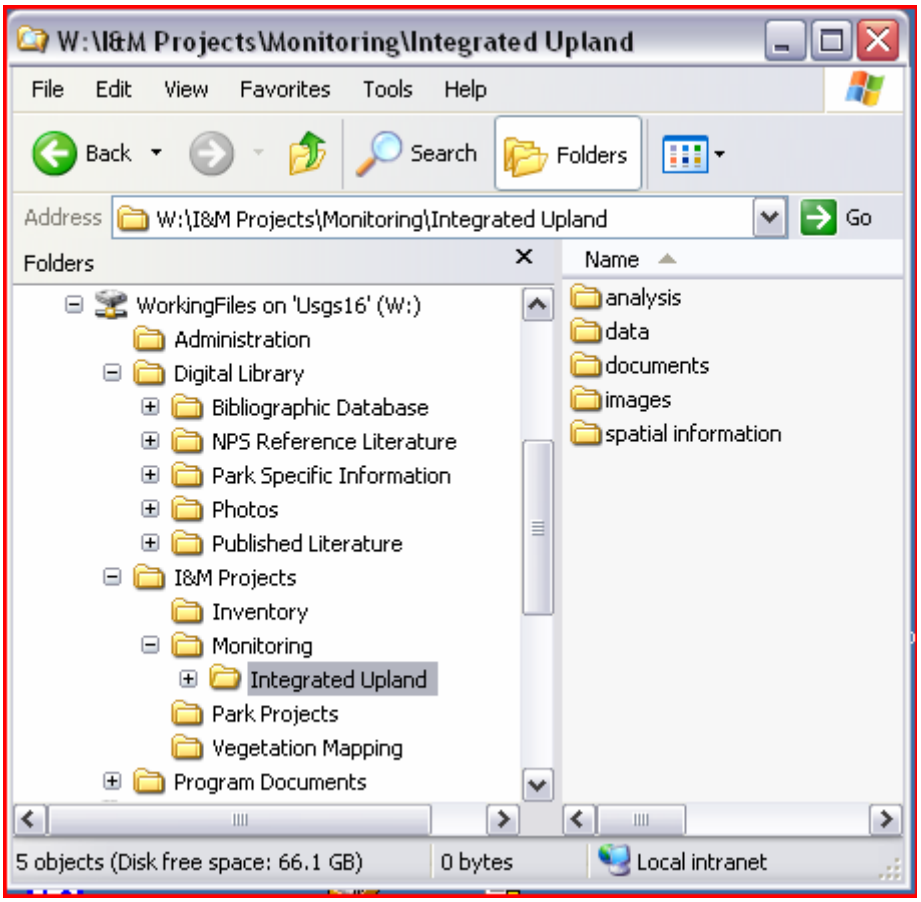


Figure 3-2. Directory organization of working files.

3.2.2 Integration of Products to Master Data Libraries

Final electronic versions of project deliverables (e.g., summary reports, final reports, databases) that have gone through all QA/QC procedures and been properly documented will be stored in several areas: archives, project directories, and database directories. The archives will have highly restricted write and read permissions and will contain final versions easily packaged for distribution or back up. The project directories will store all products from each project. The database directories will contain all project databases.

These files will be centrally stored in the SCPN file and data servers. Table 3-2 illustrates this general directory structure; exact directory structures are subject to change. Files will be reviewed for quality and standards prior to storing in the libraries.

Table 3-2. General directory structure and their roles for Master Libraries.

Location	Description	Custodian
X:\GIS_Data \Documented \Undocumented \Working	Storage of GIS data, remote sensing data and metadata.	SCPN - GIS Specialist
X:\Archives	Electronic archive of finalized project information. Includes administrative documents, reports, raw data, certified data, and distribution files.	SCPN Data Manager
X:\Databases	Storage of master project databases for use by SCPN staff.	SCPN Data Manager
X:\I&M Projects	Storage of current working documents and databases for each project, along with all project documentation.	SCPN Data Manager, Project Managers
NR-GIS Data Store \GIS_Data \Bibliography_PDFs \Biodiversity_Data	A copy of the electronic deliverables will also reside on the WASO Data Stores.	WASO

3.2.3 Integration of Long-Term Databases

SCPN currently uses Microsoft Access for all project databases. Master databases for SCPN projects will be stored on the data file server in the databases directory. SCPN may migrate to an enterprise relational database management system such as SQL Server at some point in the future. Other master databases which hold SCPN data are maintained by WASO and include NPSpecies, NatureBib and NPSSTORET.

3.2.4 Application Development

Applications are the front-end or interface with which we interact with data. The back-end, or the data itself, may be stored on a database server as describe above or as a stand alone Access database. SCPN will use Microsoft Access as the application interface for the short-term. Web-based application interfaces hold a large amount of potential for allowing wide-spread access to centralized databases. Web-based interfaces using protocols such as Cold Fusion and PHP will be evaluated for potential use in the future.

3.2.5 Internet Servers

Applications, products, and links to the clearinghouses discussed in this chapter will be available on the SCPN internet website. Information that is in draft form or that is subject to security protocols will be made available on the SCPN intranet or extranet websites, where applicable. SCPN internet, intranet, and extranet servers are hosted on the WASO servers.

3.3 LAN and WAN Infrastructure

LAN and WAN infrastructure is monitored by the NAU and CPRS IT staff and is mostly invisible to employees. Though the logical structure is stable, the actual architecture may change from day to day. Any reduction in speed of performance should be reported to the IT Staff.

3.4 IT Network Security

Over the past few years, enhanced security protocols have been or are being developed. Security procedures are being directed not only Service-wide, but also from the Department of Interior. National solutions to security are currently under development and will be applied and enforced by the regional IT staff. These procedures will affect local control of access to servers and their

respective information. Information Technology security is managed by the CPRS. SCPN staff will comply with all required security training and procedures advised by the DOI, NPS, and CPRS IT divisions. CPRS ensures that passwords and networking meet DOI standards and that firewalls and networks are secure.

Chapter 4 Data Management Process and Workflow

4.1 Types of Data and Information

Dictionary definitions of data tell us that data are factual information or values (often derived from scientific experiments) that can be used as a basis for reasoning or making decisions. Data become information through the process of analysis, synthesis and modeling. Metadata, documentation, reports and administrative records are other forms of data which are necessary to ensure the usability and longevity of monitoring data and these are included in the data management framework. The term 'data' will be used to represent the five general categories outlined in Table 4-1.

Table 4-1. Categories of data

Categories of Data	Examples
Raw data	GPS files, raw field forms and notebooks, photographs and sound/video recordings, telemetry or remote-sensed data files, biological voucher specimens
Processed data	Relational databases, tabular data files, GIS layers, maps, species checklists, output files
Documentation	Data collection protocols, data processing/analysis protocols, record of protocol changes, data dictionaries, FGDC/NBII metadata, data design documentations, quality assurance reports, catalogs of specimens/photographs
Reports	Annual progress reports, final reports (technical or general audience), periodic trend analysis reports, publications
Administrative records	Contracts and agreements, study plans, research permit/applications, other critical administrative correspondences

4.2 Data Management Framework

Within the context of a monitoring project, SCPN data management tasks can be divided into several types of activities. These activities (Figure 4-1) provide the backbone for planning and executing data management procedures for each monitoring project.

4.2.1 Monitoring Planning

Though not directly involved in the planning phase of a monitoring project, the data manager should remain aware of the initiation and progress of new projects. During the planning stage the data manager will establish a project directory on the working file server and in the archives. Each project directory will be established with a standard directory structure, which may be modified by the project manager if necessary. In addition, project team members will be given the appropriate level of access to the project directory and it will be included in backups.

4.2.2 Data Design

Data design involves design and development of field data sheets, the database structure, and the database application. The data manager and project manager have primary responsibility for this task; however, project crew members, the quantitative ecologist and the network coordinator may participate as well. Collaborating on the design and development of these three elements in an integrated manner will result in well-designed, easy to use, well-matched products for use in data acquisition, quality control and assurance, and data summary and export for the life of the project. See Chapter 5 for further details on the Data Design process and resulting products.

4.2.3 Data Acquisition and Quality Control

Collection of new data by project crew members takes place during implementation of the monitoring project. Acquisition of other NPS and external data by project managers and data management personnel may take place during the planning phase or during implementation. Once the data are collected or acquired they will be entered by project crew members or downloaded into the working project database. Quality control measures will be taken during collection and data entry to ensure

accuracy. Project crew members will verify that digital data accurately match data sheets from field collection. See Chapter 6 for further details on data collection guidelines and quality control procedures.

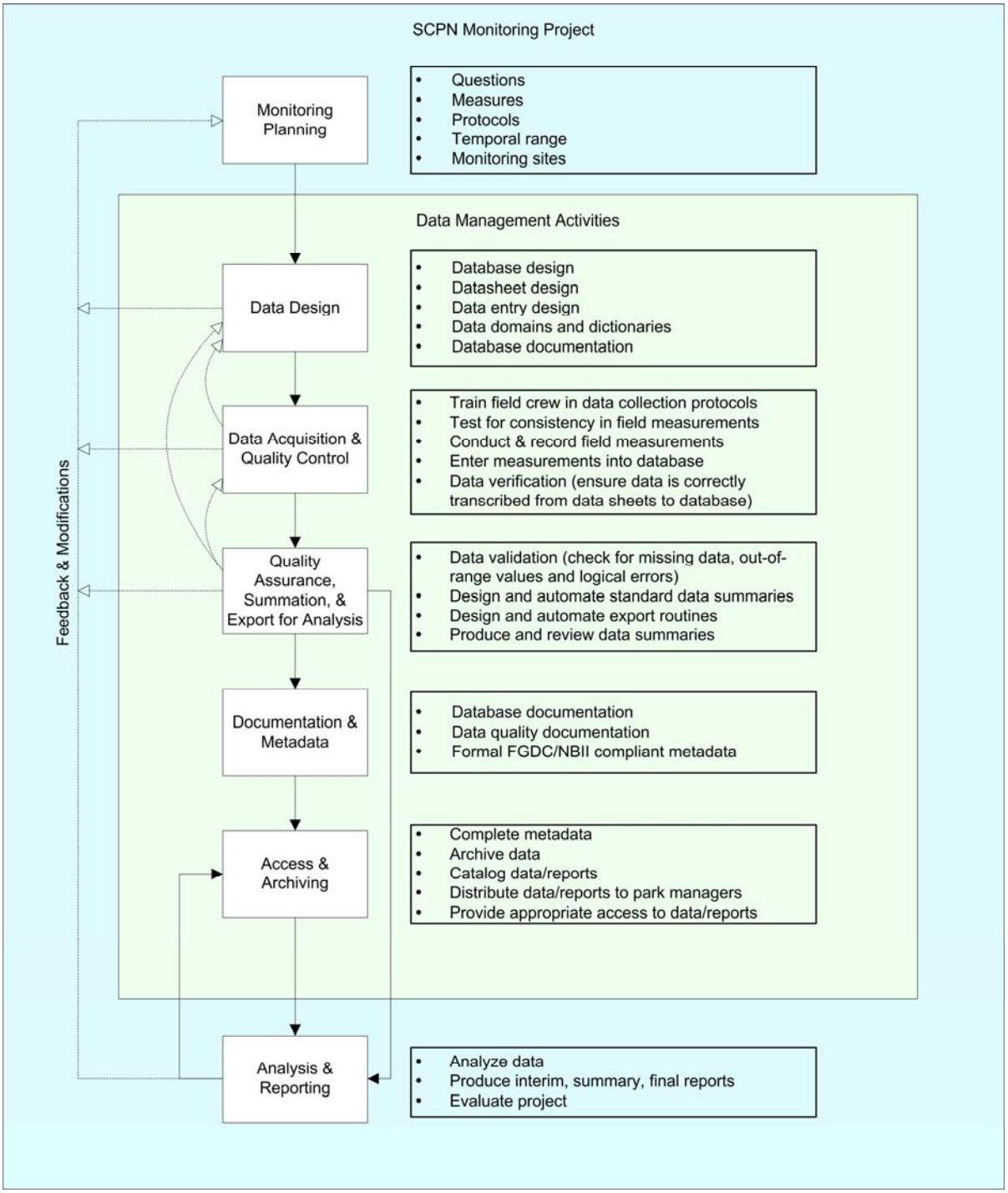


Figure 4-1. Data management activities within the context of a monitoring project.

4.2.4 Quality Assurance, Data Summarization, and Export for Analysis

Once the working database has been verified, the project manager will validate the current data collection period's data (or the entire database for short-term projects). This entails ensuring data are

logical and within normal ranges. Data summary tools will be used to complete this process. Once data are validated, routine data summaries and reports will be produced for distribution to park managers and other interested parties and data will be exported for analysis by the quantitative ecologist and project manager. See Chapter 7 for further details on quality assurance procedures and data summarization and export activities.

4.2.5 Documentation and Metadata

For short-term projects, database documentation will be completed and formal metadata produced at this stage. For long-term projects, data from the current collection period will be certified in preparation for moving data from the working database to the project master database. See Chapter 8 for further details on database documentation and data quality documentation.

4.2.6 Access and Archiving

Once verified and validated, a copy of the working database will be archived and prepared for distribution, and the data will be uploaded to the project master database. Short-term project databases will be archived and prepared for distribution. Data, summary reports, and field products (data sheets, specimens, etc.) will be packaged for distribution to parks and archived in SCPN file servers and repositories. See Chapters 9, 10, and 11 for more details related to distribution and dissemination of monitoring information products and archiving digital and analog materials.

4.2.7 Analysis and Reporting

The network will follow a schedule of project milestones and dates for production of reports. Reports will be written by project managers in cooperation with the data manager. Reports will include routine data summaries, long-term trend analysis and customized data reports to meet park management needs.

4.3 Phases of Data Processing

Project data take different forms and are maintained in different places as they are acquired, processed, documented and archived. These phases can be modeled as a sequence of events and tasks which involve interaction with the following objects:

- *Raw data* – Analog data recorded by hand on field data sheets 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 time period). This may be the only database for short-term projects with no need to distinguish current season data from the full set of validated 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 manager that the data have passed all quality assurance requirements and are complete and ready for distribution. Metadata records include detailed information about project data needed for proper use and interpretation.
- *Master database* – Project-specific database for storing the full set of validated project data, used for viewing, summarizing, and analysis. Current season data from the working database must pass all quality assurance steps prior to upload into this master project database.
- *Reports and data products* – Information that is derived from certified project data.
- *Edit log* – A means of tracking changes to certified data.
- *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.
- *Archived data* – Digital and hard-copy data stored and maintained for the long-term.

Figure 4-2 illustrates the sequence of events that a piece of data might follow from data collection to posting of final information products on public clearinghouses. The raw data, whether analog or digital, will immediately be archived prior to entering it into the working database. The working database will be verified, validated, certified and documented before it is archived and uploaded to the master data set. Data summary, export for analysis, and reporting will be accomplished with the

certified database or the master database depending on the project. Data from the certified and master databases will be extracted and uploaded to NPS clearinghouses and used to update national databases such as NPSpecies. Reports and other information products will be distributed.

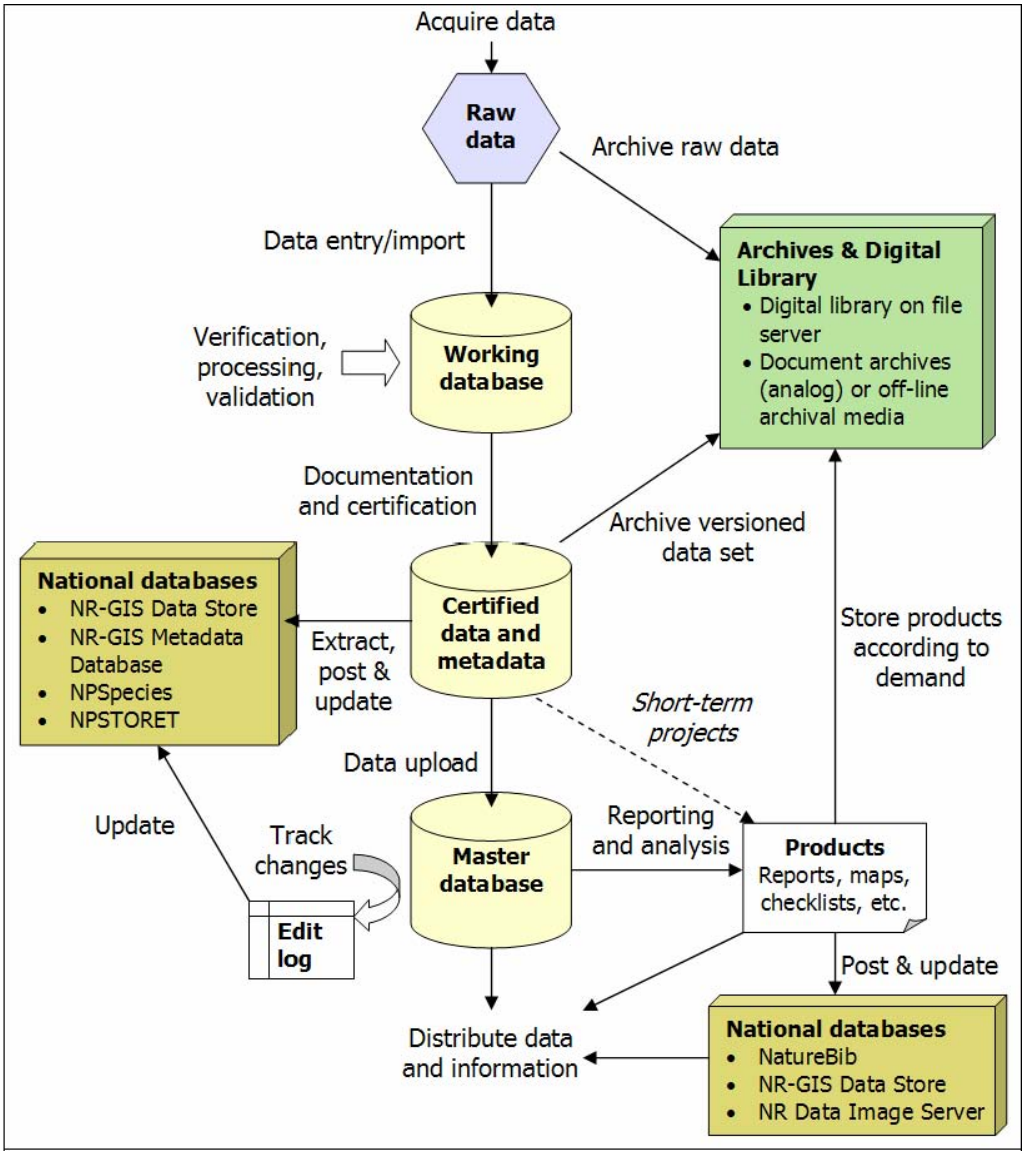


Figure 4-2. Diagram of typical phases of I&M project data.

4.4 Water Quality Data

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. SCPN will oversee the use of NPSTORET according to the network's integrated water quality monitoring protocol and will ensure content is transferred at least annually to NPS Water Resource Division for upload to the Environmental Protection Agency's STORET (STORage and RETrieval) database.

Chapter 5 Data Design

Design and development of field data sheets, physical database structure and the database application works best as a collaborative and integrated process. Primarily the responsibility of the data manager and project managers, the design process may also include project crew leaders and members and the network coordinator. Coming to a common understanding of the data collection process, necessary data elements, and the relationships between them is essential for designing data sheets and databases that will serve the monitoring project for its duration. A data modeling process will be used to facilitate communication among project stakeholders.

5.1 Data Modeling

Data modeling is a technique for defining, organizing, and documenting a project's data needs and translating those needs into a physical database into which data will be entered and managed. The process involves the project manager and data manager working together to define the items which will be represented by the data (entities), the pieces of information that describe those items (attributes), and the connections between them (relationships). Understanding the relationships between data elements is one of the most important steps in creating a data model. Data models are typically illustrated using entity relationship diagrams (ERDs), but may also include flow diagrams and associated descriptions. Data modeling is completed in three stages: conceptual, logical, and physical (Appendix G).

The data manager and project manager will work together to produce data models using the monitoring protocols and other references. Four elements are required to develop the data models:

1. *People* – The project stakeholders will need to be involved in the model development. Discussions may begin between a few with review from a larger group or start with a large group working toward the specifics with a smaller group. Stakeholders may include the network coordinator, scientists, cooperators, field crews, statisticians, and data managers.
2. *Protocols* – The protocols for the vital sign will provide the greatest substance to the models. Ideally, a conceptual data model would be included in the protocol. The protocol should clearly outline the goals, objectives, methods, standards, analysis, and desired output.
3. *Reference materials* – Reference materials such as field forms, drawings, mock-up reports, and references to classifications to be used will play a significant role in the data models.
4. *Frequent interactions* – Frequent discussions between the project manager, project stakeholders, and the data manager will be needed to develop successful data models. Detailed review of the protocols and reference materials will unravel the entities, relationships, and flow of information.

Data modeling is not a simple linear process and requires frequent communication between all parties involved. There will be several iterations and revisions until the proper data model emerges.

5.2 Database Standards

Database standards promote compatibility among data sets so that integration, aggregation and summarization can take place. Currently, there are no Service-wide standards on database design; however, the I&M Program has developed a series of recommendations. These include the Natural Resource Database Template (NRDT) and Recommended Naming Standards.

Each database must meet specific monitoring project needs. SCPN will use the above recommendations in database design as guidelines and will standardize naming according to the Recommended Naming Standards. SCPN will also develop standardized lookup tables for data elements shared across many monitoring projects.

5.3 Data Sheet Design

Development of data sheets can benefit from the data modeling process outlined in the previous section. If possible, data sheets will be developed after the conceptual model has been completed. Project deadlines may require that data sheets be developed prior to data modeling.

Data sheets will be developed by project managers and project crew leaders with input from the data manager. Standards for data sheets will be developed and will include the following key ideas:

- Project and location information will be pre-printed on data sheets used in long-term projects. Once plots are established and geographical (coordinates, place names) and static environmental (slope, geology) data are collected, they will be pre-printed on data sheets for use in subsequent sample events.

- Fields for recording quantitative measures will include units of measurement.

- Fields for recording qualitative measures will include a data domain from which to choose the correct value, preferably allowing a project crew member to check or circle the value rather than write it in.

- Comments fields will be directed to a specific aspect of the data collection. One comment field for each major section of data collection will be the goal.

- Ideally, data sheets will be designed such that all data fields are filled in when the field crew has completed data collection. Data sheets with large quantities of optional fields scattered across a page make it difficult for field crew members to quickly discern whether all required measurements have been taken. Raw data will be recorded in each field; if calculations are necessary, these will be performed by the database. Optional fields, when needed, will be clearly marked.

5.4 Database Application Design

Database design takes care of the physical database or the back-end where the data are stored. A front-end, commonly called the database application, must be designed to facilitate user interaction with the database back-end. Design of the database application is the responsibility of the data manager with input from those who will use the database – project crew members (data entry) and project managers (data review, data summarization and reporting, export for analysis).

Database applications will be designed with a common look and feel and will include these major components:

- Switchboard or Main Menu
- Data entry forms
- Queries and export forms
- Reports

Properly designed data entry forms facilitate the data entry process and reduce errors of transcribing written data into digital data. They will generally include the following features:

- Data entry forms which match field data sheets as closely as possible
- Pick lists for data fields with defined domains
- Input masks for fields with particular formatting requirements (e.g., date)
- Constraints for fields with identifiable possible data ranges (e.g., temperature)
- A design that minimizes use of the mouse, allowing movement and data entry to occur primarily through use of the keyboard

Query, export, and report forms will be designed to allow users to build semi-custom queries for data summarization and export for analysis. Routine reports will be designed to allow users to quickly

produce data summary reports for export to Microsoft Office applications. These forms will allow a user to select certain attributes and ranges to define a subset of data.

Chapter 6 Data Acquisition and Verification

This chapter describes general guidelines and procedures for acquiring or collecting data sets, entering or integrating them into a project database or digitizing them for use with a computer, and verifying that the electronic data match the original source data. The origin of data (Network, NPS, or external) dictates the procedures used in acquiring and verifying the data for use.

Quality control procedures are used during data collection and entry to prevent errors of commission - forgetting to collect a measurement or entering the wrong code in the database. Verification procedures are used after data entry to ensure data stored in the project database match source data, e.g. field data sheets.

Data managed and utilized by the network will originate from three types of sources: within the network; other NPS data collection efforts; and outside the NPS altogether. The origin of data (Network, NPS, or external) dictates the procedures used in acquiring and verifying the data for use.

- **Network Data** – any data produced from projects that are initiated (funded) by the SCPN I&M Program or projects that in some way involve the I&M Program.
- **NPS Data** – any data produced by the NPS that did not involve the I&M Program.
- **External Data** – any data produced by agencies or institutions other than the National Park Service.

6.1 Network Data

SCPN will have the greatest control over data collected by network staff or cooperators working under cooperative agreements. Network data also differ from the other two categories in that the data will generally be collected during field studies.

6.1.1 Data Acquisition

Most network data will be collected by SCPN staff or cooperators during field-based phases of monitoring projects. Specific data collection procedures for each SCPN project will be established during protocol development. The following general guidelines will be followed to reduce error during field data collection:

Field crew will be adequately trained in data collection techniques

Project specific Standard Operating Procedures (SOPs) will document field methods to be used. The project and/or crew leader will prepare a manual and conduct training for the crew members prior to the start of the field season. All crew members will be trained in proper measurement and estimation techniques and will be assessed on consistency and repeatability of each measure and estimation.

Standardized data sheets will be used and all required fields completed

Data sheets designed to specifications documented in Chapter 5 will be used when electronic data collection methods are not practical. Crew members will record all information in clearly legible printed handwriting and will complete all required blanks on the form. If any information needs to be altered, it will be crossed out with a single line; the new information written next to it; and the editor's initials and the date the changes were made included at the end of the new text.

Electronic devices will be used for data collection when possible (GPS units) and information should not be repeated on data sheet

GPS units will always be used to collect and record geographic locations unless the terrain makes it impossible to obtain a recording (e.g., deep canyons). Coordinates will be recorded on paper data forms only as a backup to GPS data collection. GPS data will be downloaded and processed (see Appendix C) upon return from the field and imported into the project database - coordinates should not be entered from data sheets. Certain projects may lend themselves to using handheld computers for data entry directly in the field. In this case, project specific SOPs

will be developed for data backup and transfer.

Automated data loggers will be used when possible

Data loggers are an efficient method for recording continuous sensor data (such as air and water quality) and will be used where appropriate. Routine inspections and calibrations will be completed on the data loggers prior to and throughout the field season, and calibration reports will be kept with the project documentation.

Field equipment will be regularly and properly calibrated

Field equipment with calibration and maintenance needs will be routinely inspected and calibrated. Calibration reports will be kept (when practical) with project documentation. Project crew leaders will ensure that all equipment is synchronized (e.g., watches, compass declination).

Project crew leader will keep a log of field-based actions and decisions

The trip log should record all major milestones crossed during field data collection; minor and major problems with data collection procedures, equipment or data loggers; and field-based decisions to alter methods or protocols. Logs will be kept in field notebooks using proper procedures (e.g., name, number and date each page).

Project crew leader will assess crew members' consistency and repeatability of measurements over the field season

Crew members will be assessed over the course of the field season for consistency and repeatability of measurements. Measurements and estimations will be repeated by crew leader and crew members to ensure that they have not drifted from the standards established during pre-season training.

6.1.2 Data Entry

Data entry is the initial set of operations for transfer of raw data from paper field forms into a computerized form linked to database tables. When data are gathered or stored digitally in the field (e.g., on a data logger), data entry consists of the transfer of data (downloading) to a file in an office computer which can be imported into the proper database. Specific data entry/transfer procedures for each SCPN project will be established during protocol development. Data entry will be the responsibility of project crew leaders and members and the data management assistant. The following general guidelines will be followed to reduce error during data entry:

Personnel completing data entry will be properly trained

The data manager will provide training to ensure proper use of the database application and any other software or hardware necessary to complete data entry.

Standard data entry forms will be used in the database application

Data entry forms designed to specifications documented in chapter 5 will be used by all data entry technicians.

Data will be entered in a timely manner

All data will be entered or downloaded into the project database as soon as possible after collection. Data entry should not be delayed until all the project data have been collected unless delay is unavoidable.

Work will take place in a clean well-organized workspace with ergonomic equipment

Desktop space near the computer should be free of clutter and distractions. 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 computer workstation should be configured to be ergonomic to reduce fatigue.

Two data entry technicians will be used, if possible.

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, s/he should work at a slower pace to avoid errors.

6.1.3 Data Verification

Quality control procedures used during data collection and entry have been discussed in the two prior sections. Despite these preventative measures, data entry errors will almost always occur. Data verification is the process by which digital data are checked to ensure that they match the source data. It is the responsibility of the data entry technician to verify that the data entered into the database match the data on the datasheets. Project managers will check a small subset of the data to ensure that data entry technicians are properly verifying data. Specific procedures for verification will be established during protocol development by the data manager and project manager. One of the following three methods will be used to verify data entry in SCPN projects in order of preference:

Visual review, printed, 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. For cases in which only one reviewer is available, someone other than the person keying data should perform 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).

Visual review, on-screen, after data entry

All records are reviewed upon the completion of data entry by comparing values on-screen against the values from the hard copy data sheets. Errors are corrected immediately. The review should be conducted by someone other than the data entry technician who was originally responsible for entering the data.

Visual review, on-screen, at data entry

The data entry technician verifies each record after input. Values recorded in the database are compared with the original values from the hard copy data sheets. Errors are corrected immediately. 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.

In addition, summary queries and tallies will be included in each project database that can be run to detect broad errors such as duplicate, omitted or unlinked records. Spatial data will also be reviewed visually for broad errors (e.g., points located outside park boundaries or in the middle of lakes).

Data transferred from a data logger or field computer do not have the same chances of transmission error as data entered from paper data sheets. These data should be reviewed to ensure that the transfer process imported the data to the correct tables and fields of the project database and that data formats are correct.

6.2 NPS Data

NPS data are data collected or generated by national, regional or park-based efforts that were not initiated or funded by the SCPN I&M Network. Data may be used for (but not limited to) selection of monitoring sites (e.g., spatial data), providing a baseline for comparisons with modern data (e.g., historic vegetation plots), and providing ancillary data for trend analysis (e.g., climate data). Data may include everything from a relational database application to raw data sheets to annual and final project reports.

6.2.1 Data Acquisition

Parks

Locating and cataloging NPS data from individual parks, called 'data mining', has been an important step in developing an information base for SCPN and will continue into the future in a focused manner. The goal of network data mining is not to acquire and store, at the network, copies of all data residing at parks. Instead, cataloging tools are used to describe and document information sources so that potential users can find them. SCPN uses NatureBib to catalog and upload digital versions of documents of interest. Datasets are tracked locally using Reference Manager – a networked bibliographic database. Data that are an exception to this rule are documents or data sets that are directly related to monitoring projects and base spatial data.

Data mining during monitoring planning has been accomplished by individuals or teams visiting parks and locating, evaluating, cataloging, and acquiring reports and data. Early data mining efforts did not focus on specific topics, but were broad sweeps of relevant natural resource information. During protocol development, data mining efforts will continue, but will focus on locating and gathering information directly related to specific protocol development projects.

SCPN will serve, for those parks that are interested, as a second library of natural resource information. Close working relationships will be established and maintained with park personnel so that sharing of data and new reports is a regular and automated process.

Regional and National NPS

NPS regional and national programs support all of the parks within the Intermountain Region and also provide a good resource for natural resources information. The following programs conduct long-term monitoring or inventories and will provide data to inform the SCPN monitoring program:

Air Quality

Ozone, visibility, and deposition data are collected in a nationwide network and processed, quality checked, archived and disseminated by the Air Resources Division (ARD). Data for individual parks, national trends, and summary information for the SCPN area are available through ARD's website.

Exotic Plant Management Team

Data collected in individual parks are processed, quality checked, archived and disseminated by the Colorado Plateau EPMT, located at Petrified Forest NP. Data are entered into a standard database and will be acquired through the EPMT data manager or liaison.

Fire Program

Fire effects monitoring data are collected and managed throughout the Intermountain region by park-based teams. Teams that are responsible for SCPN parks are based at Bandelier NM, Grand Canyon NP, and Saguaro NP. Data are collected using standard protocols (documented in the Fire Monitoring Handbook) and are entered into a standard database (the Fire Ecology Analysis Tool or FEAT). In addition to fire effects monitoring data, SCPN may acquire data about fire occurrence and severity available through the National Interagency Fire Center (NIFC) and the USGS/NPS National Burn Severity Mapping Project.

GIS

The regional GIS program assists many SCPN parks that do not have in-house GIS specialists available to collect, document and disseminate spatial data. These data are available through the NPS GIS data clearinghouse.

Water Quality

Water quality data collected by regional programs are uploaded to STORET. SCPN will acquire local updates from the national STORET database on an annual schedule.

Soils

Soil maps are being produced through a partnership with the Natural Resource Conservation Service (NRCS) and will be available through the NPS soils inventory coordinator or NRCS.

Geology

Geology maps are produced by the Geologic Resources Division and are available through the NR-GIS Metadata and Data Store webpage.

Vegetation Mapping

Vegetation maps are being produced by the NPS Vegetation Mapping Program. Maps currently in production are being coordinated by SCPN. Finished products will be available directly from cooperators and through the vegetation mapping program.

6.2.2 Data Entry

In most cases processing or entry of NPS data will be minimal. Legacy park-based data will be evaluated for modernization (in the case of legacy digital data) or digitizing (in the case of legacy non-digital data). Well-documented and current park, regional, and national data will be stored in the appropriate place on the SCPN file servers. Some NPS data processing will include:

- Enter all new park biodiversity data into NPSpecies.
- Enter all natural resource reports and publications related to SCPN parks into NatureBib. Hard copies will be scanned and electronic copies archived in the proper directory on the SCPN file server and uploaded to NatureBib.
- Ensure that all spatial data are in the proper projection, and FGDC compliant metadata are available for viewing through ArcCatalog.
- Enter all data sets into Reference Manager (a networked bibliographic database to which all SCPN staff have access) for local tracking. Create compliant metadata for appropriate data sets for inclusion in the NR-GIS Metadata and Data Store.
- Import STORET water quality data into SCPN Water Quality Database (WQD).

6.2.3 Data verification

NPS data acquired by SCPN will be completed data sets, for the most part, and as such will be verified by the original producers of the data.

6.3 External data

There are myriad data sources available with information pertinent to the SCPN monitoring program. Types and sources of external data will be identified during project planning and protocol development. Sources of information may include other federal state and local government agencies, academia and private or non-profit organizations. SCPN will only acquire those data sets with appropriate documentation and metadata. Data sets may be integrated into project databases or stored on SCPN file servers while they are in use. SCPN will not archive external data that are currently archived and maintained by other organizations (i.e., climate data).

Chapter 7 Data Validation, Summary and Export for Analysis

This chapter describes general guidelines and procedures for validating data that have been collected and entered into a database, working with the data to produce standard summaries, and exporting the data for use in statistical analysis software. Data validation, data summary, and data analysis are the responsibility of the project managers; however, the data manager will provide tools to project managers to facilitate these three activities.

7.1 Data Validation

Data validation refers to the process of ensuring that properly transcribed data (verified) are logical and within normal ranges. Data validation is the responsibility of the project manager and will be completed by a member of the project who has comprehensive knowledge about the data and is able to recognize data inconsistencies and logical errors.

Corrections or deletions of logical or range errors in a data set require notations in the original paper field records about how and why the data were changed. Modifications to the field records should be clear and concise while preserving the original data entries or notes. 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 its unique contents and domains. Specific procedures for data validation will be delineated in individual monitoring project data management SOPs. The following general methods will be used as guidelines:

Data entry application programming

Certain components of data validation, such as range checking, will be built directly into data entry forms. Database fields that represent measures with known ranges (e.g., pH or temperature) will be programmed to alert the user if a value is entered that exceeds that range. This will eliminate validation errors caused by mistyped data; however if the out-of-range value is recorded on the data sheet, the data entry technician will consult the project crew leader and/or the project manager to obtain clarification. If the out-of-range value on the data sheet is determined to be a valid measurement, it will be treated as an outlier (see next item).

Outlier Detection

GIS, database, graphic, and statistical tools for ad-hoc queries and displays of the data can be used to detect outliers. Outliers will not be immediately eliminated; extreme values can and do occur. Each outlier will be investigated by the project crew leader and the project manager who will try to determine if some contamination during measurement is responsible. If an outlying value is determined to be a valid measurement, it will be noted in a log and included in the data set documentation.

Other exploratory data analyses

Palmer and Landis (2002) suggest that in some cases, calculations for assessment 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. Normal probability plots, Grubb's test, and simple and multiple linear regression techniques may also be used (Edwards 2000).

Once validation has been completed, it will be documented. See chapter 8 for guidelines on documenting the quality of a data set.

7.2 Data Summaries

One of the values of using relational databases for storing digital project data is the ability to create and automate routine summary and reporting features within a database application. The data manager will work with the project managers, network coordinator, and ecologist to determine

appropriate summary queries for each monitoring project. These automated summaries will facilitate the production of routine data summaries and long-term trend analyses on a regular schedule according to project milestones. Customized data reports will also be prepared to meet park-specific management needs.

As the SCPN monitoring program grows, a library of automated routines and queries will be built, allowing the data manager to release each new project database application with a core set of standard summary reports. Standards for summary reports (Appendix H) will be developed as monitoring protocols are further developed.

7.3 Export for Analysis

Monitoring project data that require statistical analysis beyond simple means or sophisticated graphing for visualization will be exported so that third-party software applications can be used for these functions. Much like the data summaries, automated export routines will be developed for each database application. Many statistics software programs have specific requirements for a data file's file format, field format and length, and field order. Specifications for each third-party software program will be detailed in the Export for Analysis SOP (Appendix I) and standard automated export routines will be included in applicable project database applications.

Chapter 8 Documentation

Data documentation is the most important step toward ensuring that data sets are usable well into the future. Documentation can include information about how and why the data set was created, who created it, when and where it was created, and data quality. It also includes instructions for using database applications, details about application development and programming, and contact information for responsible parties.

8.1 Metadata

The term *metadata* is defined as information about the content, context, quality, structure, accessibility, and other characteristics of data. In addition to ensuring data longevity, metadata increase the possibility of data sharing and reuse for multiple purposes.

Mandate for Documentation

While the importance of metadata is widely accepted within the data management community, the approaches for storing this information and the levels of detail can vary. However, some established metadata strategies apply to National Park Service data, including the following:

- Executive Order 12906, signed by President Clinton in 1994, mandates federal agencies to "...document all new geospatial data it collects or produces, either directly or indirectly..." using the Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM). 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 plus 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) Clearinghouse. Although not a requirement, completion of the Biological Data Profile for appropriate data sets is recommended by NBII.
- The National Park Service (NPS) Geographic Information System (GIS) Committee requires all GIS data layers be described with FGDC standards and the NPS Metadata Profile.

The Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM) provides standard attributes to appropriately document data and provide consistency so attributes may be searched. This standard was primarily designed for spatial information and follows similar attributes as those required for a bibliographic reference. The standard, however, can be used for any kind of information, such as spreadsheets, photos, hardcopy maps, etc. More information about the FGDC can be found at www.fgdc.gov

All data, whether spatial or non-spatial, will have an associated metadata file. Spatial data will have a fully FGDC compliant metadata record. Non-spatial data will have at least a minimally FGDC compliant (Sections 1 and 7) metadata record. In the case of vital signs monitoring, each vital sign will be thoroughly documented by protocols and standard operating procedures. These other documents should be cited and hyperlinked into the metadata.

NPS Integrated Metadata System

The NPS Integrated Metadata System will serve as a clearinghouse and search tool for NPS metadata records and associated data sets for NPS purposes only. This system will include public and NPS-only metadata records. Searches done by other FGDC Clearinghouses will not have direct access to these databases. Therefore, the NR/GIS Metadata database (public records only) will be automatically imported into NPS Focus. NPS Focus is a true clearinghouse and can be accessed using NPS Focus or any other FGDC Clearinghouse search engine.

Metadata records created by the parks or networks are integrated into the NR/GIS Metadata database. Associated data sets will be stored in the NR/GIS Data Store. The locations of the associated data sets are directly linked in the metadata record.

There are several tools available for developing formal metadata. SCPN currently uses ArcCatalog as the primary tool for metadata creation for spatial datasets. Metadata creation for non-spatial datasets is completed with several software and internet tools including Dataset Catalog, Metavist, metadata parser, and the Integrated Taxonomic Information System website (see Appendix J for metadata creation procedures for non-spatial datasets).

8.2 Data Quality Documentation

Guidelines for verifying and validating data were described in chapters 6 and 7. This section describes guidelines for documenting the data quality of monitoring project data sets.

The level and format for documentation of data quality can vary from maintaining a record in the database itself stating that each field has been quality checked to maintaining a text document associated with the whole database indicating that it has been quality checked, and almost everything in between. Each technique has positive and negative aspects, and the right balance needs to be struck between time-intensiveness and obtaining the documentation that will enable longevity of the data. Specific documentation procedures for each monitoring project will be developed during protocol development, and will include these general guidelines:

Verification and validation processes used
The specific processes used to verify and validate the data set will be described in the documentation.

Name, position and affiliation
The name(s) of the person(s) who completed the quality checks, their positions, and their affiliations will be included. Each person who worked on verification and each person who worked on validation should be included.

Date
Dates the quality checks were completed. Dates should be given for each phase of verification and validation.

Scope
The domain of records which were included in the quality check should be detailed in the documentation. This will include percentage of records checked, names of fields checked (if not all fields were checked), and the dates or record numbers of the records checked (for multi-year monitoring data sets).

Description of errors and/or inconsistencies
A narrative section to describe any unresolved inconsistencies or errors in the dataset.

Whether the data quality documentation resides within the database or exists as a separate document, it will be imported into the FGDC metadata for the data set once it has been completed.

8.3 Database Application Documentation

The data manager will produce complete documentation of the database application. This will include a manual for using the database (installation, entering data, producing summaries, exporting data), an Entity-Relation Diagram (ERD) and data dictionary for the back-end, and documentation of any code used in the application front-end. Documentation will be project specific, but will follow these general guidelines:

Database Manual

- Produced prior to start of data entry.
- File names, formats and locations within project directory structure.
- Installation instructions for applications not located within SCPN networked environment.

- Description of general navigation of application.
- Line-by-line instructions for data entry.
- Instructions for producing summary queries/reports and exporting data to third-party software.

ERD & Data Dictionary

- Physical data model (as an entity relationship diagram) showing tables, primary keys, foreign keys, indexes and relationships.
- Field list that includes field name, data type and length, description or definition and data domain.

Code Documentation

- Comments will be used to describe the actions and results of each sub-routine and function in MS Access form and module VBA programming.

Chapter 9 Data Ownership and Sensitivity

9.1 Data Ownership

Significant monitoring protocol development is currently being accomplished through cooperative agreements with other federal agencies and universities. Some monitoring may be implemented through cooperative agreements. Data collected under such agreements or contracts are owned by the National Park Service, as established under Office of Management and Budget, Circular A-110, Section 36.

9.1.1 NPS Policies on Data Ownership

The National Park Service defines conditions for the ownership and sharing of collections, data, and research funded by the United States government. All cooperative and interagency agreements, volunteer agreements, and 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 Service.
- 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, and that federal funding is disbursed in accordance with applicable laws and regulations. OMB circulars establish some degree of government-wide standardization 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 the following administrative requirements pertinent to data and ownership:

> *(a) The recipient (higher institution or non-profit organization receiving federal monies for natural resource inventory and/or monitoring) may copyright any work that is subject to copyright and was developed, or for which ownership was purchased, under an award. The Federal awarding agency(ies) (in this case the National Park Service) reserve a royalty-free, nonexclusive and irrevocable right to reproduce, publish, or otherwise use the work for Federal purposes, and to authorize others to do so.*

Section 36 also states:

> *(c) The Federal Government has the right to:*
>
> *(1) obtain, reproduce, publish or otherwise use the data first produced under an award*
> *(2) authorize others to receive, reproduce, publish, or otherwise use such data for Federal purposes*
>
> *(d) (1) In addition, in response to a Freedom of Information Act (FOIA) request for research data relating to published research findings produced under an award that were used by the Federal Government in developing an agency action that has the force and effect of law, the Federal awarding agency shall request, and the recipient shall provide, within a reasonable time, the research data so that they can be made available to the public through the procedures established under the FOIA. If the Federal awarding agency obtains the research data solely in response to a FOIA request, the agency may charge the requester a reasonable fee equaling the full incremental cost of obtaining the research data. This fee should reflect costs incurred by the*

agency, the recipient, and applicable subrecipients. This fee is in addition to any fees the agency may assess under the FOIA (5 U.S.C. 552(a)(4)(A)).

(2) The following definitions apply for purposes of paragraph (d) of this section:

(i) Research data are defined as the recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues. This "recorded" material excludes physical objects (e.g., laboratory samples)…
(ii) Published is defined as either when:

(A) Research findings are published in a peer-reviewed scientific or technical journal; or
(B) A Federal agency publicly and officially cites the research findings in support of an agency action that has the force and effect of law.

(iii) Used by the Federal Government in developing an agency action that has the force and effect of law is defined as when an agency publicly and officially cites the research findings in support of an agency action that has the force and effect of law.

9.1.2 Network Policies on Data Ownership

To establish proper ownership, format, and development of network products, all cooperative or interagency work must be conducted as part of a signed collaborative agreement. Every cooperative agreement, interagency agreement, or contract involving SCPN must have the following citation from OMB Circular A-110 under a Reports and Deliverables section:

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

9.2 Data Sensitivity

All data and associated information from I&M activities will 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 will carefully identify any information that is considered sensitive, and will manage those data needing access restrictions and those to make publicly available.

9.2.1 NPS Policies on Data Sensitivity

The Freedom of Information Act, 5 U.S.C. § 552 (FOIA), stipulates that the United States Government, including the National Park Service, must provide access to data and information of interest 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. FOIA is intended to establish the right of any person to access federal agency records that are not protected from disclosure by exemptions. 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.

In some cases, public access to data can be restricted. The NPS is directed to protect information about the nature and location of sensitive park resources under one Executive Order and four resource confidentiality laws:

- Executive Order No. 13007: Indian Sacred Sites
- National Parks Omnibus Management Act (NPOMA; 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)

The following guidance for determining whether information should be protected is suggested in the draft Director's Order #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 in the park 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?

9.2.2 Network Policies on Data Sensitivity

All monitoring information products will be vetted for sensitive data prior to making them available to the general public. Classification of sensitive I&M data will be a shared responsibility that includes network staff, park resource management staff, park superintendents, and investigators working on individual projects. Park management has ultimate responsibility for deciding which information is sensitive and should not be released to the public. The network has ultimate responsibility for ensuring that sensitive data are not released to the public.

Generally, information may be withheld regarding the nature and/or specific locations of the following resources recognized as 'sensitive' by the National Park Service:

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

SCPN will complete a decision tree process that will help standardize the designation of sensitive data, and ensure that comparable guidelines are used for all parks. The decision tree will also provide documentation of the rationale behind sensitive status designations. The decision tree process is expected to include these general steps:

- 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 archiving of those records to avoid their unintentional release

Chapter 10 Dissemination and Access

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. Title II of the National Parks Omnibus Management Act of 1998 (P.L. 105-391) specifically encourages the publication and dissemination of information derived from studies in the National Park Service, as well as cooperation with other federal monitoring and information collection efforts.

To accomplish this goal, procedures must be in place to ensure that data collected, developed, or assembled by SCPN staff and cooperators are made available for decision making, research, and education. Providing well-documented data in a timely manner to park managers is especially important to the success of the program. SCPN must ensure that:

- data can be easily found and obtained.
- only data subjected to full quality control and quality assurance measures are released.
- data are accompanied by complete metadata.
- sensitive data are identified and protected from unauthorized access and distribution.

SCPN will use a number of distribution methods that will allow information collected and developed as part of the program to become widely available to park staff, the public, and other potential users.

10.1 Distribution Mechanisms

The primary mechanism for distributing data products will be the Internet. This will allow the data and information to reach a broad community of users. In support of the NPS I&M Program, web-based applications and repositories have been developed to store a variety of park natural resource information. We will use the following applications and repositories to distribute our data products:

NatureBib

The Service-wide bibliographic database for all NPS park units has the ability to store full-text documents in PDF format and it will be used to make electronic versions of SCPN inventory and monitoring reports available via the Internet. Each bibliographic record can be identified as available to the general public, all of NPS, or only the park unit from which it originates.

NPSpecies

The Service-wide species database for all NPS park units, it will be used to distribute species occurrence, abundance, residency, and nativity information. Each species occurrence can be identified as available to the general public, all of NPS, or only the park unit from which it originates. In addition, a species occurrence can be released to the general public, while withholding any general or specific geographic information.

STORET

SCPN water quality data will be made available to the general public through the Environmental Protection Agency's STORET (STORage and RETrieval) database. SCPN plans to upload certified data to STORET on an annual basis.

Biodiversity Data Store

The Service-wide digital archive of documents and spatial and tabular data files that document the presence/absence, distribution and/or abundance of any taxa in NPS park units.

Natural Resource and GIS Metadata and Data Store

The Service-wide online repository for metadata and associated spatial and tabular data products.

Southern Colorado Plateau Network Website

A repository for SCPN inventory and monitoring reports, summarized data, data subsets, and maps. The website's primary audience will be SCPN park staff and will include a mechanism for park staff to be notified by email when new products become available, to request specific data summaries or

products, and to provide feedback to SCPN staff.

10.2 NPS Access and Products

The primary location for NPS park managers and staff to obtain SCPN I&M information products will be the SCPN intranet web page; however, products will also be available via the websites listed above. The intranet web page will only be available to users who are accessing the internet from a computer on the NPS network. All reports and summary data products will be available at this site as soon as they are produced and meet the conditions above (Section 9.3). The webpage will include a mechanism for park users to sign up for notification of new products by email. At a user's request, they will receive an email indicating that the page was updated.

Each monitoring project will produce standard data summaries and interim reports on a set schedule. For projects with yearly data collection, these will most likely be produced annually. Projects with a longer data collection interval will have different schedules of interim product availability. These products will be posted on the website according to the schedule detailed in the project protocol. Some products that will be available on the SCPN intranet site include, but are not limited to:

- Interim project reports - reports produced after a data-collection cycle which summarize activities and data for a specified time period
- Trend reports - reports that may pull together data from multiple projects to report on long-term trends in a certain geographic region or topic area
- Summarized data - data summaries which are produced by automated database routines after a data-collection cycle
- Summary maps - maps portraying project information
- Data subsets - subsets of project databases, available as a static download, or as an interactive download with the user able to choose between several parameters to filter the data set
- Metadata - metadata for all project databases, allowing the user to determine whether or not to request access to the full database

At least initially, SCPN will not make full project databases available for immediate download over the Internet. Rather, park staff will be able to make a request for the full database and will be able to download it from a specified online repository upon receiving instructions. Park staff will also be able to request specific data summaries, queries or subsets from the full database. Requests will be via a form on the intranet web page and will be routed to the SCPN data manager.

10.3 Public Access and Products

According to FOIA (specifically the 1996 amendments), all information routinely requested must be made available to the public via reading rooms and/or the Internet. Network project data will be available to the public via the SCPN web site, and the public servers for NPSpecies, NatureBib, NPSFocus, Biodiversity Data Store, and the NR-GIS Data Store.

Products available at these locations will be the public versions of the products distributed to NPS park staff. In some cases, the NPS and public products will be the same. However, those information products which contain sensitive data will be edited prior to release to the publicly accessible web sites.

Full project databases will not be available for immediate download, but public users will be able to request access to these databases through a mechanism similar to that used by NPS staff. All requests will be evaluated, and the park(s) will be able to review the request prior to release. Data sets released to non-NPS individuals will be accompanied by a statement about the appropriate use and citation of data in resulting publications and a request that acknowledgement be given to the National Park Service Inventory and Monitoring Program in all resulting reports and publications.

10.4 Tracking and Feedback

Both the SCPN intranet and public distribution web pages will automatically track and log each time an information product is downloaded. At the very least, the log will contain information on how many times each product has been downloaded and the location and domain to which it was downloaded (by using the IP address). This will enable some non-intrusive tracking of who is downloading (and presumably using) certain products.

The forms to request entire project databases and specialized data products will provide another method of tracking information use. This will enable SCPN to evaluate, revise and change standard information products to meet certain demands.

In addition to the data request form, a general feedback form or contact email will be published on the web pages. This will allow any user to comment on the website layout, the availability of information products and the quality of these products.

Chapter 11 Archiving and Storage of Digital Products

This chapter describes procedures for the long-term management and maintenance of digital files that result from SCPN projects and activities. The overall goals of these procedures are to:

- avert the loss of information over time.
- ensure that network information can be easily obtained, shared, and properly interpreted by a broad range of users.

Digital products from SCPN I&M projects can include but are not limited to digital documents, database files, raw data from GPS units and data loggers, digital photos, digital images, and spatial data.

11.1 NPS and Network Standards

- The Southern Colorado Plateau Network will standardize on Microsoft products as required by all Department of Interior (DOI) agencies (DOI, Assistant Secretary for Policy, Management, and Budget, Findings and Determination, September 13, 2002).
- SCPN will update and maintain data sets to no more than two versions behind the current version of software, or store the data set in American Standard Code for Information Interchange (ASCII) format, complete with data and file documentation.

11.2 Electronic Archiving and Storage

Finite (short-term) projects

Electronic version of the deliverables will be electronically archived by three archival methods:

1. Online project archive on SCPN file server.
2. Offline project archive, files stored on removable media (backup tape, CD or DVD).
3. Digital libraries, where like products are stored together.

On-going (long-term) projects

Milestone products, produced on a schedule specific to each project, will be archived the same as short-term projects described above.

11.3 Electronic Archiving Process and Workflow

Electronic files should be consolidated and packaged for park electronic archival when a project is complete or reaches periodic milestones for ongoing projects. SCPN project managers are responsible for working with the data manager to prepare electronic files for packaging by ensuring the following:

- Files comply with file naming standards.
- Documentation is complete and follows standards.
- Draft and final products are clearly distinguished.
- Sensitive information is clearly designated and has been removed from public version.

The data manager will:

- Archive the project information on the centralized archive directory. Stored indefinitely.
- Complete a project backup on CD or DVD; media may be stored for 5 years.
- Integrate deliverables, such as final reports, into appropriate catalogs and electronic libraries, as described in this plan or in the monitoring protocols. Stored indefinitely.

11.4 Quality Control for Converted Data

Databases that are converted from one version of database software to an upgraded version will require additional quality control, particularly when the database applications are actively used for data entry or analysis. Forms, queries, reports, and data entry will be thoroughly tested during upgrades.

For data sets of limited use, the SCPN may decide to convert them to ASCII format. In this case, the data sets will no longer be upgraded. Complete documentation is necessary and should include table, field and object relationship descriptions. All ASCII files created from databases will undergo quality control activities or functions 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.

11.5 Digital Data Backups

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

This section provides a general description of backup procedures. Detailed backup procedures, software and hardware specifications, and tape rotation strategies are outlined in Appendix B, Backup Procedures

11.5.1 Program Data Backups

SCPN data files are currently stored on two servers, one for spatial data and one for all other data. Program data are stored on hard-drives set up in a RAID 1 configuration. A RAID 1 configuration automatically keeps two copies of the same data, one each on a separate disk drive. The drive is seen as a single drive.

In addition to file redundancy affected through RAID, tape backups are run on the following schedule:

Friday Night: Differential backup on all program data. Differential backups only copy those files that have been added or modified since the last backup.
Last Day of the Month: Full backup on all program data, tape archived for at least a year.

11.5.2 Personal Data Backups

A full backup of data that reside on desktop computers of networked SCPN staff occurs once a week. Each staff member also has a personal folder on the SCPN file server into which files can be copied.

11.5.3 Offsite Data Storage

Monthly backup tapes are stored offsite for at least one year before being replaced or moved back into the tape rotation as a weekly differential tape.

Chapter 12 Archiving and Storage of Physical Products

This chapter applies 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 SCPN administration, activities, and projects. This chapter also applies to physical items such as natural history specimens, photographs, and audio tapes. In most instances these documents and objects are essential companions to the digital data described in the previous chapter.

This chapter is not intended to provide a full description or procedure for archiving. For a complete description readers should refer to museum manuals and regulations. This chapter is intended to provide guidance in making the transition from completed products delivered to the project manager to a secure and long-term storage facility managed by park and regional curators.

12.1 NPS Standards

Direction for managing many of 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, that is, necessary for fulfillment of the NPS mission. 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 NPS-19 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.

The NPS Museum Handbook provides the overarching guidance for archival procedures. In particular, Part II, Appendix A: Mandates and Standards for NPS Museum Collections, lists the cultural and natural history laws, regulations and conventions for NPS museum collections and should be reviewed prior to object collections.

12.2 SCPN Standards

All specimens collected for the Network are property of the respective park in which they were collected. All specimens will be curated either by the park or a facility of the park's choosing, will be stored in a manner that meets NPS regulations, and will be properly cataloged into ANCS+, the NPS cataloging system.

12.3 Archiving Process

The project manager is responsible for obtaining an accession number and catalog numbers prior to beginning field work at the park if specimens are going to be collected. This is usually part of the research permit process.

When the project is complete or reaches periodic milestones, the project manager and data manager will work together to consolidate and package archival materials for delivery to the park curator or facility of choice.

The archival project package should contain the following items and documentation:

- A coversheet, memo, or Collection Catalog Worksheet listing contact information, project abstract and purpose, sensitivity and use of materials, and all materials included in the package.
- All materials clearly labeled with:
    - Park code
    - Date, or range of dates
    - Accession number
    - Project number
- Field notes on acid free paper (preferred) or copy paper, stored in a 3-ring binder or book box.
- Other paper materials such as reports and data printouts, stored on acid free paper (preferred) or copy paper and placed individually in an acid free, labeled folder.
- Photos should be printed and documented.
- Voucher specimens:
    - Properly labeled according to the Museum Handbook
    - Entered into Excel or text file in a format that can be imported into ANCS+
    - For specimens not residing in park archives, loan paper work is complete and a copy of the form is stored with the project package
- Archival of specimens must be addressed in Network Vital Signs Protocols and should specify:
    - Name and address of facility
    - Memorandum of Agreement, if applicable
    - List of items to be archived
- CD or DVD of all electronic materials. Jewel case should be labeled with:
    - Park Code
    - Date created
    - Range of dates for information
    - Accession Number
    - Project Number
    - List contents

A similar package should be prepared as a “working copy” for the Network and each affected park. These will be stored in the respective Network and Park Project file cabinets and will be available for day to day usage.

12.4 Network Working Files and Library

While a SCPN project is on-going, all relevant materials (reports, administrative documents, critical e-mails or correspondence, and data print outs) are stored in an organized cabinet in SCPN offices. Upon project completion, SCPN ensures that all final and pertinent project materials are submitted to the appropriate park for long-term storage and archiving. If the park lacks the resources to store and archive these materials, then SCPN stores them long-term in a centralized cabinet.

A separate file cabinet for the SCPN library will contain annual and final project reports and may contain additional reports that are related to inventories or vital signs monitoring. All materials in this library will be cataloged into the NatureBib database and will be citable. This library should not contain draft materials.

Appendices

Appendix A. Data Management Plan Revision Log
Appendix B. Backup Procedures
Appendix C. Operating Procedures for Uploading and Downloading Data to and from GPS Units (draft)
Appendix D. GIS Data Organization and Standard Operating Procedures for Data and Metadata Creation (draft)
Appendix E. Digital Photo Management (draft)
Appendix F. File Naming Conventions (to be developed)
Appendix G. Data Modeling
Appendix H. Standard Data Summary Reports (to be developed)
Appendix I. Exporting Data for Analysis (to be developed)
Appendix J. Metadata Creation for Non-spatial Data Sets

Appendix A. Data Management Plan Revision Log

Revision #	Date Issued	Chapter(s) & Section(s) Revised	Summary of Revisions	Author

Appendix B. Data Backup Procedures

Purpose

Regular backup of SCPN data and a means of storing backups offsite are necessary to ensure data are not lost due to hardware failure, user error or catastrophic events. This SOP describes the backup strategy and procedures for both program and staff data as well as tape rotation and storage schedule, backup schedule, procedures to set up backup jobs and procedures to perform backups.

Definitions and Acronyms

RAID 1 - Redundant Array of Independent Disks. Duplicate copies of all data are stored on two hard drives. This provides high data availability as two complete copies of the data are maintained on separate disks.
LTO - Linear Tape-Open. LTO technology is an "open format" technology, which allows users to have multiple sources of product and media. The "open" nature of LTO technology also provides a means of enabling compatibility between different vendors' offerings.
Program data - Commonly shared SCPN program files ranging from administrative records to data archives stored on the main file server.
Staff data - Files and data stored on individual SCPN staff computers.
Full backup – All files in the selected directories are backed up whether or not they have changed (been edited) since the last backup.
Differential backup – All files that have changed since the last backup are backed up.
Media set - A set of backup tapes that have the same overwrite protection and append periods and are used for the same backup type (i.e., daily differential).
Overwrite protection period – The period of time that must elapse before backup data on a tape can be overwritten.
Append period – The period of time during which backup data can be appended to a tape.
Selection list – The directories and/or files to be backed up.
USGS16 – SCPN work file server.
USGS71 – SCPN Data manager's computer.
USGS40 – SCPN spatial data file server and GIS Specialist's computer.

Backup Strategy

SCPN currently uses two servers, one for spatial data and one for non-spatial program data. Each server has a RAID 1 configuration which provides full disk mirroring. As a second form of redundancy and a means for storing backups offsite, each server is also regularly backed up to tape media.

Program data are backed up to tape weekly. A full backup is performed monthly on the last day of the month (or closest Friday); differential backups are performed weekly on Fridays. Sixteen tapes are used in the program data rotation schedule, providing a record of program data for the entire year. The tapes are managed in media sets by Backup Exec.

Staff data are backed up weekly using seven tapes in rotation. Staff computer backups occur Mondays through Thursdays. Two staff computers are scheduled to be backed up each evening.

Tape Backup Hardware

- 1 Hewlett-Packard Ultrium 2- SCSI Tape Drive
- 27 LTO 2 tapes, which have a storage capacity of 200 gigabytes uncompressed and 400 gigabytes compressed

Backup Software

- Veritas Backup Exec 9.1 for Windows Servers – referred to as Backup Exec throughout the rest of this document

Requirements & Procedures

1. A Backup Exec logon account with administrator privileges on all remote computers and the server it is running on is required.

2. The Backup Exec Remote Agent must be installed on each remote computer
   - From the remote computer temporarily map to USGS16\C:\Program Files\Veritas\Backup Exec\NT\Agents\RANT32 and type the following at the command prompt: setup.exe /RANT32: -s.

3. Create media sets in Backup Exec for the program data daily differential, weekly and monthly full backups, and the staff data weekly remote full backups. Label tapes and import into correct media set.

Media Set Name	Tapes (number; names)	Overwrite Protection Period	Append Period
Program – Weekly Differential	4; Week1 – Week4	6 days	27 days
Program - Monthly Full	12; Jan - Dec	infinite	infinite
Remote – Weekly Full	7; Tape1-Tape7	4 weeks	4 weeks

4. Create selection lists for program data and for each staff computer
   - Program data selection list:
     USGS16\H:\Common Files
     \C:\Program Files\Veritas
     \System State

     USGS71\D:\inventory_data
     \D:\npspecies
     \D:\Parks
     \D:\Project Data

     USGS40\GIS_Data

- Staff selection list should include Lotus Notes database, My Documents, Favorites, System State and any other folders where important files are stored (ask each staff member):

Staff Name	Computer Name
Lisa	USGS50
Rebecca	USGS83
Marguerite	USGS40
Megan	USGS84
Nicole	USGS71
Steve	USGS13
Chris	USGS69
Joan	USGS51

5. Create backup job templates in Backup Exec for each type of job. These allow you to create many jobs with the same backup parameters.

Job Template Name	Media Set	Schedule
Weekly Full Differential	Program – Weekly Differential	Fridays 7:00 pm
Monthly Full Backup	Program – Monthly Full	Last day of month, 7:00 pm
Monday Remote Weekly	Remote – Weekly Full	Mondays 6:00 pm
Tuesday Remote Weekly	Remote – Weekly Full	Tuesdays 6:00 pm
Wednesday Remote Weekly	Remote – Weekly Full	Wednesdays 6:00 pm
Thursday Remote Weekly	Remote – Weekly Full	Thursdays 6:00 pm

6. Create backup jobs from the backup job templates in Backup Exec
- Backup jobs are named after the staff member rather than the computer simply for ease of use. It is easier than having to remember or refer to a list of abstract computer names.

Job Name	Method	Media Set	Schedule
Program Data Weekly Differential	Differential	Program – Weekly Differential	Fridays 7:00 pm
Program Data Monthly Full	Full	Program – Monthly Full	Last day of month, 7:00 pm
Lisa	Full	Remote – Weekly Full	Mondays 6:00 pm
Rebecca	Full	Remote – Weekly Full	Mondays 6:00 pm
Marguerite	Full	Remote – Weekly Full	Tuesdays 6:00 pm
Megan	Full	Remote – Weekly Full	Tuesdays 6:00 pm
Nicole	Full	Remote – Weekly Full	Wednesdays 6:00 pm
Steve	Full	Remote – Weekly Full	Wednesdays 6:00 pm
Chris	Full	Remote – Weekly Full	Thursdays 6:00 pm
Joan	Full	Remote – Weekly Full	Thursdays 6:00 pm

7. Change tapes on correct schedule:

Monday — Remove Week X tape from drive insert Tape X (before 5:00 pm)

Friday — Remove Tape X and replace with Week X+1 tape (before 7:00 pm)

Last day of month — Remove Tape X (or Week X if the last day of the month is on a Saturday, Sunday, or Monday) and replace it with Monthly tape (before 7:00 pm)

First day of month — Remove Monthly tape and replace it with Tape X (before 5:00 pm)

Helpful Notes
The Backup Exec software has a schedule programmed into it—but it is more relative than absolute. For example, if no new tape is inserted for the first staff backups scheduled on Monday, it will wait until a tape is inserted—so backups can be run after the designated time if needed. If you forget to change the tape and want to run the backup the next day, you can. It is best to open the software and look at the Job Monitor tab to see which jobs are in waiting. Check the date and the time. The earliest scheduled job will be the one that starts next—not the one that has a date and time that comes after the current date and time. If a computer is not on, the job will fail and it will be taken out of line until the next week. It is up to the individual users to leave their computers on. You can also put jobs on hold if you want to run them out of order. To put a job on hold, go to the Job Monitor tab, right click on the particular job, and check "hold schedule."

Running a backup during the day if it has been forgotten the night before is ok, but it does slow the network down considerably.

Using the Remote Desktop Connection to login to the file server is the easiest option for checking backup jobs. Under the Start Menu, choose All Programs >> Accessories >> Communications >> Remote Desktop Connection. The file server is USGS16 so type that into the box that says "computer". Now type your domain user name and password (the one you use to log on to your computer everyday). You are now connected to the server.

Changing the backup administrator password in Backup Exec.
Every time you are asked by the USGS domain to change your password, you must also change the information stored by the Veritas Backup Exec software. The password you use for Backup Exec should be the same as your domain login. The reason for this is because your domain user name has been given administrative rights on all of the network computers. If the Backup Exec Software does not have the correct password stored with your user name, the backup will not be able to connect to the remote computer and it will fail. You need to change your user name in two places in the Veritas Backup Software: Network >>Logon Accounts, and Tools >> Backup Services>>Service Credentials.

Quick tour of the Software
Veritas Back-up Exec is organized with a series of tabs: Backup, Restore, Overview, Job Setup, Job Monitor, Alerts, Reports, Devices, and Media.

Backup: Hit this button if you wish to create a new backup job.

Restore: If data are lost from a computer, this is the button you would hit to get the data back from the tapes. You will likely want to check the radial button that says "View by Resource." Pick the computer or folder you want to restore from the list on the right and it will give you a list of media tapes that have data from that computer. Choose the desired one and hit "run now".

Overview: Just an overview of the software—has links to various help menus.

Job Setup: Along the top, you find each of the backup jobs. The ones called "____ Test" are what you run if the backup software is having trouble communicating with a particular computer. Running the "test" jobs tells you whether you have a problem with permissions or passwords. To run a test, double click on the desired test job, click on the word "schedule" on the right side, and select the radial button that says "run now".

Below the Jobs list is the Template list. Each backup job is based on a template. Finally, you have the Selection Lists. If you want to add a particular folder on a computer to be backed up, you modify the selection lists. This works similar to Windows Explorer.

Job Monitor: This is where you look to see if jobs were successful or not. The top of the screen lists the jobs and their scheduled dates, the bottom lists the history. You can double click on a job history to see why it failed. If you read the log and it says "Unable to attach to one of the drives," it likely

means that the computer was turned off. Another common message is “Access is denied to the remote agent”. This message means that the permissions are not correct—either the person responsible for backups does not have administrative privileges on that computer, or he has forgotten to change his password in the backup software.

Alerts: Lists alert histories. The most common is that a tape needs to be inserted for a job to continue.

Reports: Lists various reports you can run.

Devices: Lists the hardware devices used. HP1 is our backup device. You can also erase or eject media from this screen.

Media: Lists all the active tapes. If you are unsure of the numbered tape to insert, you can use this tab to figure it out. Click on “Remotes Weekly” on the left-hand side, the tapes in blue on the right-hand side are those ready for use. Select the tape that is next in line after the last tape in black. For example, if tape 1 is in blue, but tapes 2 and 3 are in black, the next tape in the sequence would be tape 4 (not tape 1). The same can be done for the weekly tapes.

Revisions

Revision	Description of Change	Author	Effective Date
1	Staff data backup now occurs Mon-Thur.	N. Tancreto	December 12, 2005
2	Updated to match current backup scheme. Added Helpful Notes, Changing backup administrator password, and Quick tour of software sections.	N. Tancreto & M. Hendrie	September 2006

Appendix C. Operating Procedures for Uploading and Downloading Data to and from GPS Units

DATA ARCHIVE LOCATIONS

Every SCPN project will have a folder on the "Common Drive" of the main server. Subfolders under each of the project folders are as follows: analysis, data, documents, images, and spatial information. Information downloaded from the GPS units should be placed in the spatial information folder.

NAMING CONVENTIONS

Files downloaded from GPS units should be given unique names based on the project, park, and date of transfer. In order to avoid software compatibility issues, spaces and dashes should be avoided whenever possible. Project codes should be at least four and no more than six letters. To construct a project code, use the first letter in each word of the project title. Add additional letters starting with the first word until a four letter code is reached (e.g. if the project title is Air Quality, use AQ, then go back and add the second letter of Air and the second letter of Quality = AIQU; Habitat-based bird communities would be HBBC). If more than four letters are needed for clarity or to distinguish between two different projects, add them starting with the first word as indicated above. The four-letter park codes should be used for each park and the date format is MMDDYYYY. The following would be a file name for the Integrated Riparian project at Glen Canyon downloaded April 5, 2005: INRI_GLCA_04052005.

GARMIN UNITS

The Garmin units use Garmin MapSource software and a serial port connection.

Downloading Data from the GPS Unit

1. Connect the GPS unit to the computer using the 4-pin connector.
2. Turn on the GPS unit.
3. Open MapSource Software.
4. Under Edit → Preferences you can select the datum and grid system you wish to use. For most projects we will use Position = UTM and Datum = NAD 83. The units tab should have values set to metric.
5. Select File →Open from Device. Select Serial Port, and select the items you wish to open.
6. Your waypoints will display on the map. You may zoom in using the magnifying glass icons.
7. Select File →Save as –and choose .txt from the pull-down menu to export your waypoints as a text file. If you want to store the waypoints for use in mapsource (i.e. you will be loading them back onto the GPS unit at a later date) Select File→Save as and select the default "Garmin database" file extension (.gdb). If there were problems or errors during data transfer, indicate such matters in the header of the txt file.
8. The text files can be imported into Excel for easier viewing or for import into GIS software applications. Open Excel and select Data → Import External Data → Import Data. Under the "files of type" pull-down menu choose "all files." Browse to the .txt file you created and double click on the name. The text import wizard will open. Choose the radial button that says "Delimited" and start importing at the row that contains the column headings. This will skip a lot of the header information, but will make GIS import easier. Accept the other defaults and click next. The next screen will ask for delimiters. Choose tab as the delimiter and check the box that says "Treat consecutive delimiters as one." You can select none for the text qualifier. Hit next and accept the defaults. Click "Finish". Accept the default location for where to put the data.
9. The x and y coordinates are now located in the same column, and this needs to be fixed in order to import data into a GIS system. The easiest way to do this is to insert three new columns to the right of the "position" column (Select the column to the right of the position column and Choose Insert → Columns from the menu). Now select the cells with the combined xy information. With those cells selected Click Data →Text to columns. These menus will look much like the import text wizard. Choose delimited as the file type and click

next. Now choose the space radial button and click finish. You will now have a column for the UTM zone number, the UTM hemisphere designation, x coordinates, and y coordinates. Make sure to provide a heading for these columns. Choose File → Save as and save this file as a dbf 4 file (using the above naming conventions). Just say yes to the following dialog boxes. Now the GPS information can easily be imported into a GIS.

Uploading Data onto the GPS Unit

1. Connect the data cable (4-pin connector) from the Garmin unit to the serial port on your computer.
2. Turn on the GPS unit by pressing the power button.
3. Press the menu button twice and scroll to Setup. Scroll over to the Interface tab and make sure the Serial Data format is set to "Garmin."
4. Open MapSource Software on your computer.
5. Open the file that you wish to download to the unit (waypoints, maps, etc.).
6. Select File→Save to Device and choose the formats of the files to upload. Click Save.

TRIMBLE UNITS

Trimble units are connected to the computer via the Trimble cradle or serial clip and a USB port on the computer. In order to download data to the computer, the user needs to have Microsoft Active Sync installed on the computer. Microsoft Active Sync is free for Microsoft users and can be downloaded from the following location: http://www.microsoft.com/windowsmobile/downloads/activesync38.mspx. Follow the instructions for set up and connection to the mobile device.

The SCPN will use ArcGIS 9.0 with the Trimble GPS Analyst Extension for downloading GPS data collected on the Trimble units. The GPS Analyst software CD also comes with a program called "Trimble Data Transfer" which will also need to be installed on the computer.

Make sure the GPS unit is in the cradle, turned on, and connected to the computer via the USB port before starting the data transfer process. Windows Active Sync should automatically detect the device. Set it up as a guest under the Active Sync dialog box.

Downloading Data from the GPS Unit

1. Navigate to Programs →Trimble Data Transfer →Data Transfer. Click on the button that says Devices…. Click on the New button and select GIS Datalogger on Windows CE. Click ok and specify a name if you wish. Click Finish. Now close the devices window. The Windows CE device should now appear in the pull-down window under Devices. Click the button to the right with the green check mark in order to connect to the device.
2. Make sure you are in the Receive tab and click on the Settings… button. Select all of the check boxes and click on OK. Now click on the Add button and choose data file from the pull-down menu. Select all the data files created during the last sampling session and browse to the spatial data folder under the appropriate project. Click on the Open button. Now click on the Transfer All button. These files will be saved with an .SSF extension. When the transfer is finished, click on the Close button.
3. Differential correction of GPS data should be done as close to the date of collection as possible. To differentially correct the Trimble data we must import it into an ArcGIS geodatabase. In ArcCatalog, navigate to the spatial data folder under the appropriate project. Right click on the folder and select New → Personal geodatabase. Give the geodatabase a name using the project and park code. Right click on the personal geodatabase and select Properties. Navigate to the GPS analyst tab and click on "GPS- enable geodatabase". You will now need to set the spatial domain and projection. Once a spatial domain is set, it cannot be changed. It is important to use one that is big enough to include the extent of your data. Use the following table to guide the spatial domain setting. All projections should be UTM NAD 83 zone 12 or 13N.

Park or Area	UTM Zone	Min x	Min y	Max x	Max y	Comments
AZRU	13N	223300	4073300	241400	4088100	Divide between zone 12 and 13 goes through park
BAND	13N	355500	3937800	410900	3983900	
CACH	12N	612000	3967000	696000	4033000	
CHCU	13N	188300	3941700	280400	4020300	
ELMA	12N	724000	3830000	815000	3903000	Divide between zone 12 and 13 goes through park, and these coordinates are big enough to also include ELMO
ELMO	12N	732500	3868000	760000	3894000	
GLCA	12N	392800	4045100	652000	4294000	Coordinates also include RABR
GRCA	12N	200000	3936500	447500	4105000	
HUTR	12N	621600	3945000	639600	3959700	
MEVE	12N	697000	4105500	748000	4144600	Coordinates also include YUHO
NAVA	12N	499600	4031800	572800	4087800	
PEFO	12N	547100	3811300	674800	3931900	
PETR	13N	298900	3855100	387000	3925800	
SAPU	13N	324000	3745700	461400	3863500	
SUCR	12N	434000	3899000	473800	3927400	
WACA	12N	434300	3877900	474000	3906200	
WUPA	12N	441800	3920200	487100	3952100	
YUHO	12N	689800	4108800	724800	4138700	
Entire Network	12N	134000	3643300	1144000	4346000	
Entire Network	13N	-459600	3612100	630700	4444000	

Spatial extent and precision are inversely related. The above values represent a spatial extent that encompasses the area of the park plus a small buffer. If you are worried about too small a reference, use the “Entire Network” values or the default values given when setting the spatial extent. This will simply mean that the precision of your points when you zoom out will be less.

4. Click on the “Set spatial reference” button. In the New Spatial Reference Window, click on the Select button. Open the Projected Coordinate Systems → UTM → NAD 83 and pick UTM Zone 12 or 13 N depending on the park. Now click Next. The spatial domain is set in the following window. Pick the spatial domain by looking at the appropriate park’s extent above or in another GIS file with the same projection. Make sure your GPS coordinates fall within this spatial reference.
5. Click next and specify the z value. It is a good practice to set the minimum z value to less than zero (-1000) although this is not a big issue in this area. Now click Finish. Know the geographic transformation must be selected because all the GPS data are collected in GCS_WGS_1984. ESRI recommends the NAD_1983_to_WGS_1984_4 to switch between NAD 83 and WGS 84. Click ok. Now click ok to exit the database properties dialog box.
6. Right-click the geodatabase and select New Feature Dataset. Enter the name in the name field and click edit. Set the spatial reference and the extent. (Should be the same as the geodatabase). Click OK.
7. Right-click on the feature dataset you have just created and select Import Trimble SSF. Navigate to the SSF file that you have downloaded from the GPS unit. Click OK. A summary screen will show the number of files, GPS positions and features.

8. Right click the geodatabase and click properties. In the GPS Analyst tab, click the Feature Class Properties button. Specify a name for the field in which the average and worst estimated accuracy for each feature will be stored. If you wish to specify the accuracy required for validation, enter the value in the bottom window.
9. To differentially correct the features, open ArcMap and add the newly created feature dataset. The feature dataset will have point, line, and area features and a GPS feature. From the GPS Analyst drop-down menu, select Start GPS Editing. Click on the differential correction button on the toolbar. The first page of the wizard will ask for the GPS session from which you would like to correct positions. GPS sessions are stored in the personal geodatabases created for the import session. Add the appropriate session in this window and the following window and click add. Click next through the next dialog window. Now the default correction settings summary will appear. Click next through this window. Select Base Provider Search in the next window and click "update list". A list of nearby base stations will appear. Select the closest station with an integrity index of greater than 80.0 if possible. Click ok and accept the rest of the defaults. The rebuild settings summary will appear. Click Change… and specify the best quality GPS positions to be used. If a minimum accuracy has been chosen for this project, input this value. When you click Start, the rebuilding will begin. New features will be created based on the previous specifications. These settings can be changed later if the results are unsatisfactory. Click on GPS Analyst → Save Edits followed by GPS Analyst → Stop Editing.

Uploading Data to the GPS Unit

1. Trimble .SSF or .COR files can be uploaded onto the Trimble unit using the Trimble Data Transfer Program. Any GPS enabled geodatabase feature class can be exported as an .SSF file by right clicking on the dataset in ArcCatalog and choosing Export. From the Export menu, choose "To Trimble SSF."
2. Open the Data Transfer program and connect to the unit using the same procedure as outlined in the download section. Click on the Send tab. Click the Add button, and choose Data File. Navigate to the desired SSF file. Click on the Transfer All button.

Revisions

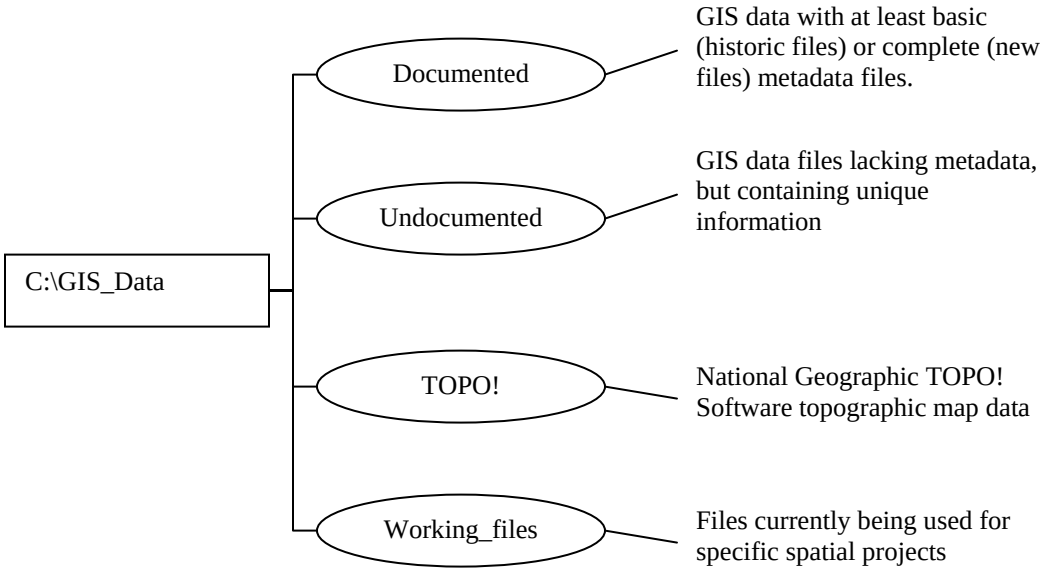
Revision	Description of Change	Author	Effective Date

Appendix D. GIS Data Organization and Standard Operating Procedures for Data and Metadata Creation

Data Storage and Directory Structure

All GIS data held by the Southern Colorado Plateau I&M Network is located on the Dell Precision Workstation 670 in the office of the Assistant Data Manager. This machine has a Level 1 Redundant Array of Independent Disks (RAID) configuration which provides full disk mirroring. This disk is differentially backed-up to the SCPN server on a daily basis with a full back-up occurring weekly.

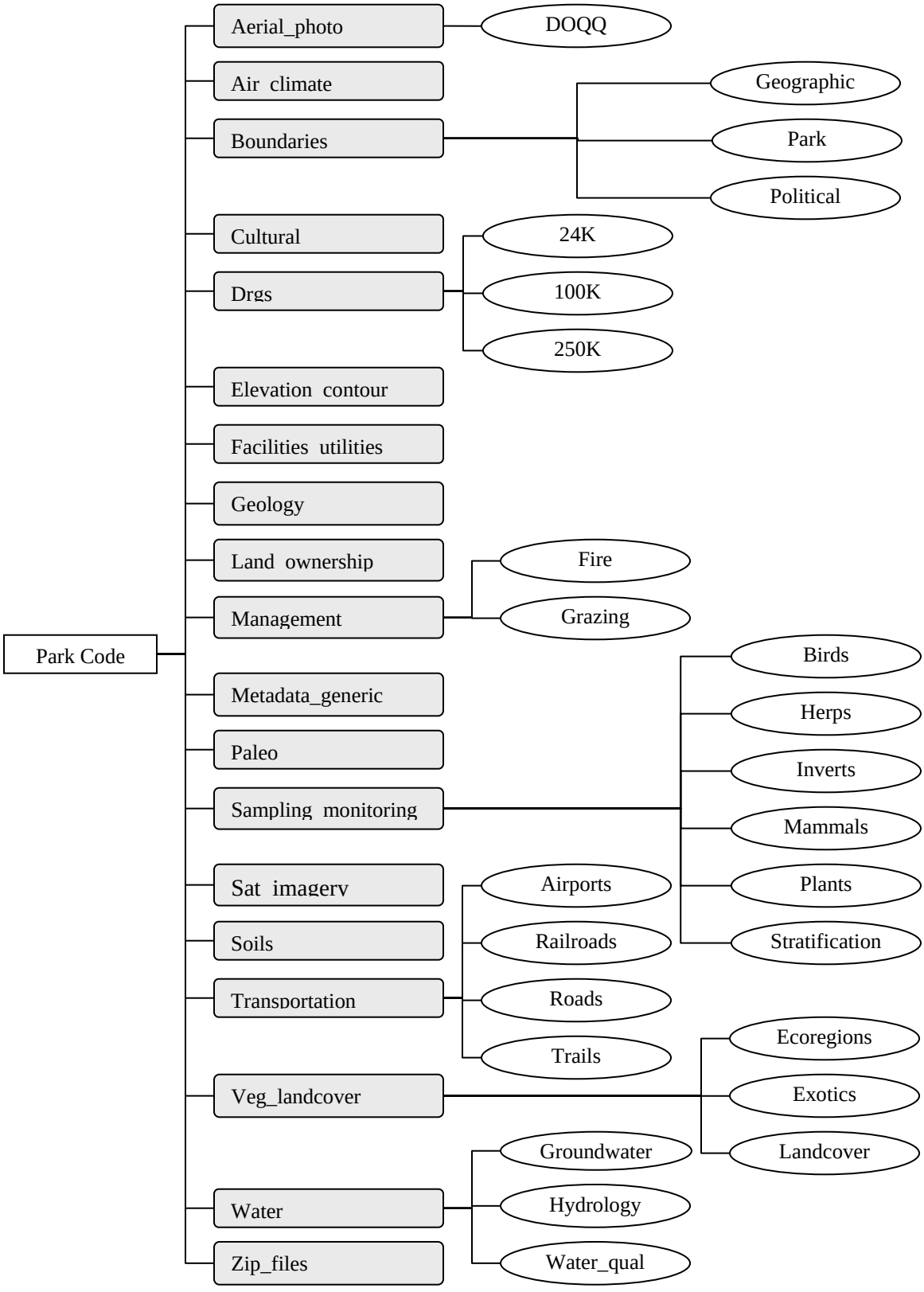
The basic directory structure of the GIS data is as follows:



Data in the “Documented” folder will be in NAD 83 UTM projection and will serve as the source of information for any I&M project work. The undocumented GIS data folder is a repository for historic GIS files that have not been matched to a report or other source of metadata. Files in this folder contain information that can not be found in other GIS files and should be used solely as a reference. These files should never be used to create new GIS layers or to produce maps for publication. The Working_files folder contains GIS files that are being manipulated, altered, or combined to create new GIS layers for the purposes of inventory and monitoring projects.

The Documented and Undocumented folders will share the following file structure.

Directory structure for Documented and Undocumented GIS folders



In addition to a folder dedicated to each park, a regional folder will also be present in both the Documented and Undocumented directories that will contain region-wide or state-wide data. The subfolder themes will be similar to those found in the individual park folders.

All data in the Documented and Undocumented directories will be cataloged in an Access database. This database will provide basic information about the data including location of the file, a link to the metadata file, format of the data, projection, datum, thematic keywords, and geographic keywords. By providing links to the metadata files, the database user will have ready access to additional information about the dataset.

Each SCPN employee will have his own subfolder in the Working_Files folder. The employee folders will be further divided by park. Folders based on the project name will be created by the employee under the appropriate park folder. At the same level as the individual employee's folder will be a folder called "Check_In".

Creation of New GIS Data Layers
Many new and intermediate data layers will be created during the implementation of various projects. New data layers that will be of use in the future for other projects or for documentation of methods will eventually be cataloged in the Access database and archived in an appropriate location in the "Documented" directory. Before moving any GIS layers, a metadata file must be created.

To create a metadata file, we will use ArcCatalog and the NPS Metadata Extension. To download the NPS Metadata Extension, visit the NPS Natural Resource GIS website, http://science.nature.nps.gov/nrgis/tools/editor.cfm. Follow the instructions for downloading and installing the extension.

Guidance for creating FGDC compliant metadata and using the National Biological Data Profile can be found in the following documents located on the SCPN server in the _____folder.

Federal Geographic Data Committee. 2000. Content Standard for Digital Geospatial Metadata Workbook Version 2.0. Federal Geographic Data Committee. Washington, D.C.

FGDC Biological Data Working Group, and USGS Biological Resources Division. 1999. Content Standard for Digital Geospatial Metadata – Biological Data Profile, FGDC-STD-001.1-1999 Federal Geographic Data Committee. Washington, D.C.

All metadata created by the SCPN should follow the National Park Service Metadata Profile. In addition, metadata for data concerning plants or animals should include information found in the National Biological Data Profile. The NPS_DataStandard and the NPS_BioProfile stylesheets in the NPS Metadata editor should be used in order to ensure all necessary information is present in the metadata record.

Once a complete metadata record is created for a dataset, the data should be copied and placed in the Check_In folder of the Working_files directory. Once per week, the Assistant Data Manager will review the files in the CheckIn folder and verify the completeness of the metadata record. Once the metadata passes inspection, the dataset will be cataloged in the Access database and filed in the appropriate folder in the Documented directory.

Appendix E. Standard Operating Procedures for Digital Photo Documentation and Archiving

Introduction

This document is designed to provide guidance for the cataloging and storage of digital photos taken by employees and cooperators of the SCPN Network. Instructions on standards and procedures for use of a digital camera will not be covered in this document, but should be covered in project-specific Standard Operating Procedures. The SCPN has chosen digital photography as the preferred method of documentation, as such; this document will not cover archiving slide or film photography. Scanned images of slides or film photography will be covered in this document.

Basic Field Photography Guidelines

While we will not cover specifics of digital photography in this document, we will ask that staff and cooperators follow a few simple guidelines in the field.

1. Make sure to verify the date and time on the digital camera with every use. Digital cameras store this information in the EXIF file associated with the picture. Poor date/time calibration will result in bad metadata.

2. Detailed notes on subject, location, date, time of day, identification of photographer, and purpose of photo should be taken every time a photo is taken in the field.

3. Digital cameras have made quality control simple and instantaneous. All digital photos should be reviewed on the viewfinder to ensure the photo meets all the desired standards. Photos that fall short of standards should be deleted and retaken when possible. If the opportunity has been missed (e.g. the animal ran off) sub-standard photos should be reviewed on a case-by-case basis. If there is no value to the photo, it should be deleted.

4. We recommend the use of a camera with 5 Megapixels or more and storing pictures with the maximum amount of resolution allowable given specific storage space limitations.

Overview of Organization Scheme

To simplify organization, we will divide all digital photos into three categories: Communication & Outreach, Data, and Special Documentation.

- **Communication and Outreach** includes photos of landscapes or scenery, plants or animals not related to a specific project, people, archaeological sites, or other topics that might be used in brochures, newsletters, websites, or other non-technical communication media.

- **Data** photos are those that will be treated as part of the data collected for a particular project. These may be photo point, voucher, or documentation photos. These photos will likely be linked to the project-specific databases.

- **Special Documentation** is a category for photos that are not related to a specific project, but were taken for the purpose of documenting a species, event, or disturbance. These might be photos taken by field workers working on projects for which this documentation has no bearing. Examples include photos of trespass livestock or rare or endangered species. Photos in this category have a value beyond aesthetic appeal, but will not be linked to a project-specific database.

Each category of photos will have a slightly different filing scheme. We will try to maintain storage efficiency whenever possible, but there may be cases in which a photo belongs in more than one

category. For example a photo point picture may be used in a web page or newsletter. In these situations, it is acceptable to place a copy of the original photo in another folder for ease of use.

Directory Structure

Originals of all photos should be stored on the SCPN main server. Storage at this location ensures the maximum back-up routines will be used.

Communication and Outreach Photos

- SCPN Server
  - Common_Drive
    - Digital_Library
      - Photos
        - Communication_Outreach
          - Landscapes_Scenery
          - Field_Work
          - Flora
          - Fauna
          - Archaeological_Sites
          - Meetings_Staff
          - Equipment_Facilities
          - Water

Special Documentation Photos

- SCPN Server
  - Common_Drive
    - Digital_Library
      - Photos
        - Special_Documentation
          - [Park Code]
            - Disturbance
            - Rare_Sighting
            - [Other]

Data Photos

- SCPN Server
  - Common_Drive
    - I&M_Projects
      - [Project Category (e.g. Inventory, Monitoring, etc)]
        - [Project Name]
          - Images

Each photo storage folder will have a subfolder called *Edited.* No image file should be edited without first making a copy and placing it in the *Edited* folder.

Camera Download Procedures

Photo downloads should take place as soon as possible after returning from the field. In order to minimize the amount of extraneous software on network computers, the preferred method of

downloading photos from the camera to the computer is through the use of a USB card reader. The SCPN Network will use ThumbsPlus 7.0 © 1993-2004 Cerious Software, Inc. as our photo cataloging software.

1. Remove the memory card from the camera and insert it into the card reader. Insert the card reader into an open USB port on the computer. The computer should eventually recognize the card reader as an external drive.

2. Open the ThumbsPlus 7.0 © software. It should be configured to open the main SCPN database at startup. If this is not the case, choose Options→Preferences and select the Startup tab. In the Initial Database section, click the radial for "specific database". Browse to the main SCPN database located Server:\Common_Drive\Digital_Library\Photos\SCPN_Master_Photo.td4. Click on the General tab. Make sure the following boxes are checked: Allow folder operations…, Use Windows recycle bin settings, Enable tracking of file and folder changes…, and Retain original file date/time …. Other settings in this tab are up to the individual user. In the Thumbnails tab, select the radial that says "For new thumbnails only" under "Automatically copy image comments to database annotations:". Finally, in the Keywords tab, check only the boxes that say "Folder names in path" and IPTC keywords. Browse through the other tabs to set settings based on user preferences. Click Apply and Close. Close ThumbsPlus and restart it.

3. The left-hand side of screen will have a directory tree similar to Windows Explorer. The card reader should show up as one of the drives. Click on the card reader and navigate to the folder in which the photos are stored. When you highlight this folder, the software will immediately begin creating thumbnails for all of the images in the folder. Allow this process to complete before proceeding. Select images you desire to move by clicking on the thumbnail (multiple images can be selected using the Shift and/or Ctrl keys). Once the correct images have been selected, choose File→Move… and browse to the desired folder using the above directory structure. It is also possible to click and drag the thumbnails, but this will copy the files rather than move them from the original location. Pressing the Shift key while dragging will move files and delete them from the original location. It is important to delete the photos from the card after they have been moved to the appropriate folders (if not done automatically).

4. Any thumbnails oriented improperly, can be rotated or flipped at this point using Thumbnail → Orient Thumbnail. This is not considered editing as it just effects the thumbnail, not the image.

Scanning Procedures

In the future, SCPN projects may involve the use of scanned photographs. It is difficult to anticipate all possible scenarios for which photographic scanning might be employed. It will fall to individual project managers to work with the SCPN data management staff to set the standard operating procedures for scanning photographs.

Naming Conventions

Communication and Outreach and Special Documentation Photos

Each photo name will have at least 4 parts: park or network code, subject, date, and sequential number. Names should not include spaces in order to avoid conflict with various software programs. Ideally photo names will be less than thirty characters. Standard 4-letter park codes will be used, and the date format is as follows DDMMYY. Examples using this naming convention are as follows:

BAND_FrijCan_031005_01 This is a photo of Frijoles Canyon at Bandelier National Monument taken 3 October 2005.

SCPN_Staff_020402_04 This is the fourth photo in a series of SCPN staff taken 2 April 2002.

If the photo is related to a project, the project code should be used in the description. All photos will have detailed metadata, so the photo names need not be extremely detailed. If a photo is edited, the suffix "ed" plus a sequential number should be added to the name.

Data Photos
Each project will set up its own naming convention for data photos. All data photo names should contain the project code.

After the photos have been moved to the proper location, each image should be given a new name. Using ThumbsPlus, it is possible to batch rename a set of photos. Select the group of photos that will have similar names and choose File→Auto Rename. A dialog box will appear. Enter the desired prefix in the Prefix box (e.g. BAND_FrijCan_031005_). In the Next # Box enter the sequence number with which you wish to start. Normally, the increment number will be set at 1. Choose the number of digits in the Leading Zeros section depending on the number of photos in the sequence. Hit Ok.

Metadata

Each photo (digital or scanned) will be required to have metadata associated. Metadata should be added immediately upon downloading or scanning a photo. The minimum metadata fields are as follows:

Original Date/Time	(Automatically saved by camera)
Native format of original	(Digital, Print, Slide, or Negative)
Storage location of original	(Where is digital or hardcopy located)
Description	(Memo field up to 255 characters)
Photographer last name	
Photographer first name	
Collection name	(If this was scanned from a special collection)
Project code	
Park code	
Subject	(Main subject of photo or names of people)
Place	(Not just park code—but general location)
Editing procedures	(Memo field up to 255 characters)
Editor last name	
Editor first name	
Equipment make	(Either camera or scanner)
Equipment model	
Keywords	(Words other than the project and park code up to 32 keywords/photo)

Data photos will have additional metadata requirements that will be set by the individual project managers.

The metadata fields listed above will be included in the general SCPN photo database. To enter metadata for a photo the procedures are as follows:

1. To assign the same keywords to multiple photos, select all the desired photos and right click while the pointer is over one of them. From the menu, select Assign keywords→Other. You may type in new keywords or select keywords that have already been used in the database (we suggest doing the latter first before new keywords are added). You can add additional keywords to individual photos in the next section.

2. Right click on the photo for which you need to enter metadata and select Properties. There will be 7 tabs. The first three tabs are filled with fields that the software automatically collects. The EXIF information embedded in the image by the camera (for digital pictures only) contains useful information about focal length, exposure, and other photography specifics. This information is found in the Info tab. In the Database tab, type in the description of the photo in the Annotation box. This should be similar to a caption on a photo. The keywords tab will have all the keywords that have been selected for this photo. The database is set to use folder names as keywords by default. At times this may result in odd keywords. Please delete these unwanted keywords.

3. The User Fields tab has the bulk of the metadata fields. Click the check box to edit a field and click Save to save the information. Enter all relevant data in the appropriate fields. Most of the fields offer the option of a pull-down menu so you won't have to type your name multiple times. After this is complete, click on the close button to exit or the >> button to go to the next photo.

Modifications to the database structure should be done only by the SCPN data management staff.

Editing Photos

Editing photos is recognized as an unavoidable procedure. It is very important to always copy a photo that is to be edited. Photos should be copied using ThumbsPlus so that associated metadata will follow the copy. Photos destined for editing should be copied and placed in the Edited subfolder of each photo folder. Intermediate or bad edits can be deleted from this folder at the editor's discretion.

ThumbsPlus has its own photo editing software. Minor edits such as cropping and orientation can be done by double clicking a thumbnail in ThumbsPlus. The Transform menu contains these basic operations. You may also reorient a picture that is slightly askance by using the Transform → Rotate to Line function. You must choose File→Save As in order to save these edits. The SCPN Network also has a copy of Adobe Photoshop for advanced editing. Notes on editing procedure should be always be taken and documented in ThumbsPlus.

Data photos should never be edited unless specifically instructed to do so by the project leader.

Special Notes on Data Photos

Information entered into the ThumbsPlus database for data photos will be exported and linked to project databases by the data management staff to avoid the need to re-enter information. This procedure will be set by the SCPN Data Manager and the project leader and will be included in individual project protocols.

GPS Information

There are a few software packages available that enable a user to link GPS coordinates to digital camera photos automatically in the field. This procedure involves using the GPS "Track Log" function and synchronizing the cameras date/time with the GPS unit date/time. At this time, the SCPN does not possess a GPS-photo link software package, but one may be made available if the need arises. For linking GPS coordinates, we advocate the use of a GPS unit to mark a point and to enter the coordinates into the project-specific photo database.

Appendix F. File Naming Conventions

To be developed…

Appendix G. Data Modeling

A three-phased data modeling process will be used to design relational databases.

Conceptual Data Model

A conceptual data model is a general model depicting the entities and their relationships to each other, but it does not include attributes or any information about database structure (Figure G-1). It is simply a model of the primary data objects necessary for the project. A conceptual model may also contain a description of how the data will be collected and processed. There are no rigid requirements for conceptual data models; the goal is to delineate the necessary data elements, understand their relationships and describe how the data will be collected in the field. Conceptual data models are tools for the project manager and data manager to ensure mutual understanding of the data needs of the project. They may be discarded once the logical data model is developed, or skipped altogether for simpler projects.

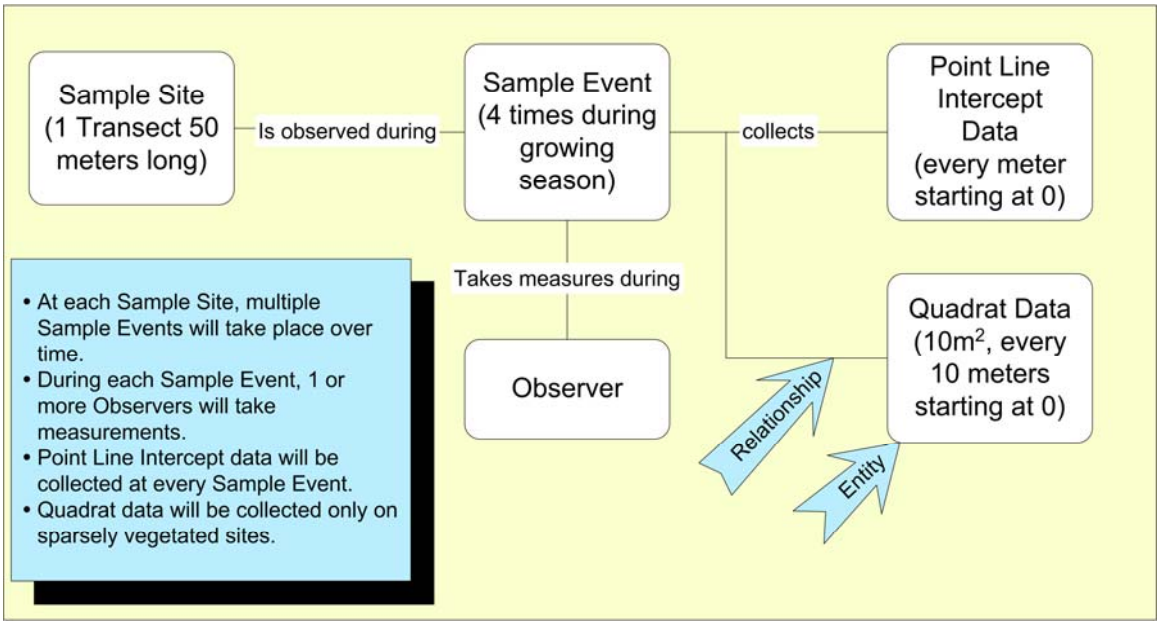


Figure G-1. Simplified example of a conceptual data model.

Logical Data Model

A logical data model is a more detailed exploration of the data entities, their attributes, and relationships (Figure G-2). It is created independent of technology considerations and is a precursor to the physical data model. In the logical model ERD, all attributes are identified and relationships are described in more detail. A primary key and foreign key for each entity should be identified. Entities and attributes are described using clear terminology, rather than names formatted for database software. Relationship cardinality and hierarchies will be delineated.

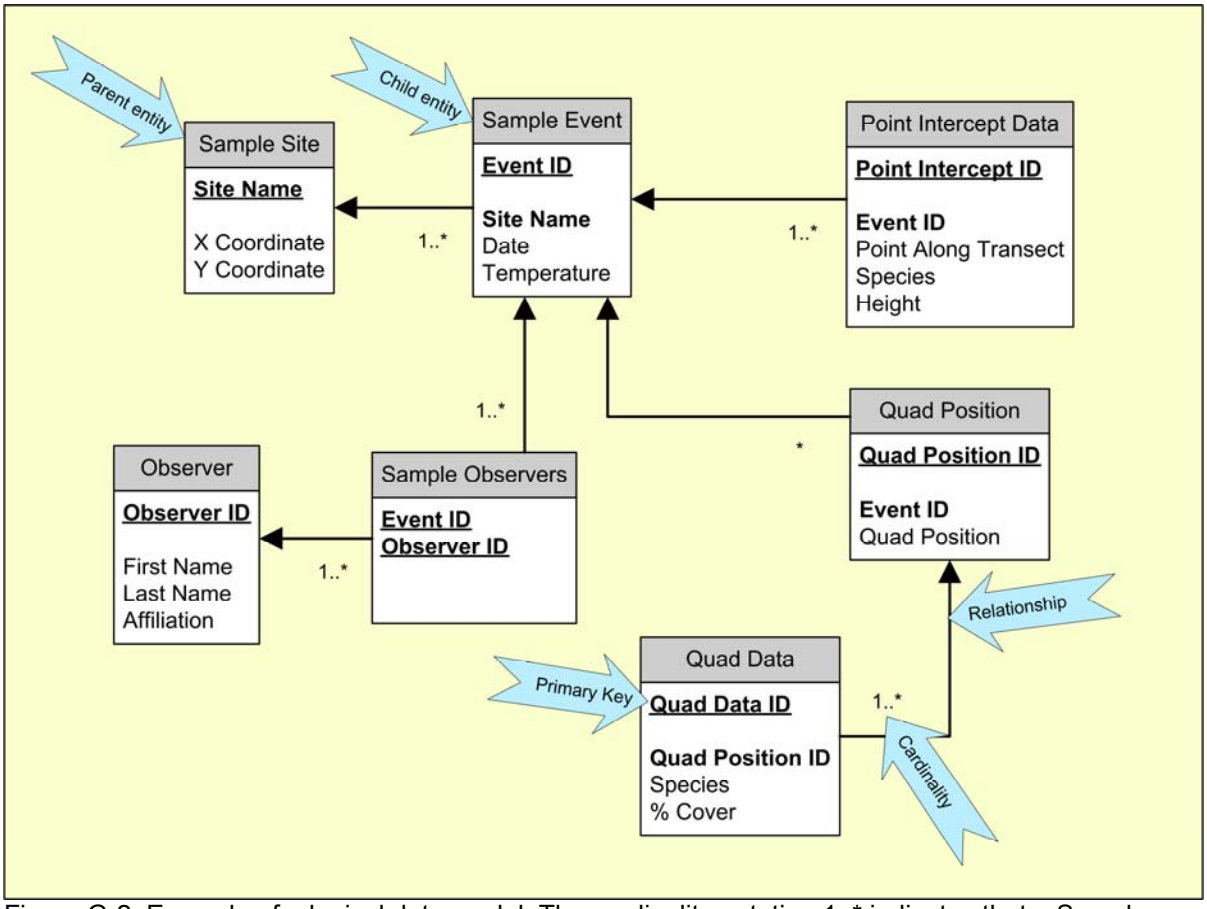


Figure G-2. Example of a logical data model. The cardinality notation 1..* indicates that a Sample Site will have at least one sample event and may have multiple events. The notation * on the relationship between Sample Event and Quad Position indicates that each Sample Event may have zero or more Quad Positions, representing that quadrat sampling is optional and may or may not occur during each sample event.

Physical Data Model

The physical data model is used to design the actual database, depicting data tables, fields and definitions, and relationships between tables (Figure G-3). Though the logical and physical data models are similar, the logical data model only provides enough detail to communicate the information to be stored in the database. The physical data model provides very specific details and definitions, such as primary keys, foreign keys and field types.

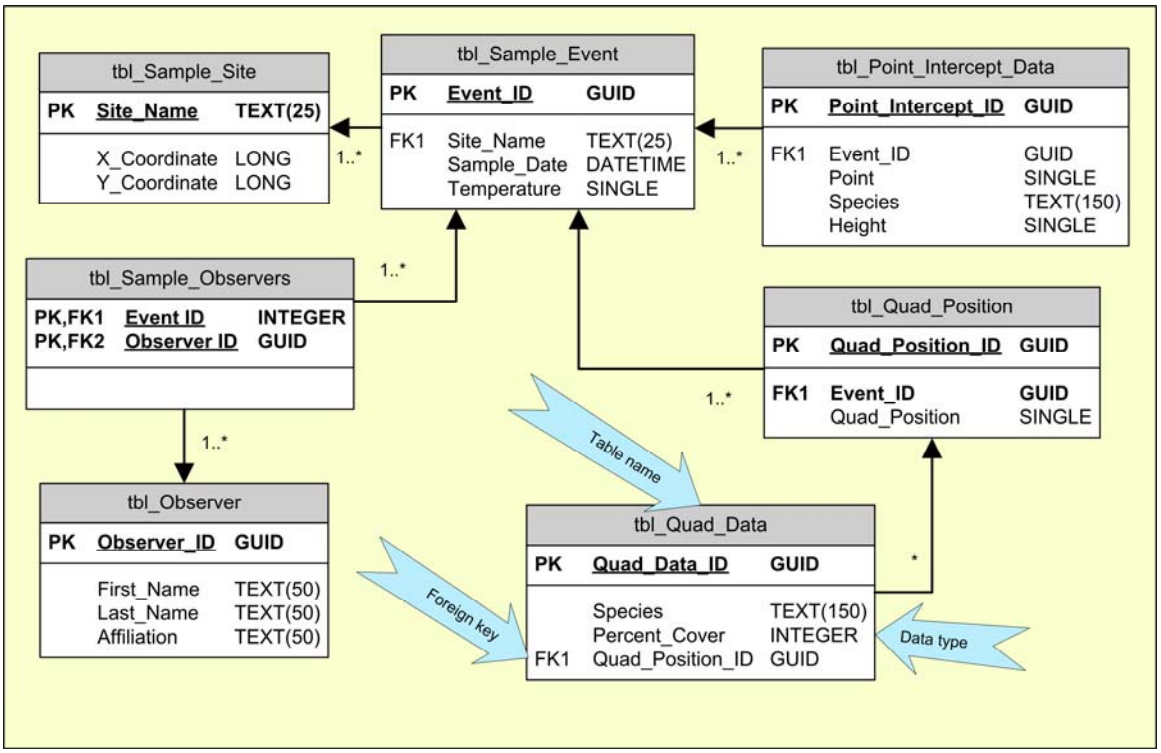


Figure G-3. Example of a physical data model. Primary keys are displayed in bold, are underlined and have the notation “PK” to the left. Foreign keys are denoted by “FK1” or “FK2” (there can be more than one foreign key in a table). To the right of each attribute, its data type is displayed. The relationships and cardinality notations are the same as in Figure G-2.

Appendix H. Standard Data Summary Reports

To be developed…

Appendix I. Exporting Data for Analysis

To be developed…

Appendix J. Metadata Creation for Non-spatial Datasets

In development

Basic workflow

1. Complete all table properties and field descriptions in MS Access.

2. Use Dataset Catalog to complete contact, spatial domain, spatial reference, entity & attribute, and distribution sections.

3. Export incomplete metadata file from Dataset Catalog in XML format.

4. Open XML file in Metavist to complete mandatory FGDC elements.

5. Obtain ITIS taxonomic report.
    - Export list of species names from Access database as comma delimited text file, entire scientific name as one field.
    - Go to http://www.itis.usda.gov/taxmatch_ftp.html.
    - Upload text file to ITIS and click Step 2.
    - Select the kingdom for the species in your list (only 1 kingdom at a time can be analyzed).
    - Select "Scientific Name (FGDC Biological Profile Report - Prototype)."
    - Use the "View Matches" and "View non-matches" to check your upload.
    - Click the Taxonomic Compare button.
    - Report summary page – to get to SGML click the "Generate FGDC Biological Profile" button.
    - Click "Download SGML data."
    - Use "SAVE AS" to save as a text file.

5. Create FGDC Biological Profile using ITIS taxonomic report.
    - Complete the first part of the Biological profile using Metavist.
    - Open the XML metadata file in EditPad Lite.
    - Open the ITIS download in EditPad Lite.
    - Paste the ITIS file into the XML metadata just below </keywtax> and Save.
    - Open the file again in Metavist and note line number of errors; use EditPad Lite to go to line with errors and edit to remove errors (errors are usually associated with common names in Spanish including accent and ~ characters).

6. Validate and correct XML file
    - Go to http://geo-nsdi.er.usgs.gov/validate.php.
    - Upload XML file and click Validate button.
    - On the "Validation and re-expression" page click the "Error report" hyperlink to see all validation errors.
    - Correct errors using Metavist and re-run the validation.
    - When XML file is error-free, save the file in HTML, SGML and text formats.