SOP 5: Data Management

This SOP provides additional details on entering data collected during the implementation of SOPs 1 through 4 and 6. There are two data recording media that will be used during 2017: initially, paper forms with post-recording entry into a digital database or direct (digital) entry into a Geographic Information System [GIS] format through a Personal Accessory Device (PAD or iPad; see SM 4). The GIS data structure uses ArcGIS Online (AGOL), a web-based map viewer that is accessible to public and private entities. It is a good practice to start with recording data on paper data sheets (Table SOP-5.1; SM 5) before attempting direct data entry via a PAD in the field. Recording of data on paper data sheets will allow familiarity with the nature and type of data before needing to learn how the digital Survey123 forms are organized and used to enter data into AGOL. Currently, SOP 4 data will need to be entered into an Excel spreadsheet file available upon request (patrick_ward@fws.gov).

Data collectors will be using the PADs with the Collector application for ArcGIS to navigate to the sampling plots in the field and for entering data into a web-accessible database (SM 4). Without an ArcGIS Online (AGOL) account, data collectors will not be able to use this application to see site locations and plot layouts, enter data into AGOL or navigate to sites without using other techniques (i.e., map and compass; Global Positioning System [GPS]).

Every person that will be collecting or viewing entered data will need to have an AGOL account. It is also important for survey coordinators to have an AGOL account. All the data is stored in the Environmental Systems Research Institute (Esri) Cloud, so account holders will be able to log on and see where the plots are and what has been collected at each refuge (SM 4).

There are three options to get an AGOL account:

FWS staff (@fws.gov email) - Please use the following Google form to apply for an account. https://fishnet.fws.doi.net/projects/gisnew/Lists/Publisher_Collaborator_Pro_Form/Item/newifs.aspx?List=1ba4ff31-665c-48de-bf77-5fcd068cba13&RootFolder=&Web=d3156e04-03c1-48c4-a667-07307f9c1a43
You do not need access to ArcGIS Pro and you are not sponsoring a collaborator.
If you already have an AGOL account with another Organization (e.g., University), that account will work too. *Ensure your AGOL account's visibility is set to Everyone (public) so that your account can be located and added to the AGOL Monarch Group.*
Collaborator - A FWS employee must input a collaborators (Non-FWS Person) information in the above referenced Google form in order to have a collaborator account created; see AGOL account creation option 1. Collaborators will only be able to Join Organizational Groups, edit existing features, and share within the organization.

Once your AGOL account has been created, contact Richard Easterbrook, U.S. Fish & Wildlife Service GIS Specialist, at richard_easterbrook@fws.gov so that you can be given access to the AGOL Monarch Group.

Recording Data on Paper

Recording data with pen and paper (or pencil on write-in-rain sheets) is a time honored method. Advantages include:

- it is economical,
- less maintenance,
- less chance of breakdown, and
- provides a record for checking errors committed during initial data recording.

Disadvantages include:

- extra time required to transfer recorded data into digital form (nearly doubles the time requirements),
- the potential for error during the transfer of data from paper to digital format, and
- the realistic possibility of damage and loss of data on paper media in wet weather.

Use of a PAD will save considerable time once users have learned how to navigate, take pictures, and record data electronically on site. The cost of the PAD is more than paper forms, but easily pays for itself within a season of data collection by saving employee time managing data.

When entering data using either a PAD or a printed data form, keep handy and refer to the metadata-instructions document (SM 5). This will save time correcting errors that are committed early in the monitoring process. Once the appropriate entries for the relevant data form fields are memorized, you may not have to use the metadata–instructions document as often, but it is advised to keep the metadata-instructions document with you when recording data.

All data sheets and associated metadata will be available in electronic form for printing, and in hard copy version (SM 5). There may be updated forms following the first month of sampling. Each SOP, excluding SOP 5, has one or more data sheets and a metadata description (Table SOP-5.1).

To Record Data Accurately—

- Reference the forms in advance and understand what type of data are entered.
- Review the metadata that describes the content for each field, including definition of codes to use and other data entry instructions.
- Ask questions and find answers from the survey coordinator if you do not understand how data are being recorded.
- Write legibly. In particular, look for distinguishing 4's from 9's, 1s from 7's, and 3's from 5's or 8's.
- Be accurate with use of codes, particularly for species names or plot labels. Have your reference sheets available in the field. Create a quick-reference sheet of codes or data field choices and tape in on your clipboard.
- Provide comments for data entries that may be unclear or questionable.

Table SOP-5.1. Directory of data sheets used to record measurements of monarch abundance and conditions of their habitat following 2017 standardized operating procedures. All data sheets and metadata descriptions can be found in SM 5.

SOP No. and Title	Paper Data sheet(s)	Sides	PAD Survey123 Form	Metadata File Label
SOP 1: Site Selection, Establishment and Description	SOP 1 Data Sheet 1 SOP 1 Data Sheet 2 SOP 3 Data Sheet 2[a]	2 1 1	*SOP1 Sampling Site Description;* *SOP3 - Datasheet 2 Miscellaneous Adult Monarch Observations*	**SOP 1**
SOP 2: Counting Adult Monarchs	SOP 2 Data Sheet 1	1	*SOP 2 SOP2 Butterfly Walk;* *SOP3 - Datasheet 2 Miscellaneous Adult Monarch Observations*	**SOP 2**
SOP 3: Counting Plants and Immature Monarchs	SOP 3 Data Sheet 1F SOP 3 Data Sheet 1B[b] SOP 3 Data Sheet 2	1 2 1	*SOP3 - Data Sheet 1 Nectar Plants, Milkweed, Monarch Eggs & Larvae;* *SOP3 - Datasheet 2 Miscellaneous Adult Monarch Observations;*	**SOP 3**
SOP 4: Estimating Monarch Survival	SOP 4 Data Sheet 1	1	**None – use Excel Spreadsheet**	**SOP 4**
SOP 6: Estimating Density and Biomass of Red Imported Fire Ants[c]	SOP 6 Data Sheet 1 SOP 6 Data Sheet 2	2 2	*SOP6 – Red Imported Fire Ants*	**SOP 6**

[a] SOP 3 Data sheet 2 is used all plot visits until front and back is filled; there is no need for a separate form for each plot-visit.

[b] SOP 3 Data sheet 1B should be copied to produce 15 pages, each with this form on the back and front; SOP 3 Data sheet 1F needs to be copied on 1 single sided page (per plot-visit).

[c] SOP 6 is only used in portions of the US where red imported fire ants occur. Use Data Sheet 1 ***or*** 2 depending on density of ant mounds and sequence in which mounds will be measured relative to other collected data.

Check Your Data—At the end of each data recording visit, trade data sheets with your partner(s) and check each other's data sheets for the following:

- all data required for a visit has a completed data sheet, especially for SOP 3 which requires one entry per subplot (or between subplot designation), which will require multiple forms; please fill out header data for all forms.
- errors in values (e.g., 56 stems of milkweed in a subplot instead of 5 or 6);
- that the plot label, visit number and monitoring area are recorded and the date is correct;
- legibility of the plant name codes;
- features of the identity of unknown plants are recorded in comments; see Appendix C for suggestions on a system for recording plant acronyms in species are not identified.

Entering Data from Paper to a Database

Data collected on paper forms will need to be manually entered into Survey123 Forms using either the Survey123 App on an iPad or the Survey123 for ArcGIS Program on a PC. If data will ultimately be archive in AGOL, using spreadsheets for digital entry of data is ***strongly discouraged*** because it cannot be imported into AGOL. Use spreadsheets for data entry only if you have no access or intent to archive data in AGOL. Data entry spreadsheets need to have the same structure (data fields and definitions) as the final "flat" files exported for AGOL

For all data entry, ***please refer to the metadata file that posts instructions and codes***, located in the same data entry spreadsheet for each monitoring area or in hardcopy form in SM 5. Check with your survey coordinator for where this file is stored if you do not know. In the metadata file, you will find that the name of a Monitoring Area is the abbreviation for the general area that you are sampling (e.g., in 2017, Balcones Canyonlands NWR – BCL; Necedah NWR – NCD; Lamoni, IA - LAM). The Monitoring Area codes are taken from the FWS Reality Database designations.

Detailed instructions on how to enter data into Survey123 and submit this data can be found in SM 4. Once your data is entered AND "submitted" via Survey123 it will be available for viewing in AGOL. These data will be backed up by NRPC staff each evening. Find out from your Survey Coordinator where paper data forms and spreadsheet entered data files should be saved and stored in the office. ***Back up spreadsheet-based data files on a flash drive after new data are added.***

Using AGOL and the Collector Application

There are 3 previously mentioned geospatial tools that are used for this survey project. These tools are described in detail in SM 4. They are:

1. ArcGIS Online (AGOL) – Cloud based geospatial data storage, viewing, and Quality Assessment (QA)
2. Collector for ArcGIS – Application used as a navigation tool on an iPad or other device for field workers to locate their survey locations. Can be used to edit survey locations as well as to add plots, transects, and subplots.
3. Survey123 for ArcGIS – Application used for digital data entry. Digital forms have been created that replicate paper forms. Entering data into Survey123 on a PAD or PC will be required for final data submission.

Quality Assessment (QA)

Survey coordinators need to check entered data early on in the monitoring trial and provide feedback to their field technicians on inaccurate entries. By doing a thorough review of the data being submitted by field technicians early in the monitoring season, data entry mistakes will be addressed and corrected before they accumulate in the entire data set. Thereafter, check the data regularly (e.g., every 3 weeks) to check for errors and progress. ***It is imperative that data recorded from paper forms are entered in a timely fashion*** because monitoring results will need to be summarized in mid-September and again in October, leaving little time for post-season checking of the entire data set. Also, check to see that data files are being saved and named by date and are being backed up and saved in the proper designated location.

QA can be performed either by viewing the data entries in Survey123 or in AGOL, but it is much easier for data collectors to review the data in exported flat files using a spreadsheet. Assessing the quality of data submitted to AGOL requires two general tasks:

- Exporting and posting of a flattened file that can be accessed and viewed by data collectors.
- Reviewing the flat files for errors and communicating/recording the changes that need to be made to the AGOL data base.

Creating the Flat File— For FWS AGOL archived data, export a flat file using the web service-based program developed by NRPC and posted at: :\Data\Projects\Monarchs\2017\SamplePoints\

1. Download to your hardrive (runs better) and execute the file 04_Monarch_2017_Data.exe.
    a. 04_Monarch_2017_Data.exe is a datasheet generator built specifically for harvesting and compiling all SOP1, SOP2, SOP3 Datasheet 1, SOP3 Datasheet 2, and SOP6 data that is stored in the AGOL Monarch Group through the submission of Survey123 forms.
    b. The datasheet generator will produce a separate Excel Comma-Separated Values (CSV) file for each survey, and automatically download the CSV files to your desktop. The difference between the CSV files of the data you are able to download directly from the AGOL Monarch Group and the CSV files produced by the datasheet generator is that the CSV files produced from the datasheet generator will have all data from all related tables for a survey compiled onto a single tab of the spreadsheet,
    c. While the CSV file downloaded directly from AGOL will have the survey data separated onto multiple tabs of the spreadsheet.
2. Rename the CSV files that are automatically downloaded to your desktop by the datasheet generator, and then store them in a secure location. It is suggested that you include the date on which the datasheets were created in the file name.

Reviewing and Editing Data—When reviewing and editing data collected by many different hands it will be efficient to devise a system for the survey coordinator to set up a data management schedule and system for communicating potential errors with the data collectors, and for understanding what and who will make the changes in the data AGOL data base.

In 2017, NWRS followed the process described below. NRPC staff acted as the Survey Coordinator and AGOL Monarch Group Data Manager for reviewing submitted data and for flagging of potential errors. Data collectors reviewed the flagged files and communicated

through a specific marking system the changes that would be needed and who would make those changes to AGOL. Beginning in 2018, Regional I&M staff will need to serve the role of reviewing and flagging potential data errors and editing recommended changes in AGOL.

A critical first step in the review process is for the data collector that submitted data via the cloud to AGOL to confirm that those data made it into the AGOL database. See SM 4 for viewing submitted data in AGOL.

The data review process can be broken down into four broad parts: 1) data submitted by data collectors (the field), 2) exported data reviewed and errors flagged by a Survey Coordinator or Data Manager, 3) flagged errors are reviewed by the field, and edited by the designated editor (in this case NRPC staff). All of these actions are recorded using a designated marking system, and this data edit record will be archived with the final data files (see ***Data Security and Archiving***).

Tracking and Scheduling Data Review—Through whatever method the survey coordinator finds most efficient, keep track of the surveys that have been submitted by the field crew through Survey123. If your organization's IT security allows, create an online spreadsheet that can be shared with and edited by multiple individuals (e.g., Google Sheets). This spreadsheet should include the plot labels of the plots the field crew is monitoring, a schedule of planned/completed monitoring visits, and space to allow for the tracking of the status of surveys for various plot visits. The status of the surveys is to be updated in the spreadsheet by the field technicians and the survey coordinator as the data collection, review, and editing process progresses throughout the season.

In 2017, a spreadsheet was shared for each SOP data collected during a specified period, generally a round of visits. A review status for each data file was used to track and name the files in the shared schedule:

1. *Needed* - Indicated that data from a period of time needed to be submitted. (The survey may have been completed by the field technicians, but until the data were submitted to AGOL through Survey123, the data were still considered as 'Needed')
2. *Submitted by the Field*- Indicated that the data were submitted through Survey123 to AGOL by the field technicians.
3. *Reviewed by the Office* - Indicated that the data were reviewed by the Survey Coordinator or Data Manager, and were provided/made available to the field technicians for their review and editing.
4. *Reviewed by the Field* - Indicated that the data were reviewed by the field technicians, and any corrections identified and requested of an AGOL Monarch Group data manager.
5. *Edited by the Office* - Indicated that the AGOL Monarch Group data manager edited the identified changes to the AGOL database.

Creating the data review and management schedule in accordance to a site visit and data collection schedule facilitated a way to validate plot labels and identify any missing data, particularly cases where submission to the AGOL cloud did not result in successful submission to the database.

In 2017, the data files provided to the field technicians for review were posted on the Google Drive as spreadsheets with three different worksheets (labeled tabs):

- Instructions and Color Key. This tab contained instructions for the data review and marking process, including definitions of cell and text colors for identifying and communicating required edits.
- Survey Data for Review. This tab was for a worksheet that contained all data that was reviewed by the survey coordinator for a given collection period. This worksheet was used to review, flag cells and missing or questionable values and then give the correct values and communicate who will make the needed changes in AGOL. Three new columns (A-C) were added to this worksheet to indicate:
    - i. the cells or records with errors and who noticed the errors,
    - ii. needed edits and who recommended those edits, and
    - iii. the date and by whom the recommended edits were made in the AGOL database (Figure SOP-5.1).
- Raw Survey Data. This worksheet contained all of the same survey data as the 'Survey Data for Review Tab,' but with no color coding or added columns. This "clean" set of survey data was included so that original data were available to the spreadsheet reviewer or data editor.

When the Survey Coordinator (office staff) reviewed submitted data in the spreadsheet, cells with potentially, incorrect or missing values were flagged for subsequent review by the data collectors. Some of the commonly occurring errors to search for are shown in Table SOP-5.2.

Data Reviewer Comment (With Commentor Initials)	Data Editor Comment/Change Made (With Commentor Initials)	Date Change Made	Person Who Made Change to Data in AGOL	Date Data Change Made in AGOL
Data needs reviewing. Survey Start time later than survey end time. (KEH 6/30/2017)	Changed end time. (EPT)	7/20/2017	Keith Hamilton	8/4/2017
Data needs reviewing. Survey Start time later than survey end time. (KEH 6/30/2017)	Changed end time. (EPT)	7/20/2017	Keith Hamilton	8/4/2017
Data needs reviewing. Survey Start time later than survey end time. (KEH 6/30/2017)	Changed end time. (EPT)	7/20/2017	Keith Hamilton	8/4/2017
Data needs reviewing. Survey Start time later than survey end time. (KEH 6/30/2017)	Changed end time. (EPT)	7/20/2017	Keith Hamilton	8/4/2017

Figure SOP 5-1. Example of the spreadsheet documenting the changes made in AGOL for SOP 3 data file following review by NRPC staff (serving as Data Manager) and data collectors in Region 2.

Table SOP-5.2. Commonly encountered errors in files with data collected to monitoring monarchs or plants during reviews of SOP specific data files.

Potential Error	SOP Data File	Comments
Incorrect Plot Label	1, 2, 3	Leads to appearance of missing records
Missing plot visit records	1, 2, 3	Based on mismatch between visits on work schedule and submitted data
Number of Subplots < minimum of 150	3.1	
Missing Verified Plot Type	1	Needed when GRTS draw points are misclassified
Missing nectar plant name when frame size is present	3.1	Use of 'Unknown' identifier' should be used
Missing records for butterfly walk segments	2	Every visit needs 10 records (1 record per segment)
Extra (erroneous) records with 0 values	3.1	Caused by misnavigating or starting in a Survey123 form

By sharing the data to be reviewed with field technicians through an online spreadsheet, multiple individuals will be able to review and edit the data through a single spreadsheet that is centrally located, and the survey coordinator will have immediate access to the edited data. The downside of posting the data on an online spreadsheet for the field technicians' to review is that they must have consistently reliable internet to access and edit the data.

It is advised that when providing the reviewed data to the field technicians for their comments and possible edits, it is best to give them a deadline when the data should be reviewed by them and returned to the survey coordinator or the AGOL Monarch Group data manager by. If given an at-your-leisure deadline to review the, there is a likelihood that the data will not be reviewed in a timely fashion.

Once the data review is completed by all of the field technicians that need to review the data, the spreadsheet of the data that the field technicians reviewed is to ultimately be sent to the AGOL Monarch Group data manager so that the data can be permanently edited in AGOL. The AGOL Monarch Group data manager should also change the Excel cell color to indicate that data they have edited in AGOL. Doing this will aid the AGOL Monarch Group data manager in tracking what data they have edited in AGOL, and will also indicate to others what data was changed in AGOL.

Starting with version 2.4, previously submitted Survey123 forms that are still on the iPAD can be edited and resubmitted to make changes in AGOL. Survey coordinators and/or Data Manager for a Region will need to determine the roles and work in changing AGOL database. In 2017, changes to the AGOL Monarch Monitoring data will be done by NRPC staff and documented in a dedicated spreadsheet record, rather than by serial submission of edited Survey123 forms. The former process does not maintain a convenient record for the edits to the AGOL data base.

Updating Nectar Plant Names—As nectar plants are identified in each FWS Region, the list (SM 3) and associated 6-character codes will need to be updated. Communication between the field crew and Survey Coordinator/Data Manager will be required for updating the choices of nectar plants in drop-down menus in Survey123 or SM 3 Tables on a regular basis (e.g., weekly). Posting the initial list of plant species name for a region as a Google Doc shared with the data collectors is one technique for facilitating rapid update. Plant scientific names need to follow those in the North American Flora series. The 6-character codes, initially comprised of first three characters of the Genus plus the first three characters of the specific epithet, should be checked with all regional plant lists to insure the code is unique.

If a plant species name already on the SM 3 lists has the same code as the newer identified plant, then the code of the newer plant name will need to be altered to something unique. This is usually done by changing the 5th or 6th character to another letter found in the specific epithet. ***Altered plant name codes need to be recorded in updated, online SM 3 tables for rapid discovery by other Survey Coordinators or Data Manager***. All codes need to be given with a scientific name and common name in the SM 3 tables and derived plant lists. Updated nectar plant lists should also be added to SOP2 and SOP3 spreadsheets being reviewed by data collectors for reducing errors in entering newly identified plants to the data files.

Data Security and Archiving

Securing and archiving final checked data files will needed to follow data management guidelines of each participating organization or agency. For example, FWS in 2017, backed up monitoring data submitted to AGOL through Suyrvey123 ***nightly***. Flat files were exported ***weekly*** by NRPC staff and stored on the T: drive of the local server at IM > Working > Monarch Butterfly > SurveyDataBackups > 2017 for the QA process. In addition, the monitoring activities and products (data and reports) will be submitted respectively to the Planning and Review of I&M activities on Refuges (PRIMR) and Service Catalog (ServCat) database.

PRIMR—The National Wildlife Refuge Systems data base of inventory and monitoring activities will need a record for this monitoring activity. When this survey is conducted on or near a station of the NWRS, a survey record should be entered into PRIMR. The survey name in PRIMR should be labeled ***MCSP Monarch and Plant Survey***. When this survey was conducted in the same area during two more years it should be labeled as ***cooperative baseline monitoring***; when never repeated in the same area, label as ***cooperative inventory***. During 2016 and 2017, NRPC recorded the appropriate information for each participating station in PRIMR and identify/link the record to the appropriate survey guidance and product documents. In 2018, Regional I&M staff should check PRIMR and determine if and by whom a new survey record is needed. Staff of a participating station that sponsored a MCSP monarch and plant survey may choose to check and update the years in which the monitoring was conducted before producing a PRIMR reports. The survey record will apply to activities conducted on and adjacent to the sponsoring station for data that will analyzed to make ***National***-scale inferences. For example, see record for Neal Smith NWR survey FF03RNLS00-040.

ServCat—The final version of checked files will be archived in ServCat as a comma delimited (.csv) flat file under the MCSP Monarch and Plant Monitoring Project (66678). Each data file should be named for the SOP and year that the data represent. Final data files will be available for use as determined and disseminated by the MCSP monitoring team and Inventory and

Monitoring staff working on the monarch monitoring trial. Any paper data forms that were completed should be copied or scanned, and mailed to the NRPC office, care of Pat Ward. Originals of the data forms should reside with the survey coordinator of the data collectors, along with copies of any relevant notes generated by the data collectors.